Linguistik Aktuell Linguistics Today

The Derivational Residue in Phonological Optimality Theory

Edited by
Ben Hermans
Marc van Oostendorp

THE DERIVATIONAL RESIDUE IN PHONOLOGICAL OPTIMALITY THEORY

LINGUISTIK AKTUELL/LINGUISTICS TODAY

Linguistik Aktuell/Linguistics Today (LA) provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.

Volume 28

Ben Hermans and Marc van Oostendorp (eds.)

The Derivational Residue in Phonological Optimality Theory

THE DERIVATIONAL RESIDUE IN PHONOLOGICAL OPTIMALITY THEORY

BEN HERMANS
Tilburg University

MARC VAN OOSTENDORP
Meertens Institute,
Royal Netherlands Academy of Arts and Sciences

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM / PHILADELPHIA

∞ TM The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

The derivational residue in phonological optimality theory / [edited by] Ben Hermans, Marc van Oostendorp.
p. cm. -- (Linguistik aktuell / Linguistics today, ISSN 0166-0829; v. 28)
Includes bibliographical references and indexes.
1. Grammar, Comparative and general--Phonology. 2. Optimality theory (Linguistics) 3. Grammar, Comparative and general--Morphology. I. Hermans, Ben. II. Oostendorp, Marc van, 1967- III. Series: Linguistik aktuell ; Bd. 28.
P217.3.D47 1999
414--dc21 99-39779
ISBN 90 272 2749 7 (EUR) / 1 55619 912 0 (US) (Hb; alk. paper) CIP

John Benjamins Publishing Co. · P.O.Box 75577 · 1070 AN Amsterdam · The Netherlands
John Benjamins North America · P.O.Box 27519 · Philadelphia PA 19118-0519 · USA

Table of Contents

List of Contributors

John Alderete
Department of Linguistics
Faculty of Arts
Buchanan E270–1866 Main Mall
Vancouver, B.C. Canada V6T 1Z1
alderete@interchange.ubc.ca

Mary Bradshaw
Linguistics Department
Ohio State University
222 Oxley Hall
1712 Neil Avenue
Columbus, OH 43210–1298
USA
bradshaw@ling.ohio-state.edu

Gene Buckley
Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104–6305
USA
gene@unagi.cis.upenn.edu

Matthew Y. Chen
Professor of Linguistics
Univ. of California, San Diego
and
Dean, Faculty of Humanities and Social Sciences
City University of Hong Kong
Tat Chee Ave.
Kowloon, Hong Kong
People's Republic of China
fhmychen@cityu.edu.hk

San Duanmu
Program in Linguistics
University of Michigan
1076 Frieze Building
Ann Arbor, MI 48109–1285
USA
duanmu@umich.edu

Chris Golston
Department of Linguistics, PB 92
California State University Fresno
Fresno CA 93740
USA
chrisg@csufresno.edu

Bruce P. Hayes
Department of Linguistics
University of California, Los Angeles
Los Angeles CA 90095–1543
USA
bhayes@humnet.ucla.edu

Harry van der Hulst
Department of Linguistics
Leiden University
PO Box 9515
2300 RA Leiden
The Netherlands
hulst@rullet.leidenuniv.nl

Ben Hermans
Department of Linguistics
Tilburg University
POBox 90153
5000 LE Tilburg
The Netherlands
b.j.h.hermans@kub.nl

René Kager
Utrecht Institute of
Linguistics/OTS
Trans 10
3512 JK Utrecht
The Netherlands
kager@let.uu.nl

Marc van Oostendorp
Meertens Institute
Joan Muyskenweg 25
1096 CJ Amsterdam
The Netherlands
marc.van.oostendorp@meertens.knaw.nl

Orhan Orgun
Department of Linguistics
University of California, Davis
Davis, CA 95616
USA
ocorgun@ucdavis.edu

Rosalind Ruth Roberts-Kohno
Linguistics Department
Ohio State University
222 Oxley Hall
1712 Neil Avenue
Columbus, OH 43210–1298
USA
rrrkohno@ling.ohio-state.edu

Introduction

Optimality Theory and Derivational Effects

Marc van Oostendorp Ben Hermans

1. Optimality Theory and Phonological Derivations

The articles collected in this volume provide an overview of the status of derivational theory within one of the most popular frameworks in present-day phonology, Optimality Theory. According to Anderson (1985), the history of phonology in the twentieth century can be seen as a sequence of periods in which the emphasis is on the structure of phonological representations, alternating with periods in which the emphasis is on phonological derivations. In periods in which representations are the focus of interest, most scholars are concerned with the internal structure of units such as segments, words and phonological phrases. In periods in which derivations are more central, people study the way in which words are phonologically related to one another. According to Anderson, taking an interest in representation is often connected to the study of *languages*; taking an interest in derivations is connected to the study of *grammars*. Of course, this does not mean that either of these topics have been completely ignored in any period of time. It is impossible to purely concentrate on one of these aspects:

> In fact, theories of rules and theories of representations deal with intimately interrelated and indisoluble aspects of the same linguistic structure. In order to understand that structure, however, both aspects must be appreciated, and this has certainly not always been the basis on which inquiry into sound structure has proceeded. (Anderson 1985:9–10)

One may wonder how the 1990s are to be evaluated in this light. Whether we like it or not, this decade will undoubtedly be seen in future historical overviews as the era of the rise of Optimality Theory (OT). It is not easy, however, to categorise this theory within Anderson's classification. The present volume tries to establish what is the derivational theory that we need under the assumptions of OT (or more generally, of constraint-based phonology). It contains contributions from scholars in many different traditions, arguing for many different positions, but all sharing a basic interest in the desired form of the theory of derivations.[1] In this way, we hope to demonstrate the complexity of the issues involved.

In this introduction, we summarise the main points of the contributions to this volume in the light of the general discussions of the past few years. In this first section, we characterise the basic notions 'derivational', 'representational', 'input' and 'output' and summarise what 'classical' rule-based theory and Optimality Theory have to say about them. In the remainder of this introduction, we then proceed from the smaller domains to the larger domains of derivationality: Section 2 discusses rule ordering and the directionality of rule application and Section 3 concentrates on issues concerning the cycle. The largest possible type of derivationalism — the distinction between lexical and postlexical phonology — is not explicitly discussed in this volume;[2] we refer to Booij (1997) and Kiparsky (1998) for discussion. Section 4 will be devoted to a conclusion.

1.1 *Optimality Theory, Representations and Derivations*

Optimality Theory arguably is not a theory of representations in the same sense in which autosegmental, metrical and prosodic phonology were theories of representations. For instance, Prince en Smolensky (1993) use a representation of the syllable which divides this constituent into onsets, nuclei and coda's, but in a footnote they note that other representational choices are not incompatible with OT, and indeed many papers in Optimality Theory divide the syllable into moras. More generally, scholars have put to use all kinds of representational assumptions in combination with Optimality Theory. Some researchers seem to even have abandoned the notion of an abstract phonological representation altogether (see Gafos 1996 and Boersma 1998, among many others)

On the other hand, one could also wonder whether Optimality Theory

1. A volume of articles with partly the same concerns is Roca (1997).

2. The device of constraint domains used by Buckley (this volume) could be naturally extended in such a way as to mimic certain aspects of Lexical Phonology.

provides us with a theory of derivations. It certainly is not a ‘theory of rules’ in the technical sense that is usually given to it, because it denies the existence of phonological rules altogether. With this, many of the derivational tools of *SPE*-type phonology such as rule ordering and the cycle are abandoned by a lot of scholars working within the OT paradigm. For most of them, a minimal derivational residue persists however: there is a function mapping inputs to outputs.

‘Classical OT’, i.e. the model that Prince and Smolensky (1993) proposed, consists of a function Gen (Generator) mapping a given input onto an infinite set of output representations called candidates and a function H-Eval (Evaluator of relative Harmony) mapping the set of candidates onto a unique representation, the phonetically realized output. This process can be formally represented in the following way (after Prince and Smolensky 1993: 4):

(1) a. Gen (In_k) $\rightarrow$ $Out_{inter} = \{Out_1, Out_2, \ldots.\}$
b. H-Eval (Out_{inter}) $\rightarrow$ Out_{real}

In this model the function H-Eval is of central importance. It determines which candidate of the infinite set of possible candidates (generated by Gen) will be phonetically realised. In order to do this, it uses a universal set of constraints, Con. These constraints may conflict, in the sense that satisfaction of a given constraint C_1 necessarily leads to a violation of a constraint C_2. The conflict is resolved in such a situation in favour of the constraint that is highest ranked: H-Eval consists of a set of violable, ranked constraints. The only representations that matter are In_k and Out_{real}: the intermediate candidates in Out_{inter} do not play a role in the formulation of constraints,[3] and as a matter of fact also the role of In_k is fairly limited: in the view of most students of OT, Con contains constraints on the desired internal structure of Out_{real} and constraints on the relation between Out_{real} and In_k, but there are no constraints on the internal structure of In_k.

1.2 *The Theory of Inputs*

The assumption that there are no constraints in the grammar that refer exclusively to inputs is usually referred to as ‘Richness of the Base’ (and sometimes as ‘Freedom of the Input’). This assumption of course does not imply that anything in the world can be input to the grammar: we need some basic assumptions

3. The only variant of OT where these representations *can* play an active role in the formulation of constraints is McCarthy’s (1998) Sympathy Theory (see below).

about the syntax of our phonological representations. For instance, we need to know whether vowel length is represented in terms of mora's, underlying x-slots or in some other way and it does not make sense to assume that x-slots can be part of underlying representation if we assume at the same time that they are represented with moras at the surface.

Although constraint-based theories such as Optimality Theory themselves are neutral with respect to the particular representational assumptions we make, these assumptions certainly have to be made in a complete linguistic theory. These decisions then will also have their repercussions on any given analysis within OT. Let us consider as an example the issue of underspecification, one of the most hotly debated topics within the generative phonology of the 1980s and early 1990s: should all redundant information be underspecified? Or should we only leave out for instance the non-redundant information that is not distinctive?[4]

A basic idea underlying this discussion was that underlying representations in linguistic theory should contain as little redundancy as possible. Under the assumption of Richness of the Base, these questions have become almost completely irrelevant: anything can be input to the grammar, including completely specified phonological representations. As a matter of fact, we could wonder whether we should not ask the opposite question to those just mentioned: are linguistic representations ever underspecified at all?[5] As we have just seen, even under Richness of the Base we should still evaluate what the linguistic objects are that the Generator function generates and the Evaluator evaluates. It is very well possible that for instance representations that are underspecified in certain ways are not valid linguistic objects at all, just like representations involving x-slots, or representations in which a syllable node dominates a foot node, are not valid linguistic objects.

Some questions relating this intriguing topic are discussed in the paper by Golston and Van der Hulst in this volume. In particular, these authors study whether syllable structure is underlyingly present, or the result of a syllabification process of strings of underlying segments. Their answer is that strings of segments are *not* valid phonological representations. Phonological processes only operate on syllable structure 'mobiles' to which phonological features are directly attached.

---

4. Cf. Steriade (1995) for an overview.

5. A similar question arises in other output-oriented theories of phonology; it receives a negative answer in Government Phonology (Kaye, Lowenstamm and Vergnaud 1985, 1990), and a positive answer in unification-based approaches (Bird 1995).

Golston and Van der Hulst thus argue against what they call the principle of IMPOSSIBILITY:

(2) IMPOSSIBILITY
Every underlying form is an impossible surface form and vice versa.

IMPOSSIBILITY is the principle underlying most underspecification theories — underlying representations are fully underspecified, surface representations have to be fully specified — and also much work on syllable structure and prosodic structure within generative phonology — in which underlying representations do not have any prosody, whereas all segments are part of a prosodic structure on the surface.

According to Golston and Van der Hulst, we do not need to make this assumption. They show that, if we have underlying syllable structure in the first place, it is no longer necessary to assume that there is anything like a segment at all. If we know that a syllable onset contains the material for a stop and a liquid, we know what the order of those segments is. Either syllable structure or the segment is superfluous, and Golston and Van der Hulst provide linguistic, paralinguistic and extralinguistic evidence that this means that the segment has no role to play. The same applies for stricture (manner) features as well: these can be read off from the structure of the syllable 'mobile' — hence the slogan 'stricture is structure'.

Golston and Van der Hulst are the only contributors to this volume who do not take an explicit standpoint for or against Optimality Theory. Indeed, their view that 'stricture is structure' is compatible with many views on phonological derivation. For instance, it could be combined with an *SPE*-type mechanism of extrinsically ordered rewrite rules. What their paper makes abundantly clear, however, is that even in a theory without independent constraints on input structure, the issue of what the input can be is still an interesting one.[6]

---

6. It should be noted that in their actual implementation of the idea that 'stricture is structure', Golston and Van der Hulst take into account the possibility that some mild form of IMPOSSIBILITY still holds true. They argue that the order of segments in a syllable is predictable and therefore needs not be specified underlyingly. Yet there are reasons to assume that linear ordering does play a role on the surface. In this sense then, there is still a derivational residue of syllabification even in the proposals of Golston and Van der Hulst (as pointed out to us by Harry van der Hulst, p.c.).

1.3 *Levels of Derivationality*

Some researchers, however, have abandoned the assumption of abstract inputs altogether. They base themselves on a lexicon of output forms, and relations among those output forms, only. In this case there also is no 'theory of derivations' anymore in the technical sense of generative phonology. Yet even in this case there is a strong need for a theory of relation between forms. As a matter of fact, many researchers use Optimality Theory in order to describe possible alternations between forms. In other words, OT is used more as a theory of grammars than as a theory of languages. In Anderson's terminology it should therefore probably be described as derivational.

As a theory of derivations, the 'classical' type of OT, as presented in Prince and Smolensky (1993), is fairly minimalist. It basically recognises only one type of derivational relation: the one between input and output. Within rule-based theory, this type of relation was established by the single rule or by an unordered block of rules. On the other hand, several additional levels of derivationality were recognised in the 1980s:

- Certain types of rules were supposed to apply directionally: on a given string for instance, application of stress feet could apply from left to right or from right to left;
- Rules were ordered with respect to one another: a form F would be undergoing rule A before it would be subject to rule B;
- Blocks of ordered rules were (in some cases) organised in cycles: such a block of rules would first apply to a small domain, and only subsequently to a larger domain;
- Blocks of rules were organised into lexical levels: all rules belonging to 'Level I' morphology would apply before rules belonging to 'Level II' morphology;
- All 'lexical' rules of phonology were supposed to apply before all 'postlexical' rules.

In each case, empirical evidence has been adduced in favour of these extra derivational devices. It seems fair to say that this evidence was accepted by most phonologists as convincing before the rise of Optimality Theory. Furthermore, cyclicity, lexical levels and the lexical-postlexical distinction are not inherently incompatible with OT in principle. The question therefore arises whether we can and should do away with all of these devices. This is primarily an empirical issue, but conceptual questions may also arise: how many theoretical add-ons do we allow ourselves in order to preserve our minimalist view of derivations? How natural are these extra devices? And are they not in the end mere notational variants of the original derivational ideas? The articles in this volume are intended to explore some of these questions, both from the empirical and from the theoretical point of view.

2. Inputs, Outputs and Directionality of Rule Application

A central claim of OT is that the significant regularities of natural language can be found in the output, not in the input (Prince and Smolensky 1993: 1). A particularly convincing argument for this position comes from syllabification in Berber. Prince and Smolensky show that, if we look at Berber syllabification from the point of view of the output, a clear relation between the degree of sonority of a given segment and the position in the syllable of that segment becomes obvious in this language. In particular, the more sonorous a segment is, the more strongly it prefers a syllable's peak position and avoids a syllable's margin. Viewed in this way, Berber is not special, for the same can be said about the syllabication of any language. What is special in Berber is just the fact that the number of discrete points on the sonority scale that are relevant for syllabification, are more numerous than in most other languages.

Looking at the output thus makes it easier to grasp the central generalization behind Berber syllabification, as well as to clarify the relation between the Berber type of syllabification and other types. The same generalizations cannot be made if we look at Berber syllabification from the point of view of the input. In that case it is a mere coincidence that the syllabification algorithm applies in such a way that the output configurations are structured as described above. In the words of Prince of Smolensky, an approach based on rules operating on an input 'suffers from the formal arbitrariness characteristic of re-writing rules when they are put to the task of dealing with … principles of output shape' (ib.:14).

Having reached this conclusion Prince and Smolensky go on to formulate Optimality Theory, specifically designed to capture the central insight that it is the output which matters, not the input. As we outlined above, OT deals with constraints that characterise output configurations rather than with procedures for getting these configurations. This again means that there are no process-specific repair mechanisms. Exactly how a given representation is repaired is decided by the constraint hierarchy of the language, nothing else (cf. McCarthy 1997; and in particular McCarthy 1996b).

In the previous section, we have shown that there are two criteria according to which we can evaluate a theory as being derivational or non-derivational.

According to one, very general, criterion a theory is derivational to the extent that it relates one representation to some other representation. In this sense, the OT-model described above is derivational, be it only minimally so, because it postulates that there is a relation between input and output. The theoretical importance of this criterion is that it allows the model to formulate principles which account for the fact that a segment's 'ontological status' can be decisive for its fate or behavior. In the 'classical' model of OT of Prince and Smolensky (1993) and McCarthy and Prince (1993), there is at least one such principle, viz. *Containment*, which forbids the deletion of any information that is present in the input (McCarthy and Prince (1993: 20). In the Correspondence model of OT, developed in Prince and McCarthy (1995), a segment's ontological status becomes relevant through the Faithfulness Constraints (furthermore, Correspondence Theory, which will be briefly outlined in Section 3 of this introduction, has been used in various ways to extend the number of representations involved in the evaluation of a given linguistic form).

There is a second, more common, way to determine to what extent a theory can be derivational in nature. This involves the criterion of serialism, or serial application of processes. The serialist position is that all representations which are relevant in the grammatical evaluation of a given form can be linearly ordered. The stages in an SPE-derivation satisfy this criterion and probably for this reason the term 'derivational' is often used as a synonym of 'serialist'. According to this definition, then, most popular approaches to OT are non-derivational. A consequence of this is that they do not recognise the existence of process-specific repair mechanisms: it is impossible to say that some process happens before or after some other process. It therefore does not have the possibility of crucially ordering one mechanism before or after the other. For this reason some people have tried to find alternatives in OT for analyses that appear to provide evidence for ordering. The basic technique that is applied is to interpret the second, serial type of derivationality in terms of the first type: the number of representations becomes larger than just 2 (input and output).

Interesting illustrations of this general approach can be found in the work of John McCarthy. An example is McCarthy (1996b), who argues that many derivational devices in Prosodic Morphology (such as positive prosodic circumscription) can be eliminated from the theory altogether, because their effects can be obtained from two independently needed faithfulness constraints. These are OO-faithfulness (cf. Section 3), which establishes relations between separate output representations, and prosodic faithfulness, which establishes relations between the prosodic structure of separate representations.

Another attempt to achieve similar effects is Sympathy Theory of McCarthy

(1998). Here, McCarthy argues in favour of a new type of faithfulness constraint, faithfulness to the sympathetic candidate. This type of faithfulness establishes a relation between an essentially arbitrary, possibly non-optimal, candidate and any other candidate. In this way, the number of representations involved in the evaluation of a linguistic form grows considerably. In terms of empirical predictions this Sympathy Theory differs just marginally from a theory that allows rule ordering, in the sense that almost anything that can be done in one theory can be done with equal ease in the other theory. To the extent that fundamental differences can be discovered (McCarthy 1998 mentions the so-called Duke-of-York Gambit as an instance), they seem to be enforced by a stipulation (viz. that only faithfulness constraints can select a 'sympathetic' representation). Furthermore, there is quite some overlap with OO-faithfulness, the more limited device to extend the number of relevant representations. If classical OT has a minimalist view on derivations, Sympathy Theory seems rather close to a maximalist assumption. It is an empirical question whether all of the predictive power of Sympathy is needed in actual practice.

2.1 *Serialism vs. Parallelism*

Although it is true that most approaches to OT are essentially non-derivational, because they do not allow the ordering of repair mechanisms, this does not necessarily imply that the OT framework is inherently non-derivational, not even according to the criterion of serialism. This point has been made already by Prince and Smolensky (1993: 4–5) and McCarthy and Prince (1993: 24). These authors explicitly recognise the possibility that every single modification of a representation carried out by Gen is followed by application of H-Eval, which is then again followed by a single modification of the representation carried out by Gen, which is then followed by application of H-Eval, etc. According to this view there is a loop between Gen and H-Eval which iterates until there is no possible operation within Gen that can increase harmony.

In practice this possibility has never been applied, as far as we know. It is standardly assumed that Gen freely performs all possible operations in one step, so that it produces all possible candidates. These are then evaluated in parallel by H-Eval. The parallelist hypothesis has been extremely successful. It has played a fundamental role in explaining top-down effects in stress related phenomena (cf. Prince and Smolensky 1993: 28–29), effects of overapplication and underapplication in reduplicative morphology (cf. in particular McCarthy and Prince 1995), and it has been the cornerstone in McCarthy's (1996b) attempt to eliminate positive prosodic circumscription, to mention just a few disparate areas.

OT is also not inherently antiderivational according to the criterion of serialism. As has been mentioned explicitly in McCarthy and Prince (1993: 24) the loop between Gen and H-Eval can be triggered by the fact that the grammar of a given language consists of serially ordered subcomponents. In such a model Gen derives all the possible candidates at level n. Before level n+1 is entered H-Eval selects the optimal candidate of level n. This candidate becomes the underlying representation of level n+1. In this way it becomes possible to combine the parallelist view with a moderate type of serialism: we incorporate one of the core ideas of Lexical Phonology into the theory.

Having made explicit to what extent OT is a derivational theory according to the serialist criterion we are ready to take a closer look at the contributions to this volume that specifically discuss ordering phenomena. All of them assume that 'classical' Optimality Theory is not well-equipped for dealing with evidence for rule ordering. But taking this as a starting point, three strategies are chosen: Alderete shows that more sophisticated representations make certain opacity problems disappear; Chen argues that ordered rules can be incorporated into the theory; and Bradshaw and Roberts-Kohno claim that rule ordering phenomena are a reason for abandoning constraint-based phonology and OT altogether.

2.2 *The Representational Alternative*

Alderete discusses the phenomenon that in some languages epenthetic vowels are invisible for stress assignment. At first sight, these languages offer a paradigm example for rule ordering. The rules inserting the vowel can simply be ordered after the rule assigning stress in a rule ordering theory. A good example of a language where epenthetic vowels are invisible is Dakota. In this language the second syllable of the word is normally stressed. However, when this vowel is epenthetic, as in the following examples, this syllable is neglected, and stress is transferred to the preceding syllable.

(3)	underlying representation	output	
	ček	čéka	'stagger'
	čap	čápa	'lazy'

On the other hand, in those languages where epenthetic vowels are visible for stress, the stress rule is ordered after epenthesis. An example of such a language is Swahili. In this language stress is normally located on the penultimate syllable. Epenthetic vowels are no exception to this pattern.

Alderete argues that it is quite easy to explain the hybrid relation between epenthesis and stress assignment without invoking ordering. He does so by reinterpreting this instance of serial derivationality as an instance of relational derivationality (cf. above). This strategy takes the following form. Alderete argues that invisibility of an epenthetic vowel can be explained by prosodic faithfulness, more in particular by a constraint HEADDEP. This constraint is formulated as in (4).

(4) HEADDEP
Dep(S1,S2); A segment located in head position in S2 has to be present in S1

If S2 is the output and S1 the input, HEADDEP says that a vowel that is not present at the underlying level should not be stressed in the output. In languages, like Dakota, where epenthetic vowels are not stressed, HEADDEP is higher ranked than the constraints that account for the regular stress pattern. On the other hand, in languages, like Swahili, where epenthetic vowels follow the same pattern as normal vowels, HEADDEP is lower ranked than the constraints that account for the regular stress pattern. Alderete shows that faithfulness to prosodic heads can be motivated independently. His evidence comes from the phonology of vowel reduction; a vowel which is located in a prosodic head is more faithful to its underlying correspondent than a vowel which is not located in a prosodic head. This is the reason why reduction is usually restricted to metrically nonprominent positions.

Alderete argues that an explanation of invisibility of epenthetic vowels in terms of prosodic faithfulness is in fact superior to an account in terms of rule ordering. His argument is based on languages, where epenthetic vowels do not behave in a uniform way. In these languages epenthetic vowels may be stressed in a specific environment, yet remain consistently invisible elsewhere. A good example of such a language is Yimas. An explanation of the Yimas facts based on rule ordering would have to split up epenthesis into two rules. In between these two rules, stress assignment is ordered. This, however, leads to a significant loss of generalisation, because the two rules are formally identical. In short, rule ordering requires the bifurcation of a unitary process.

An account based on prosodic faithfulness does not suffer from this problem. Alderete shows that a system like Yimas can be characterised in the following way. Some of the constraints accounting for the distribution of stress are ranked below HEADDEP, explaining why in some environments epenthetic vowels are invisible. However, other constraints dealing with stress are ranked higher than HEADDEP, explaining why in other environments epenthetic vowels are visible for stress. In this way epenthesis can be treated in a uniform way.

Alderete's approach is to claim that there is no such thing as 'rule opacity' in the world of natural languages. Phenomena are not inherently opaque; only (rule-based) analyses are, or can be. Analysing the Dakota or Yimas facts as opaque is the result of a fairly superficial analysis of the relevant facts. Once we allow ourselves a more sophisticated view of the structures involved, and distinguish between heads and non-heads, the need to refer to more than two representations (input and output) in a serial derivation disappears completely. Approaches such as this one have sometimes been criticised, e.g. by McCarthy (1998), who points out that analyses such as these cannot be extended to other types of phenomena usually treated as 'opaque', such as Hebrew spirantization. Yet the issue of whether 'opacity' really should be considered an independent analytical category for the study of natural language may well be an open one. There is no a priori reason to believe that Hebrew spirantization and Dakota stress are similar in any way. Opacity may well be a mixed bag of unrelated phenomena, resulting from independent general properties of linguistic structures and the way these structures are related to one another.

2.3 *Rule Ordering in Tableaux*

We have seen that OT deals with constraints that characterise output configurations, not with procedures for getting these configurations. To put it differently, the output constraints as targets must be separated from the processes responding to these targets. Exactly how a given ill-formed representation is resolved is decided by the constraint ranking of the language and by nothing else. The OT model in its most widely accepted form thus claims that there are no repair strategies associated to specific (ill-formed) representations. As a result of this it can never be the case that a given repair strategy must crucially be applied in a specific way, for instance by applying it from left to right rather than from right to left, or by applying it before, rather than after, some other repair strategy. The article by Matthew Chen in this volume presents an analysis of an interesting tone sandhi phenomenon, showing that both these claims are in fact disputable. There are situations in which specific repair strategies are paired with specific ill-formed representations, and, in addition there are situations in which these strategies must be applied in a specific, directional way.

Chen's argument is based on data from the Mandarin Chinese dialect of Tianjin. With the exception of high tones, a sequence of adjacent, identical tones is not allowed in this dialect. Violations of the OCP — a good candidate for an output-based constraint — are resolved in one of three ways: (i) a sequence of two low tones is resolved by changing the first low into a rising tone; (ii) a sequence of two rising tones is resolved by changing the first one into a high tone; or, (iii) in a sequence of two falling tones the first tone is changed into a low tone. These are the three dissimilation rules of Tianjin. There is also an absorption rule, which operates specifically on a sequence consisting of a falling tone followed by a low tone. This rule changes the falling tone into a high tone.

Chen shows that the exact order in which these repair mechanisms are applied is crucial in no less than three different ways. First of all, although the rules can apply in both directions in principle, the left-to-right direction is favoured, and the opposite direction is taken only under pressure, that is, if left-to-right application leads to a result which still violates the OCP. Chen accounts for this in terms of constraint ranking. Specifically, he formulates a constraint on the preferred way in which processes should apply to an underlying form. Chen assumes that the preferred ranking for this is left-to-right for reasons of processing and calls the relevant constraint TEMPORAL. This constraint is ranked below tonal wellformedness, explaining why e.g. a sequence RRR (three rising tones in a row) is changed into HHR (two high tones followed by a rising tone), rather than into RHR. Although from a strictly representational point of view there is nothing wrong with the output RHR (it satisfies the OCP even better than the actual output HHR), this form could only be derived by applying the relevant repair mechanism from right to left. This would violate the constraint TEMPORAL. Left-to-right application is not enforced at all costs. Some underlying configurations are modified by right-to-left-application of the repair mechanisms, but only if left-to-right application would not lead to a tonally sound representation.

Furthermore, after a string has been scanned from left to right, the repair mechanisms may not apply again to remove any left over violations of the OCP. Any such derivation is avoided. Chen formulates the blocking of such a derivation as another derivational constraint, which he calls NOBACKTRACKING. In order to satisfy both the WFC and TEMPORAL backtracking is sometimes necessary. Yet Chen shows that backtracking seems to be avoided as much as possible.

The third constraint evaluating the way in which a repair strategy is applied is called PREEMPT. Chen shows that the dissimilation rules must always be applied before the absorption rule, even if this requires right-to-left application of the repair rules: dissimilation preempts absorption. Also this effect should be formulated as an OT constraint, in Chen's view.

All in all, Chen thus proposes three constraints that regulate the way in which a given repair mechanism is applied. Two constraints control the direction in which the repair mechanisms apply, and one keeps track of the order in which they are applied. All three constraints state that derivations should be as economical as possible. This makes Chen's proposals similar to certain ideas in minimalist syntax (Chomsky 1995): forms can undergo a certain (limited) number of derivations, all of which give an acceptable outcome from a representational point of view. A number of constraints on ‘economy of derivations’ then pick out the optimal one.

Chen’s proposals are somewhat harder to accommodate with mainstream Optimality Theory. Other proposals to incorporate derivationalism into the OT model can be characterised in terms of the Gen — H-Eval loop. These suggestions still leave intact OT’s basic premise; all that matters is the output. In the terminology of Chomsky (1995), there is economy of representation, but no economy of derivation. The only tool available to decide which reparation is most appropriate in a given situation are ranked constraints on phonological representations (which may sometimes compare the structure of an input representation to that of an output representation).

From this point of view, Chen’s proposal is difficult to accommodate with the line of research that is pursued by most other researchers in the OT tradition. The three relevant constraints specifically evaluate the derivational history of a given output. Chen’s careful analysis of the intriguing Chinese facts is challenging and interesting: it remains to be seen whether (and how) Chen’s facts could be analysed succesfully in an approach without reference to derivational economy.

2.4 *In Defense of Ordered Rules*

Chen tries to incorporate aspects of serialism into Optimality Theory. Two other contributors to this volume, Bradshaw and Roberts-Kohno, on the other hand, argue that certain phenomena of natural language necessarily need to be analysed in terms of serially ordered rules, because they can not be analysed in a constraint-based framework, or only be analysed in such a framework at some cost.

Mary Bradshaw analyses a complex tonal phenomenon that occurs in Suma. She claims that the tonal alternations applying in the associative construction of this language present evidence for rule ordering and therefore constitute a major problem for OT.

In a nutshell, the Suma facts are as follows. In the associative construction a final low tone becomes mid. This mid tone then spreads further to the left. It affects all tone bearing units, except the initial one. According to Bradshaw raising to mid of the final L is caused by the association of the associative morpheme, which consists of a floating [+upper]. The spreading to the left is done by a rule, called Upper Doubling. Bisyllabic and trisyllabic nouns thus are subject to the following derivation:

(5)	bisyllabic nouns		trisyllabic nouns
	LL	underlying representation	LLL
	LM	Upper Docking	LLM
	n.a.	Upper Doubling	LMM

The mid tone feeds another rule, Raised Spread. This rule spreads a high tone to the second syllable if that syllable has a mid tone. A bisyllabic noun with the underlying HL tone pattern will undergo the following derivation:

(6)	bisyllabic HL nouns	
	HL	underlying representation
	HM	Upper Docking
	HH	Raised Spread

The evidence for rule ordering comes from the interaction between Upper Docking and Raised Spread. In nouns with an initial H tone that are longer than two syllables the spreading of mid does not feed spreading of the high tone; Upper Doubling counterfeeds Raised Spread. Words of this type thus will undergo the following derivation:

(7)	trisyllabic HLL nouns	
	HLL	underlying representation
	HLM	Upper Docking
	---	Raised Spread
	HMM	Upper Doubling

The fact that Upper Docking counterfeeds Raised Spread indicates that the former must crucially be ordered before the latter. This is a problem for any theory that does not recognize rule ordering, at least as long as no alternative account (for instance in terms of more sophisticated representations) could be given. Until that moment, Bradshaw's facts pose a challenge to any adherent to a purely surface-based analysis of natural language.

Another contributor to this volume, Roberts-Kohno, similarly raises objections against a purely output-based approach to phonological theory. Roberts-Kohno shows that Kikamba has many processes that are activated before a vowel. One example is the elimination of vowel hiatus. A sequence of adjacent vowels is avoided by a process which turns the first vowel of the sequence into a glide, and lengthens the second vowel. An example illustrating this process is given in (8).

(8)	underlying representation	output	
	ko-ak-a	kwa:ka	'build'

Yet there is a class of morphemes that does not undergo this type of vowel coalescence. Kohno argues that this exceptional behavior can be explained by postulating an empty root node before the vowel. Since the empty root node separates the vowels, the environment of vowel coalescence is not met. This analysis is illustrated by the following example:

(9)	underlying representation	output	
	ko-Cind-a	ko-inda	'submerge'

Roberts-Kohno now shows that in the phonology of Kikamba a sharp division must be made between two types of process. For some processes, like vowel coalescence sketched above, the distinction between the two morpheme classes is crucial. Other processes, however, do not make the distinction. With respect to these processes, then, the morphemes with the underlying empty root node behave as if they begin with a vowel. Roberts-Kohno accounts for the distinction between the two types of process in terms of rule ordering. There is a rule deleting an empty root node. Furthermore, all the rules for which the distinction between the two morpheme classes is relevant are ordered before the deletion rule. On the other hand, all the rules for which the distinction is not relevant are ordered after the deletion rule.

OT could mimic this analysis quite easily, provided the removal of the empty root nodes takes place at a specific point in the grammar, namely at the break between two subcomponents of the grammar, for the assumption that all that matters is the output can be combined with a stratification of the grammar along the lines of lexical phonology. Of course, Roberts-Kohno is aware of this alternative. Her claim, however, is that the point where deletion of the empty root nodes takes place cannot be identified as a break between components, like the lexical and postlexical components. Although it is true indeed that none of the postlexical processes makes the distinction between the two morpheme classes, this does not mean that deletion of empty root nodes takes place at the break between the lexical and postlexical components. The reason is that, according to Roberts-Kohno, there is one lexical rule, Prefix-/k/ Deletion, which doesn't make the distinction either. Hence, deletion of the empty root nodes must apply in the lexicon, where it must be ordered before Prefix-/k/ Deletion.

Roberts-Kohno's argument that Prefix-/k/ Deletion is a lexical rule is based on the fact that the rule is sensitive to morphological structure in the sense that it applies only in specific prefixes. Since postlexical rules are not sensitive to morphological structure, Prefix-/k/ Deletion must be a lexical rule, with the concomitant effect that in the lexicon it must be ordered after deletion of the empty root nodes.

On the other hand, the claim that postlexical rules cannot see properties of individual morphemes is by no means unanimously agreed upon (for a critical assessment see Orgun 1995b, 1996). This being the case, we could just as well interpret the Kikamba facts in a different way. We could say that Prefix-/k/ Deletion is a postlexical process, even though it makes use of morphological information.[7] Then it would become possible to locate deletion of the empty root nodes at the break between the lexicon and the postlexicon. This analysis can easily be translated into an OT type of approach. In the lexical component the constraints are ranked in such a way that empty root nodes are kept intact. In the postlexical component, however, constraints are ranked in such a way that empty root nodes are eliminated. As a result, all and only lexical processes can see the distinction between the two morpheme classes; postlexical processes can no longer see the distinction.

Also Roberts-Kohno's article therefore shows that we are in need of a good theory about the interaction between morphology and phonology, either replacing Lexical Phonology or combining the frameworks of LP and OT in one way or another. Whether such a theory could be formulated within OT in such a way that also this theory could account for the facts is, we believe, an open question.

3. The Cycle

The cycle has for a long time been considered one of the hallmarks of derivational phonology. It was first formulated in Chomsky, Halle and Lukoff (1956), and reformulated several times since then. Cole (1995: 70) distinguishes three types of phenomena in which cyclicity has been invoked in the generative literature:

1. the failure of rule application in nonderived monomorphemic environments;
2. the application of a rule to a morphological environment which is a substring of a word;
3. rule ordering paradoxes — apparent violations of the strict linear ordering hypothesis, which requires all phonological rules to apply in a sequence, with each rule applying only once.

The first type of phenomenon is commonly known as *Strict Cyclicity.* It is not

---

7. We would off course still need to incorporate some amount of derivationalism, i.e. a distinction between a lexical and a postlexical component of phonology.

explicitly discussed by any of the contributors to this volume (see Polgárdi 1998 for discussion). The third type of phenomena — rule ordering paradoxes — do not play a role in the OT literature for obvious reasons: if rule ordering does not exist, rule ordering paradoxes cannot arise. (Rule ordering paradoxes are problems for rule theory, not for OT. This of course also means that the argument for the cycle is inherently weaker in OT than it is in other generative theories.)

Most attention in this volume is directed towards the second type of phenomenon: the application of rules to morphologically defined substrings of the word, apparently disregarding the elements that belong to higher-order structure. We will call this type of phenomenon 'simple cyclicity' so as to distinguish it from 'strict cyclicity'; it will be the topic of the remainder of this section.

3.1 *Cyclic OT*

Technically the simplest way to describe cyclicity effects within Optimality Theory of course is to simply incorporate the mechanism into the theory. It is as a matter of fact quite easy to do this. One could, for instance, assume that Gen is restricted in such a way that it can get at most two morphemes as an input at the same time. If a word is composed of more than two morphemes, one first has to generate a structure on one or two parts of that word, evaluate that structure, and input the resulting optimal form with another morpheme. The resulting model differs only minimally from 'classical' Optimality Theory (there is one fairly trivial restriction on Gen) but it immediately gives us a literal version of cyclicity.

More sophisticated versions of course are also possible, and may even be needed to account for the distinction between cyclic and non-cyclic processes. One of those more sophisticated approaches is represented in this volume by the article by San Duanmu. This article discusses compound stress in Shanghai Chinese. These facts are textbook examples of cyclicity: the stress patterns of complex words can only be understood if we understand the structure of simplex words first.

Duanmu argues that these facts cannot be analysed in a classical view of Optimality Theory. Such a theory would need to take recourse to an Alignment analysis of the relevant facts; one in which output-oriented constraints would align morphological structure to phonological structure, but in which we would still have only one level of representation (at which morphology and phonology would be present). Duanmu shows convincingly that such an analysis faces hard empirical problems: it cannot account for the inside-out effects that are so typical for cyclicity. For this reason, Duanmu presents a cyclic analysis in OT: one in which the components of a compound are recursively evaluated before they are submitted for evaluation as a compound. Because the Shanghai Chinese stress facts are restricted to compounds — like other Chinese dialects, Shanghai does not have productive affixation —, cyclicity is naturally restricted to the word level; as we will see below, this is an important restriction also for other scholars.

3.2 *Simple Cyclicity: Possible Non-Derivational Solutions*

Not everybody is satisfied with the adoption of cyclicity within Optimality Theory. For one reason or another, many scholars have been trying to get rid of the derivational residue that is constituted by the cycle. This problem has inspired many authors (Burzio 1994; Benua 1995; Kaye 1990; McCarthy 1996b, 1997) to adopt various strategies to this problem. We will distinguish three such strategies:

- Denial of the relevance of cyclicity for synchronic phonology (cf. Cole 1995). Authors who adopt this position argue that superficially cyclic effects are mere lexicalized relicts of historical processes. As far as they are concerned, one does not find them in synchronic phonologies of natural languages.
- Cyclicity as a *paradigmatic effect:* output forms of words in a paradigm have to be related to one another. This strategy is used in this volume by Buckley, Kager and Hayes.
- Cyclicity as a result of *representational phonology-morphology interleaving*: morphological structure and phonological structure are set up in such a way that all cyclic steps are represented at one and the same level of representation. An ingenious solution of this sort is offered by Orgun in this volume.

The first strategy seems too radical. We believe that many examples (e.g. those of Palastinian Arabic) discussed below show that the empirical claim it makes is incorrect. We do indeed find quite a range of phenomena that are best explained using some mechanism to derive synchronic cyclicity. None of the contributors to this volume seems to be willing to adopt the extreme position that cyclicity does not exist. The other two approaches are far more popular; they are defended by some of the contributors in this volume. Kager, Hayes and Buckley take the second option; Orgun the third one.

3.3 *Simple Cyclicity: Paradigmatic Approaches (Base-Oriented)*

A popular strategy in the recent literature is to assume a *paradigmatic approach* to cyclicity effects: there is a principle requiring paradigmatically related forms to be phonologically as similar as possible. This strategy can be implemented in various ways. The most popular implementation within Optimality Theory requires the phonological output of a morphologically derived form to be maximally similar to the phonological output of the 'base' of the morphological derivation. The origins of this approach within the OT framework can be found in the influential paper of McCarthy and Prince (1995). In this paper it is argued that faithfulness constraints — i.e. the constraints which require output forms to be maximally close to the input — can be formally related to constraints governing the relation between the stem and the form of the reduplicative morpheme. A reduplicative morpheme has to possess as many segments of the base as is possible given the other constraints of the language, just like a phonological output form has to remain segmentally maximally similar to the phonological input.

This is the Correspondence Theory we informally referred to above. McCarthy and Prince formalise a set of constraint schemas called *correspondence*. The two most important constraint schemas from this set are given below:

(10) For two phonological representations S_1, S_2:
Max(S_1, S_2): A segment in S_1 has to be present in S_2
Dep(S1, S_2): A segment in S_2 has to be present in S_1

If we fill in 'input' (I) for S_1 and 'output' (O) for S_2, we get ordinary faithfulness: Max(I, O) is a constraint against deletion of input material, Dep(I, O) a constraint against epenthesis of material that is not present in the input. If we fill in 'base' (B) or 'stem' for S_1 and 'reduplicant' (R) for S_2, we get constraints on the similarity between those two strings: Max(B,R) wants the reduplicant to be a complete copy of the base, whereas Dep(B,R) states that the reduplicant should not contain any segments not present in the input.

McCarthy and Prince (1995) suggest in a footnote that other values for S_1 and S_2 might also be considered. One instance would be related output forms in a paradigm. This suggestion has been followed up almost immediately by a number of papers that implemented this idea, notably the ones by Benua (1995, 1997), McCarthy (1996b) and Kenstowicz (1995).

A relatively simple example of the type of analysis provided within a theory of output-output correspondence is English stress as discussed by Benua (1997) (in essence following Burzio 1994, who argued for a paradigmatic approach within a slightly different framework of ideas). In long monomorphemic English words, we usually find secondary stress on the first syllable of the word (11a). This is probably caused by an alignment constraint (‘all words should start with a foot’) that is undominated for words of this type. Yet derived words of the same length often seem to violate this constraint, as can be seen in (11b). Chomsky and Halle (1968) have established that this is due to a cyclicity effect: derived forms preserve the stress feet of the underlying bases as much as possible. The primary stress foot of the simple forms in (b) survive as secondary stress feet in the complex forms in (11c).

(11) a. àppalàchicóla, wìnnepesáukee, lòlapalóoza, àbracadábra
b. arìstocrátic (*àristocrátic), orìginálity (*òriginálity), theàtricálity (*thèatricálity)
c. arístocrat, oríginal, theátrical

Benua uses an output-output correspondence relation on feet in order to express the relation between the forms in (11b) and those in (11c). Since the constraint referring to this relation is ranked above left alignment, the forms in (11b) cannot start with a foot. Output-output correspondence however is irrelevant in (11a), since these forms are not related to simpler words in any way.

Two of the contributors to this book adopt the correspondence approach, and develop it in interesting ways. In this subsection and the next we will briefly evaluate the proposals by René Kager and Bruce Hayes respectively.

According to Kager, derivational cyclicity is inherently wrong. He substantiates this claim by pointing at facts from Palestinian Arabic, first presented in a paper by Brame (1974, defending classical derivational cyclicity):

(12)	a.	/fihim-na/		[fhímna]	‘we understood’
	b.	/fihim-na/	[fihím-na]	*[fhímna]	‘he understood us’

The forms in (12a) and (12b) are both derived from a root or stem *fihim* plus a suffix *na*. Yet in (a) the first /i/ of the root has been syncopated, whereas this has not happened in the second form. The reason for this is that this second form is related to *fíhim* ‘he understood’. In [fihím], the first /i/ is not deleted, because stressed vowels are exempt from the syncope process. The /i/ in [fihím-na] is also not deleted, in order to preserve the similarity between the base and the correspondent.

On the other hand, the form [fhímna] in (12a) is not morphologically related to any ‘base word’ in the same way. It does not make sense to say that ‘we understood’ is derived from ‘he understood’. Therefore, correspondence is not at play, and the syncope process works unrestrictedly.

Kager's approach largely resembles Brame's (1974), except that the latter is framed in a traditional, derivational, cyclic approach. According to Kager, Brame's approach faces a problem of arbitrariness: it has to be stipulated that the addition of object clitics involves an internal cycle, whereas subject clitics are added to the root without the latter constituent undergoing its own cycle first. In essence, cyclicity predicts that the phonological form of any subconstituent in a word W can influence the phonological form of W. Kager claims that this prediction is too broad. Only subconstituents that are independent words cause cyclicity effects.

Kager's own approach, on the other hand, does not suffer from this problem, since it is based on correspondence between derived form and base, and the notion 'base' has a specialised interpretation:

(13) 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 semantic sense. (The meaning of the affixed form must contain all grammatical features of this base.) Moreover, the base is a free form, i.e. a word. This second criterion implies that a base is always an output itself.

Correspondence does not hold in the form [fhihm-na], because there is no possible base. The only 'free form' that could be relevant (/fíhim/) contains syntactic and semantic features that are not present in the affixed form — notably the 'feature' [3d person singular subject]. The situation is very different in the case of [fihím-na]; here, all the syntactic and semantic features of the independent word /fíhim/ are indeed present in the derived form as well.

Kager's proposal gives substance to the abstract notion of output-output correspondence. It restricts the number of possible correspondence relations: correspondence can only hold between a word and a subconstituent of that word (see also Kenstowicz 1994), and not even all possible subconstituents count as possible triggers for cyclicity effects. It is predicted, for instance, that output-output correspondence could not influence the phonological shape of affixes, since these by definition are not independent words.

It should be noted that (13) is a stipulation. Nothing in the theory forces us to believe that output-output correspondence should only hold between elements that occur as independent words, or that the relation should be one between a constituent and a subconstituent. Furthermore, we could essentially add the same stipulations to a rule-based theory, such as Brame's (1974). We could submit that only elements that exist as independent words in the language undergo their own cycle; the restriction on subconstituents does not even have to be stipulated in a cyclic theory (other derived forms in a paradigm cannot influence stress by definition). As a matter of fact, we have seen above that such a restriction is implicit in Duanmu's work as well. Kager's argument therefore does not really affect derivational cyclicity in an essential way, as far as we can see, although he shows that a nonderivational alternative is available.[8] It is difficult to find empirical evidence on which to compare Kager's proposals to a possible analysis along the lines of Duanmu.

In any case, this restriction has well-defined empirical consequences; it makes Kager's proposal more restricted than most others, both within correspondence theory and outside of it. This is what makes the proposal interesting. (Another interesting aspect is that Kager's proposal is very much kindred in spirit to Alderete's: not just because of the special role Head Faithfulness has to play, but also because opacity phenomena are shown to disappear if we take a more sophisticated look at linguistic structure.)

3.4 *Simple Cyclicity: Paradigmatic Approaches (Not Base-Oriented)*

The papers by Gene Buckley and Bruce Hayes in this volume are somewhat more radical than the one by Kager. Both authors develop an analysis in which paradigm uniformity is not just oriented towards the structure of the base. Buckley, for instance, defines UNIFORMITY in the following way:

(14) UNIFORMITY
If the first foot is stressed in one instantiation of a root, then it must be stressed in all instantiations of that root.

This constraint could of course easily be formalised in terms of OO-Correspondence defined on heads of feet (thus combining the formal apparatus used by Alderete and Benua), but the question arises whether Buckley's UNIFORMITY could not be restricted by Kager's observation in (13). As far as we can see, there is no *a priori* reason why such a move should be considered infeasible.

Buckley's crucial cases all involve Kashaya Foot Flipping. In Kashaya, an initial long vowel does not get stressed, if it is parsed in a monosyllabic foot; word stress is on the foot immediately following this syllable (it is always on the first foot of the word). Now, if morphological material is added to a root containing such a long vowel, the long vowel may become incorporated in a bisyllabic

---

8. Kager provides several additional arguments against a derivational approach to the facts under discussion, but these do not involve the element of cyclicity. Most of the points he notes may be problematic for a theory involving ordered rules, but not specifically for cyclicity.

foot (which actually becomes an iamb for independent reasons). But even though the word now starts with a foot, this foot does not get primary stress. According to Buckley, UNIFORMITY is responsible for this: if stress would fall on the newly formed foot, the stress pattern of the forms with Flipping would be markedly different from those in which Flipping does not apply.

As far as we can see, there are a few indications that the flipped forms (which are the main targets for UNIFORMITY) are morphologically more complex than the non-flipped forms; the least we can say is that they are longer. It therefore does not seem impossible that Buckley's facts eventually are no counterexamples to Kager's claims. (They certainly are not represented as such by the author.)

In his contribution to this volume, Bruce Hayes points to more serious problems for a standard view of cyclicity (or Kager's hypothesis) in the analysis of Yidin^y. Hayes points out that the phonological structure of paradigms is predictable 'from the inside out': given the phonological structure of a simple, underived form one can predict the form of a derived word, but not necessarily vice versa. In this context, Hayes refers to the so-called 'Wug test': "in the classic case for English, one is asked "What is the plural of [wʌg]?" and replies "[wʌgz].""

Although the classic concept of cyclicity seems to implement the idea behind the Wug test to some extent by providing a way to derive [wʌgz] from [wʌg], the inside-out effect is not represented in classical derivational theory in any way. The existence of postcyclic rules can make the Wug test fail. For instance, Dutch has a postcyclic rule of Final Devoicing which makes it impossible to answer the question "What is the plural of [vlat]" with certainty. This answer could be either "[vladə]" or "[vlatə]".

Within standard derivational theory, this disobedience to a principle of inside-out predictability is without further consequences, but Hayes argues that such a principle should be incorporated into the theory. His example derives from Dixon's (1977) work on Yidin^y. In this language, vowels in the final position of a word can get deleted under certain circumstances. This vowel only shows up in derived contexts, where it is exempt from deletion:

(15) ginda:n 'moon' — gindan**u**-ŋgu 'moon-erg.'

In the non-ergative case, the vowel is lengthened by an independent process. It is important that in this case, as in the case of Final Devoicing in Dutch just discussed, the 'bare' stem is predictable from the derived case, rather than the other way around.

Hayes now argues that language learners have largely restructured the language in such a way that they could pass the Wug test. He claims that to a large extent the quality of the alternating vowel has become predictable. The actual constraints involved and their interactions can get complicated, but in (15) the [u] is predictable because that segment almost always appears after a nasal consonant. Hayes shows that in contexts such as this an [u] almost always appears, even if cognate languages have a different vowel in this position.

In this way, the derived ergative form has become predictable from the simpler form, as well as the other way around: given any form in a paradigm, we can predict the other forms. (If Hayes is right, Kager's theory may be too strict.) Hayes claims that this cannot be expressed in classical rule-based phonology, and that even a standard view of Optimality Theory in which at least the derivational levels of input and output are recognized, has problems. It is impossible, for instance, to work with Morpheme Structure Constraints, which would state for instance that nasal consonants have to be followed by [u]. The problem is that there are also stems in which the final vowel does not get deleted; it would be a coincidence that exactly these stems would not confirm the Morpheme Structure Constraints in question.

Hayes' position on derivationality therefore is the most radical one among the authors in this book: he suggests that input-output relations may not play a role in the grammar at all. If we have correspondence relations among forms in a paradigm, we might not need to set up an independent and uniform underlying representation at all.

It should be noted that taking this step is fairly radical indeed. We would need to revise many ideas that have always been central to the generative view of phonology. For instance, it is no longer possible to predict whether a given phonological form is wellformed on the basis of the grammar of a language alone: we also need to know the whole paradigm in which the form in question is supposed to fit. Hayes himself acknowledges the fact that this position is radical. It is certainly worth exploring the consequences of such an approach.

3.5 *Simple Cyclicity: A Representational Approach*

However radical paradigmatic proposals to cyclicity may seem, they are still inherently derivational: maybe not in the technical sense usually attributed to this term, and therefore not in the same sense in which Duanmu's contribution to this volume is derivational, but certainly in the sense of Anderson (1985), for the emphasis in all of these proposals is on the way in which phonological representations can be related to one another.

Paradigmatic proposals may or may not have advantages over 'serial' views of the derivation, such as Duanmu's, they all share one important problem: they have to stipulate that the most common way in which a base form and a derived form are related is that the structure of the base form influences the structure of the derived form, and not the other way around. This follows immediately from a serial account: in order to build the derived form, one first has to build the base form; but it is not necessary to build the derived form in order build the base form. This is actually the core idea of cyclicity. In a parallel approach to derivations, on the other hand, there is no a priori reason why the output of a derived form could not influence the structure of a base form. If this is the case, it should be stipulated.

The position represented in this volume by the paper of Orhan Orgun (based on work presented more extensively in Orgun 1996) is the most representational in the strictest sense. Or rather, the boundaries between derivation and representation are vague in the tradition in which Orgun's work can be situated, that of Head-Driven Phrase Structure Grammar (Pollard and Sag 1994). Orgun points out that many aspects of serial derivation are unnecessary once we take the concept of morphological structure seriously. The morphological structure of a complex form contains by definition the information of the base, but not vice versa. The direction of influence is readily explained in more or less the same way as in the serial accounts. In many ways, Orgun's proposals can be seen as a restatement of classical insights within a representational framework. A representational theory is about the internal structure of a linguistic form; a derivational theory is about comparing two structures. If we can still 'see' the structure of a base in the structure of a derived form, the boundary between the two theoretical modules disappears.

It is difficult to compare Orgun's proposal to the alternatives presented by Buckley, Kager and Hayes. The difference with Kager seems mainly technical: if Kager is empirically right, Orgun should incorporate the assumption that cycles can only be built on complete words. If Buckley's and Hayes' predictions turn out to be right, on the other hand, the problems for Orgun would be much more severe: but the same may be true for Kager, if Hayes is right. Also for this reason, this issue is worth to be explored in more detail.

4. Conclusion

A complete theory of phonology should account for the regularities of structure within one given phonological representation and for the way in which phonological representations can be related to each other: a complete theory of phonology needs both a theory of representations and a theory of derivations.

Views of both representations and derivations on first sight have changed quite dramatically over the past few years, with the advent of Optimality Theory and similar ideas of constraint-based phonology. For instance, some researchers have abandoned 'abstract' representational assumptions about autosegmental and prosodic structure and embraced a more 'concrete', phonetically-oriented, view of sound structure. OT's initial assumptions about derivations, as presented in Prince and Smolensky (1993), McCarthy and Prince (1993a, 1993b) and related work, offered a revolutionary change. Taking the position that only two separate representations were involved in the evaluation of a phonological form — input and output, of which only the former was subject to independent wellformedness constraints — is a minimalist hypothesis, and in this respect it offered an interesting alternative to the convoluted derivational models of previous times. The empirical claim was fairly strong, and it was challenged almost immediately. The response to this has been to weaken the claim and allow Optimality Theory to consider more representations in parallel.

Indeed, if one wishes, one can incorporate so many derivational mechanisms into one's optimality-theoretical engine, that one gets a theory that is at least as powerful as the theories of the 1980s. In our view, the question is whether this is a desirable state of affairs. We probably need a theory of derivations that is more powerful than Prince and Smolensky's (1993), but less powerful than one that incorporates all of OO-Correspondence, Sympathy, lexical levels, Chen's constraints on derivational economy, etc. It remains to be seen what the exact compromise should be. We hope that the present volume will help to clarify this issue.

Acknowledgments

The fact that the names of the authors of this introduction appear in reversed alphabetical order and the names of the editors of this volume in alphabetical order, is arbitrary and does not reflect a hierarchical ranking between the two researchers involved. The collective list of references for this introduction and for all other articles can be found at the end of this book.

Head Dependence in Stress-Epenthesis Interaction

John Alderete

1. Introduction

What is the nature of the interaction between stress and epenthesis? Do epenthetic syllables count in word stress, or not? This paper will study these questions from various angles and discuss the theoretical issues they raise.

In SPE style phonology (Chomsky and Halle 1968), stress-epenthesis interaction depends on rule ordering. If vowel insertion is ordered before stress assignment, epenthetic vowels will be counted and stressed according to the regular pattern; conversely, if stress precedes epenthesis, then the inserted vowels will be inactive in stress. While the Rule Ordering theory can account for virtually any pattern of stress-epenthesis interaction, this theory fails to offer an explanation of the phenomena. The behavior of epenthetic vowels in stress is described by stipulating the required rule ordering, leaving us to wonder why the state of affairs could not be different.

Working within theories of Prosodic Phonology, some researchers have tried to improve on the Rule Ordering theory by considering the role of prosodic representations in deriving stress-epenthesis interaction (see Broselow 1982 for example). In this approach, a class of epenthesis rules are identified as syllabically conditioned (as in Selkirk 1981, 1984; Itô 1989), and this kind of epenthesis *must* be counted in stress because of general principles of prosodic organization. In particular, syllables must be dominated by prosodic feet, and this prosodic layering works 'from the bottom-up'. As an example of how the Bottom-Up theory works, consider Broselow's analysis of the interaction between stress and epenthesis in Swahili.

Swahili regularly stresses the penult (1a). Further, epenthetic vowels introduced in loanwords are counted and stressed according to the canonical pattern (1b); in the examples below and throughout, epenthetic vowels are underlined.

(1) *Swahili (Ashton 1944; Polomé 1967; Broselow 1982)*

a.	jíko	'kitchen'
	jikóni	'in the kitchen'
	nilimpíga	'I hit him'
	nitakupíga	'I shall hit you'
b.	tíket ~ tikéti	'ticket'
	rátli ~ ratíli	'pound'

In the Bottom-Up theory, the explanation goes like this. Epenthesis of *i* is syllabically conditioned because it applies in order to syllabify obstruents as onsets (coda obstruents are generally avoided in the language). Working from the bottom-up, epenthetic syllables are inserted to parse the unsyllabified obstruents, and stress feet are then built over these syllables. With this order of events, the interaction between stress and epenthesis could not be otherwise: epenthetic vowels must be active in word stress because they are an automatic by-product of inserting an epenthetic syllable, which in turn forms the building blocks for stress feet.

While the Bottom-Up theory makes a significant improvement on the Rule Ordering approach, there is an empirical problem with this theory. It cannot account for syllabically conditioned epenthesis which is invisible in stress; the assumptions inherent to this theory predict that this class of behavior does not exist. Consider the following example from the Siouan language Dakota as a counterexample to the Bottom-Up theory.

In Dakota, stress regularly falls on the second syllable from the beginning of the word (2a). Yet syllabically conditioned epenthesis into the second syllable (2b) creates exceptions to canonical second syllable stress (see Kennedy 1994 and Sietsema 1988 on a syllable-based analysis of epenthesis in Dakota).

(2) *Dakota (Shaw 1976, 1985)*

a.	č^{h}ikté			'I kill you'
	mayákte			'you kill me'
	wičháyakte			'you kill them'
	owíčhayakte			'you kill them there'
b.	/ček/	→	čéka	'stagger'
	/khuš/	→	khúša	'lazy'
	/čap/	→	čápa	'trot'
	cf. /kte/	→	kté	's/he, it kills'

This pattern of stress-epenthesis interaction presents a clear counterexample for the Bottom-Up theory: *a*-epenthesis is syllabically conditioned; and since the organization of syllables into stress feet proceeds from the bottom-up, the epenthetic syllable should be stressed according to the regular pattern of peninitial stress. But this is not correct for Dakota, which calls into question the explanation that the theory offers for other languages. In order to account for the Dakota pattern, stress assignment must be ordered after *a*-epenthesis, and once rule ordering is admitted in the theory, the account of the Swahili pattern is no different from the Rule Ordering approach.

We are left, it would seem, with some version of the Rule Ordering theory, and there is a reason for rejecting this theory as well. Epenthetic syllables do not always behave in a uniform way in relation to stress. They can be ignored in some environments, and yet incorporated into the stress pattern in others. Stress-epenthesis interaction in the Papuan language Yimas is like this, and as we will see in detail below, such patterns point to a real flaw in the Rule Ordering approach.

In Yimas, the main stress regularly falls on the initial syllable of a word (3a). Epenthesis into this position, however, creates exceptions to regular initial stress, pushing stress forward a syllable (3b).

(3)	*Yimas (Foley 1991)*				
	a.	wáŋkaŋ			‘bird’
		kúlanaŋ			‘walk’
		wúratàkay			‘turtle’
		mámantàkarman			‘land crab’
	b.	/pkam/	→	pɨ̠kám	‘skin of back’
		/tmi/	→	tɨ̠mí	‘say’
		/kcakk/	→	kɨ̠cákɨ̠k	‘cut’
		/nmpanmara/	→	nɨ̠mpánmara	‘stomach’

There is a further complication on this pattern, which is that if the vowels of both the first and second syllable are derived by epenthesis, then main stress defaults to the initial syllable (4). (See Foley 1991:44 ff. motivation of an epenthetic analysis in words with long strings of consonants such as these.)

(4)	/tkt/	→	tɨ̠́kɨ̠t	‘chair’
	/klwa/	→	kɨ̠́lɨ̠wa	‘flower’
	/krmknawt/	→	krɨ̠́mkɨ̠nawt	‘wasp’
	/tmpnawkwan/	→	tɨ̠́mpɨ̠nàwkwan	‘sago plam’

In sum, epenthetic vowels are generally invisible to stress (3b), but in a phonologically defined context, epenthetic vowels are stressed according to the regular pattern (4).

The derivational theory needs to say that Yimas has two rules of epenthesis. One process of ɨ-insertion must apply before the assignment of initial stress in order to account for the fact that an epenthetic syllable is stressed when followed by another epenthetic syllable. A second rule of ɨ-insertion, on the other hand, must follow stress assignment because of the fact that epenthesis, as the elsewhere case, creates exceptions to regular initial stress. The problems with the Rule Ordering theory, therefore, run deeper than simply failing to explain stress-epenthesis interaction. In cases like Yimas, the rule-ordering approach leads to loss of generalization in the analysis of the epenthesis process itself. Concretely, the epenthesis process yields a uniform structural change and it is motivated as a means of syllabifying consonant clusters according to the phonotactics of the language. But the analysis of Yimas in terms of ordered rules misses these generalizations by positing two rules of epenthesis. The observations regarding the output of epenthesis, and that epenthesis is syllabically motivated, are stated more than once in the grammar.

In this paper, I propose to account for the problematic cases of Dakota and Yimas as constraint interaction within Optimality Theory (Prince and Smolensky 1993; McCarthy and Prince 1993b). The idea developed below is that Universal Grammar has a well-formedness constraint, HEAD-DEP, that bans the stressing (and footing) of epenthetic segments. When HEAD-DEP dominates a set of constraint responsible for 'regular stress', the result is that epenthetic vowels are invisible in stress, as in the case of Dakota. However, if HEAD-DEP is low ranking, a different pattern of stress-epenthesis interaction is predicted, i.e., metrical activity of epenthesis, as in the case of Swahili. Moreover, the constraint interaction theory developed here provides a clear line of analysis for the more complicated cases exemplified above with Yimas. For such cases, the precise details of the system can be directly characterized by interleaving HEAD-DEP with the set of constraints deriving the regular pattern.[1]

The remainder of this article is structured as follows. In Section 2, I will develop and motivate HEAD-DEP, the constraint which plays a central role in the OT analysis. The constraint will then be applied to the examples of Dakota and Swahili as exemplification of the basic proposal. Section 3 gives a nonderivational analysis of stress-epenthesis interaction in Yimas, and the advantages of

---

1. Two further constraint-based approaches to stress-epenthesis interaction have been proposed in unpublished work: (i) Kennedy (1994), who proposes to explain cases like Dakota through the alignment of morphological and prosodic structure, (ii) and Ikawa (1995), who deals with the avoidance of stressing epenthetic vowels within the theory of Local Conjunction structured in Smolensky (1993). Unfortunately there is no space in this paper to review these proposals in detail.

the constraint-based approach are discussed. In the final section, Section 4, the main results reached in this paper are summarized and some theoretical issues raised in the preceding sections are discussed, namely the role of serial derivation in phonology, the character of input-output faithfulness, and position-sensitive faithfulness.

2. Head Dependence in Stress-Epenthesis Interaction

This section begins with a restatement of stress-epenthesis interaction in Dakota, and then, as a means of motivating the notion of Head Dependence, a comparison is made to stress-related vowel reduction in Russian. The necessary constraints are then formalized and applied to the analysis of particular languages.

Recall from the above discussion that epenthesis in Dakota is invisible to stress: when epenthesis inserts a vowel into the regularly stressed syllable, stress is shifted to avoid stressing the epenthetic vowel. More generally, it seems that in Dakota, the stress system avoids stressing vowels which are not present underlyingly. As a first approximation of the constraint HEAD-DEP, we can say that noncanonical stress in Dakota is due to the following principle: the stressed vowel must have a lexical counterpart in the input. Applying this constraint to the examples below, stress may be assigned to the canonically stressed syllable, i.e., the syllable to which stress is assigned by the regular pattern, if it contains a lexical vowel (5a). But if the second syllable contains an epenthetic vowel, stress falls elsewhere in the word because of the requirement that stressed vowels have lexical counterparts (5b).

(5) *Stress-Epenthesis Interaction in Dakota*

	a.	b.	
	/ č^{h} i k t e /	/ č e k /	INPUT
	↑	↑	
	[č^{h} i k t é]	[č é k $\underline{a}$]	OUTPUT

The stressed vowel is said to be ‘input-dependent’ in the sense that it must have a counterpart in the input — even though this may lead to exceptions to the canonical pattern of stress, as in *čék*$\underline{a}$.

This notion of input-dependence makes it possible to construct a clear parallel in the domain of segmental processes sensitive to stress. Stronger faithfulness requirements on stressed vowels are also essential to the characterization of a common form of vowel reduction. Consider the case of Russian, one of a wide range of similar languages (see Beckman 1998 and Flemming 1993 for comprehensive surveys). In tonic positions, Russian licenses six full vowels, i.e.,

/i ɨ e a o u/; but in unstressed positions, only the three peripheral vowels surface.[2] This observation holds of lexical forms, and is supported by morphophonemic alternations. For example, the stem-internal mid vowel surfaces under stress in the nominative form *stól*, yet in forms where stress is moved off the stem vowel, underlying /o/ lowers to *a*, e.g., *stal-óf*. This process of vowel reduction, referred to as 'A-Kanje', is exemplified with the nominal and verbal paradigms below.

(6) *Russian A-Kanje (Jones and Ward 1969; Zubritskaya 1995)*

a.	Nom. Sg.	stól	slóv-o
	Gen.	stal-á	slóv-a
	Dat.	stal-ú	slóv-u
	Nom. Pl.	stal-ý	slav-á
	Gen.	stal-óf	slóv
	Dat.	stal-ám	slav-ám
		'table'	'word'
b.	glaž-ú	važ-ú	1 per. Sg.
	glóž-iš	vóz-iš	2 per.
	glóž-it	vóz-it	3 per.
	glóž-im	vóz-im	1 per. Pl.
	glóž-it'i	vóz-it'i	2 per.
	glóž-ut	vóz-ət	3 per.
	'gnaw'	'carry'	

The fact that the stressed vowel resists the general pattern of reduction suggests a position-sensitive requirement on a par with the one employed above for Dakota. In particular, suppose that the quality of the stressed vowel must be identical with its lexical counterpart. As illustrated in the input-output mappings below, mid vowels lower generally because of a context-free ban on mid vowels (discussed below), as the stem vowel does in (7b). But this lowering process does not apply if the vowel occurs in an accented position. In such cases, mid vowels remain faithful to their input specification because of the high-ranking identity requirement for stressed vowels.

2. Russian vowel reduction is more complicated than this, requiring the distinction between three distinct domains (i.e., the tonic syllable, the pretonic syllable, and the complement set of syllables), as mid vowels reduce to a peripheral vowel in the pretonic position, but to a schwa elsewhere (Jones and Ward 1969). See Alderete (1995b) for discussion of the theoretical implications of this three-way pattern of vowel reduction.

(7) *Vowel Reduction in Russian*

a. / s t o l /	b. / s t o l — o f /	INPUT
‖	‖	
[s t ó l]	[s t a l — ó f]	OUTPUT

Characterizing vowel reduction in Russian as such paves the way for relating this observation to the metrical inactivity of epenthesis in Dakota. Both cases involve a constraint on the relation between the stressed vowel and its input counterpart, as restated directly below.

(8) a. Stressed vowels must have counterparts in the input.
b. Stressed vowels must be identical to their input counterparts.

Furthermore, the above constraints have the effect of suppressing general phonological patterns. Hence, the requirement in (8a) effects noncanonical initial stress in Dakota, and the requirement in (8b) characterizes the fact that stressed vowels fail to undergo vowel reduction. The parallels observed here are striking, and call for a formal basis for relating the two phenomena.

Both of the requirements given in (8) assert stronger requirements for stressed vowels, and in doing so, they require reference to 'counterparts' in related structures. This notion of a counterpart is fundamental to the theory of faithfulness proposed in McCarthy and Prince (1995, M&P henceforth). As a direct account of the parallels observed between reduplicative copying and faithfulness of input to output, M&P generalize the notion of correspondence developed in McCarthy and Prince (1993b) to input-output faithfulness. Correspondence between input and output provides a formal characterization of the concept of a counterpart (read as correspondent) employed in the above descriptions.

(9) *Correspondence (McCarthy and Prince 1995)*
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. Segments α (an element of a string S_1) and β (an element of a string S_2) are rereferred to as correspondents of one another when $\alpha\mathfrak{R}\beta$.

With this characterization of correspondence, the requirements driving the apparently aberrant patterns in vowel reduction and stress-epenthesis interaction can be stated more formally. Faithfulness of input to output is embodied in a set of constraints on correspondent elements, which in the case of the present study, involves correspondent segments. The constraints given in (8) also involve input-output faithfulness, with special reference to metrically prominent positions, i.e., prosodic heads like the main stressed syllable or the main stress foot of a prosodic word.

(10) HEAD-DEPENDENCE (Alderete 1995b)
Every segment contained in a prosodic head in S_2 has a correspondent in S_1.
If β is contained in a prosodic head in S_2, then $\beta \in$ Range($\Re$).

(11) HEAD-IDENTITY[F] (McCarthy 1995; Alderete 1995b; Zubritskaya 1995, Beckman 1998)
Correspondent segments contained in a prosodic head must be identical for F.
If β is contained in a prosodic head in S_2, and $\alpha\Re\beta$, then α and β agree in the feature F.

The proposed constraints employ the basic faithfulness constraints of M&P, and simply refine their application to certain metrically strong positions. Hence, HEAD-DEPENDENCE (HEAD-DEP henceforth) is a refinement of the anti-epenthesis constraint DEPENDENCE; and HEAD-IDENT(ITY)[F] employs the same modification for the class of featural faithfulness constraints IDENT[F]. The consistent modification to these constraints is therefore the specification of a prosodic target, and with this modification, and nothing more, the two classes of phenomena are explained.[3]

Starting with the first constraint, the effect of HEAD-DEP is that prosodic heads are input-dependent. That is, only segments with input correspondents may occur in metrically strong positions. Because epenthetic vowels are introduced in the mapping from the input to the output, they have no input correspondents (M&P), and so parsing them internal to the prosodic head of a word will constitute a violation of HEAD-DEP. This notion of Head Dependence will be applied to several cases of stress-epenthesis interaction below.

Similarly, HEAD-IDENT[F] explains resistance to vowel reduction in stressed positions. Vowel reduction is part of a larger distributional pattern whereby a wider range of contrasts are licensed in strong positions than those allowed in metrically weak positions. With HEAD-IDENT[F] high-ranking in the grammar, this distributional asymmetry is explained; and what is more, it extends to the

3. It is highly likely that faithfulness constraints defined for prosodic heads have a functional basis in pyscholinguistic theories of lexical access. The prosodic faithfulness constraints employed here ensure preservation of the lexical specification for stressed units. Roughly speaking, both of these constraints protect lexical information from being destroyed in the surface form by regular processes of the language. This accords nicely with pyscholinguistic evidence that strong syllables play an important role in segmentation for lexical access (see Cutler and Norris 1988 for crucial experimental results, and Beckman 1998 for a comprehensive review of the psycholinguistic literature and discussion of its theoretical implications).

above morphophonemic alternations which involve stress shift. Returning to the example of Russian A-Kanje, /o/ lowers to *a* generally, but this regular pattern of vowel reduction is suppressed in the accented syllable, e.g., /stol-of/ → *stalóf*. As shown in the informal analysis below, HEAD-IDENT[F] plays a crucial role in deriving this fact.

Sketching the basic components of the analysis, I assume a theory, pursued in Beckman (1995) for vowel harmony, that phonological processes can be motivated as a means of minimizing structural markedness. Specifically, reduction of a mid vowel can be seen as a way of avoiding a violation of the featural markedness constraint *MID, a context-free constraint which yields a "*" for every mid vowel in a form. However, mid vowels fail to undergo vowel reduction in stressed syllables because of high-ranking HEAD-IDENT[F]: faithfulness to the vowel features characterizing mid vowels is ensured by this position-sensitive constraint. This is illustrated in the following OT tableau.

(12) *Head Identity in Russian A-Kanje: /stol-of/ → stalóf*

/stol-of/	HEAD-IDENT[F]	*MID	IDENT[F]
a. stolóf		**!	
b. staláf	*!		**
c. ☞stalóf		*	*

The first candidate is fully faithful to the input, but it is ruled out by *MID because it has more violations of this constraint than its competitors, and *MID dominates the context-free faithfulness constraint IDENT[F]. The candidate in (12b) obeys *MID completely by lowering both mid vowels, but in doing so, this form is unfaithful to the featural specification of the stressed vowel, and this results in a fatal violation of top-ranked HEAD-IDENT[F]. The optimal form (12c), therefore, is the one which is both faithful to the features of the stressed vowel, and minimally violates the featural markedness constraint *MID by lowering all vowels elsewhere in the word. To summarize, the driving force behind vowel reduction is given a general account, while avoidance of vowel reduction in stress syllables is described with the head-sensitive faithfulness constraint.

Returning to stress-epenthesis interaction, an application of HEAD-DEP to a concrete example will serve to clarify its interpretation, and to establish the parallel between the two classes of phenomena under discussion. Recall from the introduction that epenthesis is inactive in Dakota stress: surface stress is realized on the peninitial syllable, yet epenthesis into the second syllable correlates with initial stress. In the analysis developed below, noncanonical stress is characterized by ranking HEAD-DEP above a constraint responsible for deriving regular stress.

Before describing the pattern of initial stress, we start first with the constraints governing the canonical stress pattern. Following Shaw (1985) and Hayes (1995a), second syllable stress is derived by forming an iambic foot which is properly aligned with the left edge of the word.[4] Iambic structure is ensured by the rhythm type constraint, RHTYPE = I (Prince and Smolensky 1993), and this iamb must be binary by high-ranking Foot Binarity (McCarthy and Prince 1986; Hayes 1995a). The fact that stress prominences are not found subsequent to the second syllable suggests that foot construction is noniterative (Shaw 1985). And following McCarthy and Prince (1993a), nonrepeating stress is derived by a high-ranking alignment constraint, ALIGN-L (F,PrWd), which prohibits iterative footing by requiring that the left edge of all feet coincide with the left edge of some prosodic word.

These constraints on the location and form of feet enter into conflict with Head Dependence when epenthesis inserts a vowel in a regularly stressed position. The stress foot constraints posit the head syllable of the word peninitially, but stressing an epenthetic vowel in this position leads to a violation of HEAD-DEP. If HEAD-DEP is top-ranked, however, noncanonical stress will be the result, as illustrated in the following tableau.

(13) *Metrical Inactivity of Epenthesis in Dakota: /ček/ → čéka*

/ček/	HEAD-DEP	RHTYPE = I
{če ká}	á!	
☞{čé ka}		*

In the candidates above, epenthetic *a* has no input correspondent; this is because epenthetic vowels by definition do not stand in correspondence with underlying vowels. Therefore, parsing *a* internal to the syllable head of an iambic foot, as in the first candidate, fatally violates HEAD-DEP. The optimal candidate is thus the form which satisfies the input-dependence constraint by reversing the rhythm

4. Dakota stress is not sensitive to syllable weight, which according to standard foot typologies constitutes evidence against iambic rhythm. But see Shaw (1985) for three empirical arguments in favor of the iambic analysis of second syllable stress, and discussion in Hayes (1995a) for a historical account for how this ‘defective iambic system’ could have developed.

type of the stress foot.[5] In summary, the notion of Head Dependence developed here permits an adequate analysis of Dakota stress-epenthesis interaction in parallelist OT.

We are now in a position to make the comparison between the two phenomena exemplified in Dakota and Russian on a more formal level. In Dakota, the avoidance of epenthetic vowels to participate in stress is derived by ranking the position-sensitive faithfulness constraint HEAD-DEP above the constraint yielding a canonical iambic foot. Likewise, in the case of Russian, the resistance on the part of stressed vowels to undergo vowel reduction is due to high-ranking HEAD-IDENT[F]. Thus, the two classes of phenomena, radically different in surface form, are described with a consistent modification of context-free faithfulness. Some further implications of position-sensitive faithfulness are considered at the end of this paper.

Returning to the role of Head Dependence in stress-epenthesis interaction, in the analysis of Dakota, the avoidance of stressing epenthetic vowels is an effect of HEAD-DEP defined for the syllabic head of the prosodic word. It is a straightforward matter to extend this result to cases in which the stress system also fails to *count* the epenthetic vowel in the assignment of stress. In such cases, as exemplified below with the Austronesian language Selayarese, the prosodic head relevant for the meaning of HEAD-DEP is the main stress foot. Failure to count an epenthetic vowel is thus described as a failure to foot the epenthetic material. In other words, the segmentism of the main stress foot is input-dependent.

In Selayarese, surface stress regularly falls on the penultimate syllable (14a). But epenthesis into the final syllable, for the purpose of syllabifying certain consonants as onsets, yields irregular antepenultimate stress (14b).

(14) *Selayarese (Mithun and Basri 1985)*

a.	állo	‘day’
	allónni	‘this day’
	pá:o	‘mango’
	paó:ku	‘my mango’
b.	ká:tala	‘itch’
	pó:tolo	‘pencil’
	maŋkássara	‘Macassar’
	lámbere	‘long’

5. See Prince and Smolensky (1993) for a similar approach to Southern Paiute stress in which the constraint responsible for extrametricality effects, NONFINALITY, conditions a rhythm type reversal.

Summarizing the facts here, final syllable epenthesis creates exceptions to the canonical pattern of penultimate stress by pushing stress back one syllable.

To sketch an account of this pattern, I assume first that regular penultimate stress in Selayarese is the result of positing a disyllabic trochee at the right edge of the word. To be more concrete, the rhythm type constraint, RhType = T, and Foot Binarity are high-ranking, and the relevant alignment constraint, Align-R, enforces alignment of the right edge of the prosodic foot with the right edge of the word. In sum, the constraints on the form and position of prosodic feet posit the main stress foot at the right periphery of the word to yield the canonical pattern.

Noncanonical antepenultimate stress arises from a different form of constraint interaction, namely constraint conflict between Head-Dep and Align-R. The two constraints enter into conflict in a context where epenthesis inserts a vowel within the domain of the stress foot. Therefore, by defining Head-Dep for the main stress foot, and giving this constraint high rank in the system, the invisibility of final epenthesis in stress is directly obtained.

(15) *Metrical Inactivity of Epenthesis in Selayarese: /katal/ → ká:tala*

/katal/	Head-Dep	Align-R
ka {tá:la}	a!	
☞{ká:ta} la		la

The result illustrated here is that the final syllable is skipped, in violation of Align-R, because parsing it as a weak member of the trochee would violate Head-Dep set for the stress foot. Thus, this pattern of metrical inactivity of epenthesis is characterized by a constraint ranking in which Head-Dep is ranked above a constraint deriving a regular stress pattern, as in the case of Dakota stress-epenthesis interaction.

To foreground an important point, the approach to stress-epenthesis interaction which employs the notion of Head Dependence differs fundamentally from the Rule Ordering theory in the way that phonological activity of epenthesis is characterized. In the derivational model, the behavior of epenthesis in stress is a matter of serial derivation: metrically active epenthesis is early in the derivation, while invisibility of epenthesis in the stress system is indicative of a later rule. The parallelist theory proposed here does not allow intermediate stages in the derivation, and so phonological activity of epenthesis cannot be characterized in this way. Rather, activity of epenthesis in stress is simply a matter of constraint ranking, a fundamental property of Optimality Theory. If Head-Dep is high-ranking relative to a set of constraints responsible for deriving regular stress, i.e., 'C_{Stress}', then epenthesis is metrically inactive (16a), as is the case in both Dakota and Selayarese. On the other hand, if HEAD-DEP is low-ranking in the constraint system, then epenthetic vowels will be active in the system (16b).

(16) a. HEAD-DEP » C_{Stress}: metrical inactivity of epenthesis
b. C_{Stress} » HEAD-DEP: metrical activity of epenthesis

The point here is that the constraint interaction theory characterizes behavior of epenthesis with the position of HEAD-DEP in the constraint system. All patterns of stress-epenthesis interaction are thus predictable on the basis of ranking of HEAD-DEP relative to the constraints governing stress. To complete the typology of stress-epenthesis interaction then, let us consider how the schematic ranking given above applies to the case of metrical activity of epenthesis in Swahili.

Swahili has canonical penultimate stress, so for the present purposes, the regular pattern will be derived by the same constraints employed above in the analysis of Selayarese. A syllabic trochee is formed at the right periphery of the word, showing that both ALIGN-R and RHTYPE = T are at play in the system. In contrast to Selayarese, however, epenthetic vowels are active in Swahili stress; they are counted and stressed according to the regular pattern of penultimate stress. Within the framework of ideas developed here, this observation entails that the two stress related constraints dominate HEAD-DEP. Thus, ALIGN-R dominates HEAD-DEP, as illustrated in (17a), to account for the fact that word-final epenthesis fails to invoke improper alignment of the stress foot. Furthermore, RHTYPE = T also outranks HEAD-DEP to account for the lack of a rhythm type reversal (17b), as was seen to be the case in Dakota.

(17) a. *Metrical Activity of Epenthesis in Swahili: counting of epenthetic vowels*

/tiket/	ALIGN-R	HEAD-DEP
{tíke} ti̲	ti̲	
☞ti {kéti̲}		i̲

b. *Metrical Activity of Epenthesis in Swahili: stressing of epenthetic vowels*

/ratli/	RhType = T	Head-Dep
ra {ti̲lí}	*!	
☞ra {tí̲li}		i̲

In sum, the activity of epenthesis in the regular stress pattern is accounted for with the schematic ranking proposed above: metrical activity of epenthesis follows from low-ranking Head-Dep.

To summarize the empirical results reached in this section, the notion of Head Dependence was developed as a means of modeling stress-epenthesis interaction in parallelist OT. The case of metrical inactivity of epenthesis in Dakota was handled with a constraint ranking in which Head-Dep dominates the constraint governing the canonical foot. This result was then extended to the example of Selayarese, where the invisibility of epenthesis in stress was derived by ranking Head-Dep above the alignment constraint responsible for positioning the stress foot within the word. Finally, the opposite rankings were employed in the analysis of Swahili, where the constraints of the structure and position of feet dominate Head-Dep. These rankings account for the fact that epenthetic vowels are stressed and counting according to regular stress in Swahili. In closing, the theory of stress-epenthesis interaction developed here handles straightforwardly the cases shown to be problematic for the Bottom-Up approach discussed in the introduction. In the next section, we turn to the cases which point to a problem for the derivational Rule Ordering theory, namely cases where epenthetic vowels are only partially visible to stress.

3. Nonuniformity in Stress-Epenthesis Interaction

The interaction between stress and epenthesis is not always a uniform and across-the-board phenomenon. Thus, epenthetic vowels may be stressed in a specific context, and yet consistently inactive elsewhere in the stress system; epenthesis may be only partially active in word stress. For example in Spanish, epenthesis into initial *sC* clusters is ignored by the stress system (Harris 1970; McCarthy 1980), and yet the same process, applied as a way of resolving triconsonantal clusters, is active in stress (Harris 1977; Alderete 1995b). A second example of partial metrical activity of epenthesis in stress is found in the

Iroquian language Mohawk. In this language, syllabically motivated epenthesis generally breaks up obstruent + resonant clusters, but the sensitivity of the process to the basic accent pattern is mixed: epenthesis into biconsonantal clusters is inactive in the system, yet epenthetic vowels which surface in a closed syllable are stressed according to the regular pattern (Michelson 1981, 1988; Piggott 1995). Similar cases of partial metrical activity of epenthesis are observed in the Malayo-Polynesian language Lenakel (Lynch 1978), in Arabic dialect phonology (see especially Farwaneh 1995), and in the Papuan language Yimas discussed in the introduction. In this section, the complicated interactions between stress and epenthesis in Yimas will be studied and its implications for the role of derivationalism in phonology will be discussed.

To review the facts fleshed out in Section 1, words in Yimas are regularly stressed on the initial syllable, but epenthesis into initial clusters causes two complications for the regular pattern. First, if the initial syllable contains an epenthetic vowel, stress is shifted to the second syllable, e.g., *pɨ̱kám*. This observation shows that the system avoids stressing epenthetic vowels.

However, epenthetic vowels may be stressed in a particular context. When both the first and second syllable contain epenthetic vowels, stress defaults to the canonical position, i.e., the initial syllable, as in *krɨ̱́mkɨ̱nawt*. Any theory of stress-epenthesis interaction needs to account for the mixed behavior of epenthetic vowels in cases like this.[6]

Let us begin the analytical work by considering how partial metrical activity of epenthesis is derived within a derivational model. The basic premise of the Rule Ordering theory is that activity in stress is derived by the presence of some structure at the derivational instant at which stress is assigned. In Yimas, this entails a characterization of epenthesis as a pair of rules along the following lines. One rule of epenthesis, Epenthesis$_1$, applies before the assignment of initial stress, operating essentially in the context of triconsonantal clusters which cannot be incorporated into well-formed syllables. An independent rule of vowel insertion, Epenthesis$_2$, which is also motivated in contexts of unsyllabified consonants, must follow stress assignment. This is illustrated for two crucial forms in the following derivation.

6. Secondary stress is assigned to the third syllable from the beginning of the word, suggesting that stress obeys a binary pattern. Discussion of the interaction of secondary stress and epenthesis is postponed to the end of this section.

(18) *Partial Metrical Activity in a Derivational Model*

Underlying Representations	/pkam/	/krmknawt/
Epenthesis_1	DNA	krɨ̱mknawt
Initial Stress	pkám	krɨ̱́mknawt
Epenthesis_2	pɨ̱kám	krɨ̱́mkɨ̱nawt
Surface Representations	[pɨ̱kám]	[krɨ̱́mkɨ̱nawt]

As stated in the introduction, the problem with the derivational approach can be summed up as follows: stress-epenthesis interaction as rule ordering requires the bifurcation of a unitary process of epenthesis. The epenthesis process itself yields a single structural change and is conditioned in essentially the same phonological environments, motivated as a means of syllabifying consonant clusters according to the phonotactics of the language (Foley 1991: 48). Therefore, the rule-ordering approach to stress-epenthesis interaction leads to loss of generalization because the observations which characterize the epenthesis process are stated more than once.

Stress-epenthesis interaction in Yimas shows that a unitary process has divergent effects in the stress system: the behavior of epenthetic vowels in the pattern of primary stress is nonuniform. Patterns of nonuniformity of this kind are well-studied phenomena within Optimality Theory,[7] and I will argue that nonuniformity in stress-epenthesis interaction receives a natural interpretation in parallelist OT. To give a brief sketch of this approach, let us review the schematic rankings employed in Section 2. The cases examined above were rather straightforward, involving essentially the ranking of a set of stress constraints relative to HEAD-DEP. Giving HEAD-DEP high rank in the system yields metrical inactivity of epenthesis, as for example in Dakota. On the other hand, by assigning HEAD-DEP low rank, the opposite result is obtained, as in the case of Swahili. In the more complicated case of Yimas, the intricacies of the system can be directly obtained by combining both of these schematic rankings. In particular, partial activity of epenthetic vowels in stress is derived by interleaving HEAD-DEP with the stress related constraints as shown below. In this ranking, C_{Stress} represents a set of constraints which yield a regular pattern of stress, and C'_{Stress} is the complement of that set.

7. See Prince (1993) for the first characterization of the phenomena in OT, and Prince & Smolensky (1993), Alderete (1995b), and Pater (in preparation) for discussion of a variety of examples.

(19) *Partial Metrical Activity of Epenthesis*
C_{Stress} » HEAD-DEP » C'_{Stress}

Partial activity in stress is therefore a direct consequence of the basic assumption in OT that grammars are modelled as a total ordering of constraints. One component of the constraint system derives the metrical inactivity of the epenthesis, namely HEAD-DEP » C'_{Stress}. However, an independent set of constraint rankings, i.e. C_{Stress} » HEAD-DEP, forces a violation of HEAD-DEP , with the effect of compelling metrical activity of epenthesis in just those contexts governed by C_{Stress}. This schematic ranking will be applied directly below in my analysis of Yimas.

Starting with the regular stress pattern itself, initial stress with alternating secondary stress on subsequent syllables diagnoses Yimas as a trochaic language. Hence RHTYPE = T is high ranking, relative to the analogous rhythm type constraint requiring iambic rhythm. Further, Foot Binarity is enforced at the level of the syllable, which together with high-ranking RHTYPE = T, yields a syllabic trochee. Further, the alignment constraint, ALIGN-L(F, PrWd), ensures left-to-right foot construction; iterative footing is accounted for by ranking the syllable-to-foot parsing constraint, PARSE-SYL, above ALIGN-L, which asserts that all prosodic feet coincide with the left edge (following McCarthy and Prince 1993a). In total, RHTYPE = T and Foot Binarity gives the syllabic trochee, and the alignment constraint, interacting with the syllable parsing constraint, conspire to yield left-to-right iterative foot construction.

With this component of the constraint system fleshed out, we can now move to the constraint rankings which account for noncanonical second syllable stress. Recall that this stress pattern correlates with initial epenthesis. Since the constraints deriving the canonical pattern posit the syllabic head of the main stress foot on the initial syllable, epenthesis into this position puts the stress constraints in conflict with HEAD-DEP. Hence, parsing the first two syllables as a syllabic trochee, as in the first candidate given below, violates HEAD-DEP for the head syllable. By ranking HEAD-DEP above the alignment constraint, therefore, the right result is obtained.

(20) *Metrical Inactivity of Epenthesis: /kcakk/ → kɨ̠cákɨ̠k*

kcakk	HEAD-DEP	ALIGN-L
{kɨ̠́ca} kɨ̠k	ɨ̠!	
☞kɨ̠ {cákɨ̠k}		kɨ̠

Because the epenthetic vowel *ɨ̱* has no input correspondent, parsing it as the head syllable of the word incurs a violation of the top-ranked HEAD-DEP. Therefore, the optimal form is the one which violates ALIGN-L as a means of satisfying HEAD-DEP. Furthermore, in disyllabic forms such as *pɨ̱kám*, HEAD-DEP compels a rhythm type reversal, by domination of RHTYPE = T. Thus, consistent with the approach taken to metrical inactivity of epenthesis in Section 2, failure to stress an epenthetic vowel is derived by ranking HEAD-DEP above stress-related constraints.

In one context, however, epenthetic vowels are recruited in stress assignment, indicating that HEAD-DEP is itself dominated in the system. As mentioned above, epenthesis into both the first and second syllable correlates with regular initial stress, e.g., *krɨ̱́mkɨ̱nawt*. While the system shows an general avoidance for stressing epenthetic vowels, it would seem that this imperative cannot compel post-peninitial stress because of a hard constraint enforcing a two syllable stress window. With the assumed trochaic foot structure, the stress window amounts to a general ban on two adjacent unfooted syllables, as defended by the scholars listed in (21).

(21) PARSE-SYL-2 (Kager 1994; Alderete 1995b; cf. Green and Kenstowicz 1995)
In adjacent syllables, avoid more than one unfooted syllable.

Failure to foot both the first and second syllable constitutes a violation of PARSE-SYL-2, and since laying down metrical structure has the effect of assigning stress, the complex syllable parsing constraint will suffice as our stress window constraint.

Applying this to the account of the stressed epenthetic vowels, all that is required is to rank the stress window constraint above HEAD-DEP, as illustrated in the following tableau.

(22) *Metrical Activity of Epenthesis: /krmknawt/ → krɨ̱́mkɨ̱nawt*

krmknawt	PARSE-SYL-2	HEAD-DEP	ALIGN-L
a. krɨ̱m{kɨ̱́.nawt}		ɨ̱	krɨ̱m!
b. krɨ̱m.kɨ̱ {náwt}	*!		krɨ̱m.kɨ̱
c. ☞{krɨ̱́m.kɨ̱} nawt		ɨ̱	

Among the candidates provided above, (22b) is not acceptable because by pushing main stress beyond the second syllable, the first two syllables are left unfooted, resulting a fatal violation of top-ranked PARSE-SYL-2. This leaves the two alternatives, (22a) and (22c), which tie on HEAD-DEP because both forms posit a syllabic head over a nonlexical vowel. The decision therefore falls to the low-ranking ALIGN-L, which chooses in favor of (22c) because it is perfect with respect to left-edge alignment.

Before moving on, it's worth considering the role of HEAD-DEP in the placement of secondary stress, and how the rankings employed thus far extend to this pattern. Secondary stress is assigned to every other syllable following the primary stressed syllable, which is the third syllable from the beginning of the word in most cases.[8] As noted by Foley (p. 77), this pattern of secondary stress is disrupted precisely when the third syllable is derived by epenthesis. In such forms, stress falls on the fourth syllable from the beginning of the word, as exemplified by the following data.

(23)	/tŋkmpɲawa/	tɨ̠́ŋkɨ̠mpɨ̠ɲàwa	'wild fowl'
	/kntkcki/	kɨ̠́ntɨ̠kɨ̠cɨ̠̀ki	'bird (sp)'

This shows us that the role of HEAD-DEP in the stress system is a rich one, extending beyond the exceptions to primary stress placement. That is, it seems that HEAD-DEP induces a noncanonical pattern of foot parsing by requiring an epenthetic syllable to be skipped when it would otherwise be stressed. The constraint ranking here simply involves ranking HEAD-DEP above the alignment constraint, ALIGN-L, which encourages all prosodic feet to be leftmost in the word. With HEAD-DEP top-ranked, the syllable containing the epenthetic vowel will be skipped in the assignment of the pattern of secondary stress, as shown in the following tableau.

(24) *Inactivity of Epenthesis in Secondary Stress: /tŋkmpɲawa/ →tɨ̠́ŋkɨ̠mpɨ̠ɲàwa*

tŋkmpɲawa	HEAD-DEP	ALIGN-L
{tɨ̠́ŋkɨ̠m} {pɨ̠̀ɲa}wa	ɨ̠!	**σ σ**
☞{tɨ̠́ŋkɨ̠m} pɨ̠ {ɲàwa}		**σ σ σ**

The analysis based on constraint interaction with HEAD-DEP, therefore, extends to the pattern of secondary stress in Yimas as well.

8. Secondary stress is only assigned in words greater than three syllables, providing further support for an analysis assuming a syllabic trochee with no mechanism for degenerate feet.

To review the basic components of the analysis, partial metrical activity of epenthesis is derived within the proposed ranking schema given at the outset of this section.

(25) *Stress-Epenthesis Interaction in Yimas*
PARSE-SYL-2 » HEAD-DEP » ALIGN-L, RHTYPE = T

The constraint rankings in which HEAD-DEP is in a dominating position yield the noncanonical pattern: in these rankings, HEAD-DEP compels violation of the stress constraints responsible for deriving regular initial stress and secondary stress on the third syllable. A different ranking in the system involves the domination of HEAD-DEP by the stress-window constraint, with the effect that epenthetic vowels are recruited in the assignment of stress just in forms which begin with two epenthetic syllables. In conclusion, the theory of stress-epenthesis interaction proposed here meets the challenge of deriving partial metrical activity of epenthesis in a rather straightforward way.

4. Summary and Implications

In this paper, a theory of stress-epenthesis interaction was developed which relies crucially on the notion of correspondence between inputs and related outputs and the OT assumptions that constraints are ranked and violable. The properties inherent to this theory were shown to have a set of advantages, which I will now summarize.

First, these properties of the theory permit a nonderivational treatment of stress-epenthesis interaction. Correspondence between related strings is essential in the formal characterization of HEAD-DEP, and by reranking this constraint in relation to the constraints governing stress, the range of different patterns are derived without the use of serial derivation. The theory is therefore consistent with recent research which outlines the strengths, both empirical and theoretical, of the parallelist approach (McCarthy 1993, 1996b; Benua 1997; Alderete 1995a; see Potter 1994 for a different view).

The second advantage of the approach taken here is an empirical one, stemming from the principles of OT. The properties of constraint ranking and domination were used effectively in the analysis of Yimas, which exemplified the common pattern of partial metrical activity of epenthesis. The mixed behavior of epenthesis in this case was handled straightforwardly by ranking HEAD-DEP both above and below the constraints deriving regular stress. The important point is that the OT approach contrasts sharply with the rule-based theory, which was shown to lead to loss of generalization in the characterization of the epenthesis process itself.

The theory developed here also has a theoretical advantage over plausible alternatives in that it paves the way for relating patterns of stress-epenthesis interaction to other phenomena, namely segmental processes like metrically-conditioned vowel reduction. In the analysis of Dakota, noncanonical stress pattern is the result of ranking HEAD-DEP above a constraint which derives a regular pattern of stress. Resistance to vowel reduction in stressed syllables is derived in a parallel fashion by employing the related constraint HEAD-IDENT[F]. In both cases, a head-sensitive faithfulness constraint is given high rank in the constraint hierarchy, with the effect of suppressing regular phonological patterns. The theory of faithfulness to prosodic heads therefore covers considerable empirical ground, with very limited resources.

Before closing, it is worth mentioning some related work that has surfaced in the past years. In the above analyses, two classes of faithfulness constraints are specified for a prosodic target, i.e., the anti-epenthesis constraint, DEP, and the featural faithfulness constraint, IDENT[F]. A number of researchers have modified other faithfulness constraints along similar lines, and I will review them briefly here as a way of sketching some further implications of this basic idea.

One interesting application of position-sensitive faithfulness is developed in Kager (this volume) as an account of the resistance of stressed vowels to undergo syncope. Rather than characterizing vowel deletion as a process that specifically operates on unstressed vowels, a more general account of syncope is given in this work, with the stability of stressed vowels explained as an effect of a position-sensitive HEAD-MAX constraint. Thus, consistent with the approach taken here to vowel reduction, the resistance on the part of stressed vowels to undergo a regular linguistic process is handled with high-ranking head faithfulness.

A second modification of the MAX family of constraints is explored in Beckman (1998) in an account of various syllable structure asymmetries. Different from Kager's HEAD-MAX, Beckman formulates a constraint MAX-*Position*, which requires that all underlying segments be realized in a prominent position in the output. The result of this constraint, when it is properly ranked in relation to the constraints governing syllable shape, is that certain prominent syllables may license a wider range of syllable shapes than the complement set of syllables in a word.

A third application of position-sensitive MAX is given in Yip (1996) for the distribution of tone in Chinese dialects. Developing a strong parallel to the case of Russian examined above, the preference for certain tonal units to dock with a head syllable are explained with a HEAD-MAX[F] constraint for tonal features.

Thus, the important innovation here is that Yip's theory completes a positional faithfulness family in input-to-output correspondence by proposing a MAX-type constraint and employing it as a featural faithfulness constraint.

A different form of faithfulness, originally proposed in McCarthy (1996b), involves 'matching' of correspondent segments in head positions. McCarthy's HEAD-MATCH asserts that if two segments stand in correspondence, and one segment is in a prosodic head, then its correspondent must be in a head as well. Variants of this constraint have been used in nonderivational approaches to prosodic circumscription effects (McCarthy 1996b; McCarthy, to appear; Itô, Kitagawa, and Mester 1996) and in the characterization of faithfulness to underlying prosody in lexical stress systems (McCarthy 1996b; Alderete 1996; Pater in preparation).

Acknowledgments

Thanks to the participants at the Rutgers/UMass Joint Class Meeting and the Tilburg conference on the Derivational Residue in Phonology for helpful comments and suggestions. This paper has also benefited from conversations and correspondence with John McCarthy, Lisa Selkirk, Roger Higgins, René Kager, Marc van Oostendorp, Ben Hermans, and Dale Hartkemeyer. I am responsible for any errors in fact or interpretation. This work was supported in part by the National Science Foundation grant SBR-9420424.

Unrecoverable Origins

Mary M. Bradshaw

1. Introduction

It has generally been claimed that ranked constraint analyses can handle data as well as derivational analyses can, but without the use of rule ordering. However, some data remain intractable to analysis within an OT-type framework precisely because ordering is crucial in arriving at surface forms. One kind of ordering that is problematic is counterfeeding order. An example from Chukchee will illustrate this problem. In Chukchee, there is a process by which preconsonantal *k* becomes *γ*, as shown in (1a). By another process, *γ* becomes *w* before a labial, as shown in (1b). These processes are in a counterfeeding order with *w*-formation preceding *γ*-formation. Therefore underlying /*km*/ becomes [*γm*] instead of [**wm*], as shown in (1c). *wm* would be expected if a feeding relationship existed.

(1) *Chukchee (Krause 1979)*
a. /kC/ → [ɣC][1]
b. /ɣ Lab/ → [w Lab]
c. /km/ → [ɣm] (*wm)

These facts might be handled in an OT-type analysis through the use of two constraints: No Preconsonantal *k* (NO*k*C) and No Sequence of a Velar Continuant before a Labial (NoVelarCont-Lab). The Tableau in (2) shows that the incorrect form **wm* would be chosen as optimal rather than the attested ɣ*m* form.

1. Forms like *ekak* ‘son’ make it clear that this is not a constraint against coda *k*.

(2)

/km/	NO*k*C	NOVELARCONT-LAB
km	*!	
ɣm		*!
☞*wm		

The problem posed by the counterfeeding processes in Chukchee can be handled by distinguishing between underlying and derived properties. Thus, if the constraint against a sequence consisting of a velar continuant before a labial is modified such that it is specifically a velar with an *underlying* continuant feature that cannot occur before a labial, then *ɣm* in the tableau no longer incurs a violation of any relevant constraint.

The solution of distinguishing derived properties from underlying properties in the application of constraints will handle some, but not all, of the data which is analyzable in terms of counterfeeding order in a derivational framework. That is, there is a residue of data which relies on other distinctions. For example, counterfeeding can involve data with properties derived by one process which must be distinguished from properties derived by another process. An OT-type analysis cannot make reference to properties of output forms which are present to satisfy one constraint vs. those which are present to satisfy a different one. Suma, an Ubangi language from the Central African Republic, is a language in which such a distinction becomes necessary in deriving the correct tones for head nouns in the associative construction.

The tone system of Suma (following Bradshaw 1995) consists of 3 tones which result from the interaction of 2 binary tonal features: [upper register] and [raised pitch]. A H tone can be described as [+upper, +raised]; a M tone as [+upper, –raised] and a L tone as [–upper, –raised], as in (3).

(3) *Tone System (from Bradshaw 1995)*

	H	M	L
upper	+	+	–
raised	+	–	–

When a tonal feature is unspecified, default rules, given in (4), fill in [–upper] and/or [–raised].

(4) *Default Rules*
[ ] → –raised
[ ] → –upper

In keeping with these rules, L is the default tone.

2. Associative Construction

The associative construction consists of a head noun and a following complement in a noun phrase. It generally can be translated as ‘X of Y’ in phrases like ‘mouth of Bonam’ and ‘basket of manioc’. Among other things, it is used for inalienable possession and in questions with *gèé* ‘which’.

The tone pattern of a noun in Suma changes in a systematic way in the associative construction due to the addition of the associative morpheme which consists of a floating tone feature. Only head nouns with a final L are affected by these tone changes, because these nouns have a final toneless TBU before the filling in of default values. Head nouns with final H or M tones do not alternate, as shown in (5) for H-toned nouns and in (6) for M-toned nouns. (Underlining is used to mark nasalization on vowels.)

(5)	nú̱	‘mouth’	nú̱ Bōnà̱m	‘Bonam’s mouth’
	zí̱	‘fly’	zí̱ gɔ̱̀rɔ̱̀	‘bee (honey fly)’
	yéré	‘basket’	yéré gɛ̀ɗà	‘basket of manioc’
	sárá	‘spine’	sárá nú̱	‘whiskers (of mouth)’
	zàká	‘horn’	zàká sàɗè	‘animal horn’
	běy	‘person’	běy àlà	‘person of suffering’
(6)	mbārā	‘money’	mbārā zòrò	‘price of fish’
	gūn	‘waist’	gūn běy	‘waist of person’
	kōy	‘remains’	kōy kàm	‘leftover food’
	dɔ̱̄ɔ̱̄	‘wine’	dɔ̱̄ɔ̱̄ mbúrú	‘palm wine’
	mbōŋgō	‘top of head’	mbōŋgō tān běy	‘top of person’s head’

2.1 *Tone Alternations in Toneless Nouns*

When the final tone is L in isolation, it is realized as a M tone in the associative construction, as shown in (7). The associative morpheme docks to the final toneless TBU and spreads leftwards.

(7)	rì	‘water’	rī bèrè	‘breast milk’
	bèrè	‘breast’	bērē bágàrà	‘cow’s teat’
	wɔ̱̀	‘hunger’	wɔ̱̄ sàɗè	‘hunger for meat’
	sàà	‘game’	sāā bêm	‘children’s game’
	ɓàà	‘arm’	ɓāā běy	‘person’s arms’
	tàn	‘head’	tān Bōnà̱m	‘Bonam’s head’
	tɛ̱̀n	‘thorn’	tɛ̱̄n tè	‘tree thorn’
	nì̱k	‘tendons’	nī̱k gɛ̀ɛ́	‘tendons of neck’

zà̱ŋ	'insides'	zā̱ŋ Bōnà̱m	'Bonam's insides'
mà̱rà̱	'tail'	mā̱rā̱ sàɗê	'animal's tail'
gbà̱rà̱	'hard part'	gbā̱rā̱ kùrà	'arrow-head'

In (8), toneless nouns are shown in the associative construction in phrases using 'which'. Again, there is a surface alternation between L and M.

(8)	ɔ́ nɛ̱̄ rī gèé	'which water is it?'	from rì	'water'
	ɔ́ nɛ̱̄ wēn gèé	'which dispute is it?'	from wèn	'word, dispute'
	ɔ́ nɛ̱̀ bērē gèé	'which breast is it?'	from bèrè	'breast'
	ɔ́ nɛ̱̄ dūwā gèé	'which goat is it?'	from dùwà	'goat'
	ɔ́ nɛ̱̄ gɔ̱̄rɔ̱̄ gèé	'which honey is it?'	from gɔ̱̀rɔ̱̀	'honey'

The tone alternations in the associative construction differ depending on the mora count of the target noun. Toneless nouns of 2 moras, as seen in (7) and (8), surface with M tones on every mora. Toneless nouns longer than 2 moras, shown in (9) and (10), are also characterized by the leftward spread of a M tone, but they differ from shorter nouns in that the spreading does not target the word-initial mora. In other words, only non-initial L alternates with M in nouns longer than 2 moras. Toneless nouns with 3 moras (and the tone pattern LLL in isolation) surface as LMM in the associative construction, as in (9).

(9) LLL → LMM (*MMM)

yàwùndù	yàwūndū gèé	'which red peanut?'
sɛ̀rɛ̀ŋgɛ̀	sɛ̀rɛ̄ŋgɛ̀ gèé	'which beefeater bird?'

Toneless nouns with 4 moras (and all L tones in isolation) surface with an initial L followed by M tones in the associative construction, as in (10).

(10) LLLL → LMMM (*MMMM)

yàmàndìyò	yàmāndīyō gèé	'which red peanut?'
dɔ̀kɔ̀ɗ ìlè	dɔ̀kɔ̄ɗ īlē gèé	'which caterpillar?'
à̱kà̱dèrè	à̱kā̱dērē gèé	'which toad?'

Depending on the number of moras in the noun, the first vowel of the toneless noun may be realized with a L tone or with a M tone. In order to capture this distinction without resorting to adhoc mora counting, foot structure is assumed in which a right-headed binary foot is located at the left edge of the noun.

2.2 *Tone alternation in nouns with an underlying H*

Nouns with a final toneless mora preceded by a mora with a H tone also alternate. When these nouns are bimoraic, only a HL pattern is possible in isolation. Their tone alternates with HH in the associative construction, as in (11) and (12). Thus, they can be characterized by the rightward spreading of H.

(11)	kpá̱nà̱	'pot, jar'	kpá̱ná̱ rì	'water jar'
	kpárɛ̀	'seeds for sowing'	kpárɛ́ fón	'millet seeds'
	ɓólò	'pouch'	ɓóló ná̱ŋá̱	'animal skin pouch'
	kúrì	'egg'	kúrí gɔ̌k	'serpent's egg'
	ndîŋ	'dirt, filth'	ndíŋ wèn	'foul language'
	súɓè	'stomach contents'	súɓé ndārā	'contents of buffalo's stomach'

(12)	ɔ́ nɛ̱̄ kpá̱ná̱ gèé	'Which pot is it?'
	ɔ́ nɛ̱̄ kpárɛ́ gèé	'Which sowing seeds are they?'

When longer nouns with final L preceded by a H alternate tonally in the associative construction, H shows no evidence of spreading rightward. In (13), the LHL tone pattern alternates with LHM rather than with LHH. In (14), the HLL pattern alternates with HMM rather than HHH. In (15), the HHL pattern alternates with HHM.

(13)	LHL → LHM (*LHH)		
	ŋgàálà		'ctn. plant'
	ŋgàálā gèé		'which plant?'
	sìílì		'ctn. insectivore'
	sìílī gèé		'which insectivore?'
	sùmá̱rì̱	(from Banda)	'secret society'
	sùmá̱rī̱ gèé		'which secret society?'
	wárá zɔ̀ɔ́rɔ̀		'fruit of *zɔ̀ɔ́rɔ̀* tree'
	wárá zɔ̀ɔ́rɔ̄ gèé		'which fig?'

(14)	HLL → HMM (*HHH)	
	yá̱kɔ̀ɓà	'Ascaris, the intestinal worm'
	yá̱kɔ̄ɓā gèé	'which Ascaris?'
	bágàrà	'cow'
	bágārā gèé	'which cow?'
	ná̱kà̱rà̱	'shoe'
	ná̱kā̱rā̱ gèé	'which shoe?'
	pá̱pà̱nà̱	'amniotic fluid'
	pá̱pā̱nā̱ gèé	'which amniotic fluid?

(15) HHL → HHM (*HHH)
kpéɗésè ‘squirrel’
kpédésē zân ‘bush squirrel’

The tone alternations in Suma in the associative construction are summarized in (16). The final TBU, if it is L in isolation, always becomes M except in the case where it is preceded by an initial H. Initial L only becomes M in words of 2 moras or less. Although the shorter nouns seem to deviate from the overall pattern, they are far more typical of Suma than the longer nouns, since the overwhelming majority of nouns are bimoraic.

(16)

L → M	HL → HH
LL → MM	HLL → HMM
LLL → LMM	LHL → LHM
LLLL → LMMM	HHL → HHM

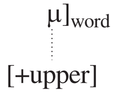
3. Analysis of [+upper] Spreading

The pattern characterized by the leftward spread of M in toneless nouns in the associative construction can be analyzed in both a derivational and a ranked constraint analysis. The associative morpheme, which is realized as a M tone, is assumed to be the floating feature [+upper].

3.1 *Derivational Analysis of [+upper] Spreading*

Derivationally, the analysis of the spreading of the associative morpheme can be done with four processes, which target only toneless moras. When the final TBU of the head noun is unspecified for tone, the associative morpheme becomes linked to it by a process of Upper Docking, illustrated in (17). This accounts for the fact that in the associative construction, [+upper] always appears at the right edge of the target word.

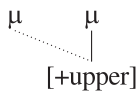
(17) *Upper Docking*



The operation of Upper Docking alone accounts for the alternation of L and M in monomoraic words, such as *rì* → *rī bèrè* ‘breast milk’.

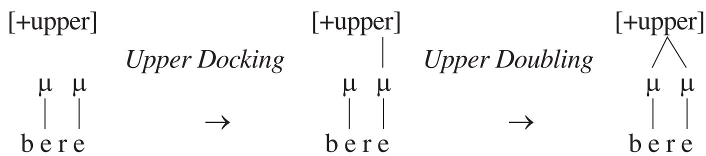
The rule accounting for the spread of the associative morpheme is Upper Doubling, given in (18). After [+upper] has docked by Upper Docking (17), it spreads leftward to an adjacent mora unspecified for tone. This process corresponds to the generalization that the associative morpheme always spreads at least once if an unspecified TBU is available.

(18) *Upper Doubling*



The processes of Upper Docking and Upper Doubling can account for the alternations seen in bimoraic toneless nouns, as illustrated in (19).

(19) /bere+ASSOC geé/ 'which breast' → [bērē gèé]



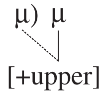
In order to account for longer nouns in the associative construction, two more processes are needed: Foot Construction and Upper Spread. Foot Construction has been justified by the need to distinguish nouns of two moras from longer nouns. By this process, shown in (20), a right-headed bimoraic foot is constructed at the left edge of the word. Foot Construction need not be extrinsically ordered with respect to Upper Docking and Upper Doubling. However, it is crucially ordered before the final step in the derivational analysis, Upper Spread. Later, it will be seen that a foot is also necessary in an account of [+raised] spreading in nouns where a H precedes the final toneless mora.

(20) *Foot Construction*

$_{word}$[(μ μ)

Finally, the process of Upper Spread, given in (21), consists of the spreading of [+upper] as far as the head of the foot. It is the mora at the right edge of the foot that constitutes its head. The attraction of tone features to a prominent metrical position has been observed in other Niger-Congo languages, such as the Nguni languages (Downing 1990) and Tanzanian Yao (Odden 1998). It has also been pointed out by Van der Hulst and Smith (1988) for a number of languages.

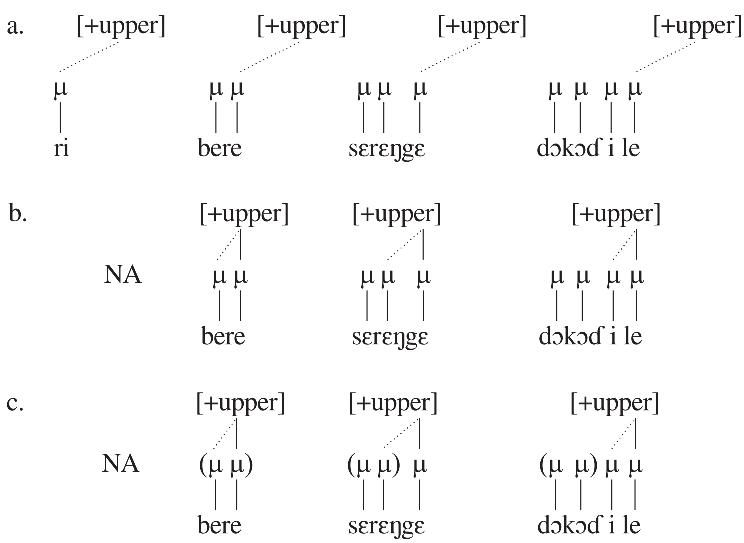
(21) *Upper Spread*



Two spreading rules, Upper Doubling (18) and Upper Spread (21) are needed to account for the fact that bimoraic nouns have an initial M in the associative construction, while longer nouns do not. If Upper Doubling is ordered before Upper Spread, then Upper Doubling will ensure that [+upper] spreads once, regardless of the location of the head of the foot. Upper Spread will ensure that if [+upper] spreads more than once, it is sensitive to the location of the head of the foot. With only one process of spreading, the distinction between bimoraic nouns and nouns of 4 moras would not be captured, as is clear in the derivations in (22).

The four processes discussed above will derive the correct surface forms of underlyingly toneless nouns, as shown in (22), where derivations for nouns having from 1 to 4 moras are given. In (22a), the associative morpheme, a floating specification of [+upper], docks. In (22b), [+upper] spreads once by Upper Doubling. In (22c), a foot is constructed. In (22d), [+upper] spreads left to the head of the foot by Upper Spread.

(22)



d.

				[+upper]
	NA	NA	NA	(μ μ) μ μ
				dɔkɔɗ i le
SR	rī	bērē	sɛ̀rɛ̄ŋgɛ̄	dɔ̀kɔ̄ɗīlē
	'water'	'breast'	'beefeater bird'	'caterpillar'

3.2 *Ranked Constraint Analysis of [+upper] Spreading*

The ranked constraint analysis of the leftward spreading of the associative morpheme relies on five constraints, not counting the constraints needed to get foot structure. With the exception of PARSE[FEATURE] all constraints will refer to grammatical tone features to the exclusion of lexical tone features. Again we must account for the observation that the associative morpheme is realized in shorter nouns on both TBU's and in longer nouns on all, and only, noninitial TBU's. The constraints needed to capture these generalizations include PARSE[FEATURE] and an alignment constraint that plays a role in determining on which mora(s) [+upper] is realized. These two constraints correspond to Upper Docking in the derivational analysis. PARSE[FEATURE], given in (23), requires that all features be parsed. This ensures that the associative morpheme will be parsed to a toneless head noun.

(23) PARSE[F]
All features must be parsed.

In the Tableau in (24), the candidate in which the associative morpheme is unparsed is eliminated in favor of the candidate with a parsed associative morpheme. The optimal candidate in this Tableau surfaces with a M tone because the feature [−raised] is filled in by default. The reader will recall that [+upper, −raised] specifies a M tone.

(24)

ri + [+upper]	PARSE[F]
[+upper] ☞ | ri	
[+upper] ri	*!

The alignment constraint, referred to as ALIGN[U]-R in (25), captures the generalization that all floating [+upper] features are associated to the rightmost TBU. It is violated when [+upper] is not aligned to the right edge of a word.

(25) ALIGN[U]-R
ALIGN([+upper],R,word,R): Align [+upper] to the right edge of a word.

The Tableau in (26) demonstrates a violation of ALIGN[U]-R for *bere* 'breast'. The M tones in (26) represent the feature [+upper], which is multiply linked to the moras of the vowels in the first candidate.

(26)

bere + [+upper]	ALIGN[U]-R
☞bērē	
bēre	*!

Another constraint must account for the doubling phenomenon, whereby bimoraic toneless nouns in the associative construction realize a M tone on the initial TBU (cf. *bērē* 'breast') in contrast to longer nouns which do not (cf. *dɔ̀kɔ̱̄dīlē* 'caterpillar'). This constraint, MINIMAL TONE ASSOCIATION or MTA, penalizes a singly linked [+upper] specification. The version of MTA in (27) and (28), which states that [+upper] must be multiply linked, is a variation of the constraint proposed in Poletto (1998).

(27) MINIMAL TONE ASSOCIATION
[+upper] must be linked to more than one TBU (from Poletto 1998).

A formal representation of MTA is given in (28).

(28) *[+upper]
|
μ

Violations of MTA, ALIGN[U]-R and PARSE[F] for a bimoraic noun are seen in (29). The attested surface form violates none of them. Thus, for a bimoraic toneless noun in the associative construction, the violation of any of these constraints is fatal. It will be noted that an unparsed tone feature is assumed not to incur violations of constraints that refer to associations.

(29)

bere + [+upper]	PARSE[F]	ALIGN[U]-R	MTA
☞bērē			
bēre		*!	*!
berē			*!
bere	*!		

Although no rankings of these constraints can be established with reference to bimoraic nouns, monomoraic nouns demonstrate that MTA is ranked below PARSE[F]. The optimal candidate in (30) violates MTA while satisfying the other constraints. Since ALIGN[U]-R is never violated by grammatical tones and is therefore undominated[2], it is also assumed to rank above MTA (and above PARSE[F]), although there is never a direct conflict between the two[3].

2. It is not entirely clear that because a constraint is unviolated, it is highly ranked. That is, if it accidentally happens that there is another constraint that affects the same set of words that the constraint in question affects and the former constraint is highly ranked, then the latter constraint need not be. In this case, there is no evidence for a second constraint that accidentally affects the same set of forms. Furthermore, this constraint ranking is not crucial for my point.

3. PARSE[F] does not qualify as undominated. It is violated when a noun does not end in a toneless mora underlyingly. It is not possible to show that a noun with a HH, MM or HM does not exhibit docking of the associative [+upper] morpheme. However, if PARSE[F] forced 'docking', a LH noun is expected to surface as MH because of MTA. This is because a grammatical H tone must be multiply linked. A LH form would violate this constraint while a MH noun would not incur this violation, nor any others. However, it is a LH tone that surfaces in the associative. This is evidence that the associative morpheme has not docked.

(30)

ri + [+upper]	ALIGN[U]-R	PARSE[F]	MTA
☞rī			*
ri		*!	

The constraints given so far will account for the attested surface forms of monomoraic and bimoraic nouns in the associative construction. They will not account for the surface forms of longer nouns in which [+upper] can be associated to more than two TBU's and in which the initial TBU is left toneless. Another constraint is needed to account for spreading further to the left. ALIGN[U]-HD, given in (31), states that the left edge of [+upper] aligns to the left edge of the head mora of a foot. Since feet are rightheaded in Suma, this entails that the edge of [+upper] is optimally aligned to the rightmost mora of a foot.

(31) ALIGN[U]-HD
ALIGN([+upper],L,$head_{foot}$,L) (Align the left edge of [+upper] to the left edge of the head mora of a foot.

This constraint interacts with the others proposed in this section to select the optimal candidate in the case of toneless nouns with four moras, as shown in (32). All possible combinations of L and M tones are considered in the tableau, except those in (33) which are eliminated by a constraint against gapped constructions.

(32)

dɔkɔɗile + M	ALIGN[U]-R	PARSE[F]	MTA	ALIGN[U]-HD
☞(dɔkɔ̄)ɗ īlē				
(dɔkɔ)ɗ īlē				*!
(dɔ̄kɔ̄)ɗ īlē				*!
(dɔkɔ)ɗ ilē			*!	*
(dɔkɔ̄)ɗ īle	*!			
(dɔ̄kɔ̄)ɗ īle	*!			*
(dɔ̄kɔ̄)ɗ ile	*!*			*
(dɔkɔ̄)ɗ ile	*!*		*	
(dɔkɔ)ɗ īle	*!		*	*
(dɔ̄kɔ)ɗ ile	*!**		*	*
(dɔkɔ)ɗ ile		*!		

(33) dɔkɔ̄ɗilē dɔ̄kɔɗilē dɔ̄kɔɗ īle dɔ̄kɔɗ īlē dɔ̄kɔ̄ɗilē

ALIGN[U]-HD is ranked below PARSE[F], ALIGN[U]-R and MTA, as shown in (34), because the optimal candidate in bimoraic nouns will violate it. Both MTA and PARSE[F] are shown to cause fatal violations in contrast to ALIGN[U]-HD, and it has already been mentioned that ALIGN[U]-R is unviolated.

(34)

bere + [+upper]	ALIGN[U]-R	PARSE[F]	MTA	ALIGN[U]-HD
☞(bērē)				*
(berē)			*!	
(bēre)	*!		*!	*
(bere)		*!		

The crucial rankings which have been argued for are listed in (35).

(35) Crucial Rankings: PARSE[F] » ALIGN[U]-HD
MTA » ALIGN[U]-HD
PARSE[F] » MTA
ALIGN[U]-R undominated

Another way to represent these rankings is presented in (36).

(36) Summary: ALIGN[U]-R_G
PARSE[F]
|
MTA
|
ALIGN[U]-HD

In this section, the spread of the associative morpheme has been analyzed in both a derivational and ranked constraint framework, revealing few advantages or disadvantages in either approach. But these analyses are necessary background for the more problematic analysis of [+raised] spreading, which involves counterfeeding. In the next section, we will see how one might try to account for [+raised] spreading in both approaches.

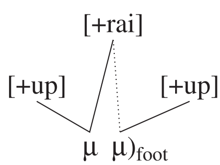
4. Analysis of [+raised] Spreading

In this section, the spreading of [+raised] in bimoraic nouns and its failure to spread in longer nouns (cf. *ná̱kā̱rā̱* 'shoe' vs. *kpá̱ná̱* 'pot') will be explored in a derivational and a ranked constraint analysis. The spreading of [+raised] interacts with the spreading of the associative morpheme, since [+raised] spreads when the associative morpheme has docked to the head noun in the associative construction. This [+raised] spreading occurs only in bimoraic nouns with the tone pattern HL in isolation, and not with longer nouns in which the same environment seems to occur. That is, it does not occur with nouns of the tone pattern HLL, LHL or HHL. If [+raised] spreading targets the head of the foot, we can account for its failure in longer nouns.

4.1 *Derivational Analysis of [+raised] Spreading*

Within a derivational analysis, the rightward spread of [+raised] can be accounted for by a process of Raised Spread, formulated in (37). Accordingly, [+raised] spreads rightwards to an adjacent TBU with a [+upper] specification which forms the head of a foot.

(37)

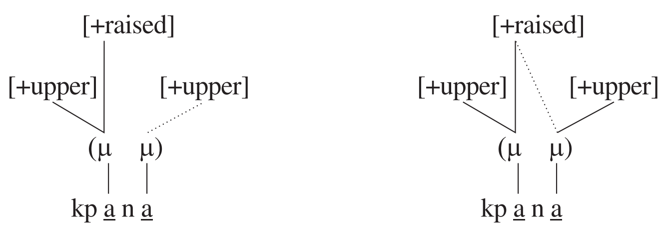


In the case of bimoraic nouns, Raised Spread operates after foot construction and Upper Docking to produce HH nouns in the associative construction, as in (38).

(38) /kpá̱na̱ + ASSOC geé/ 'which jar' → [kpá̱ná̱ gèé]

Foot Construction *Raised Spread*

Upper Docking

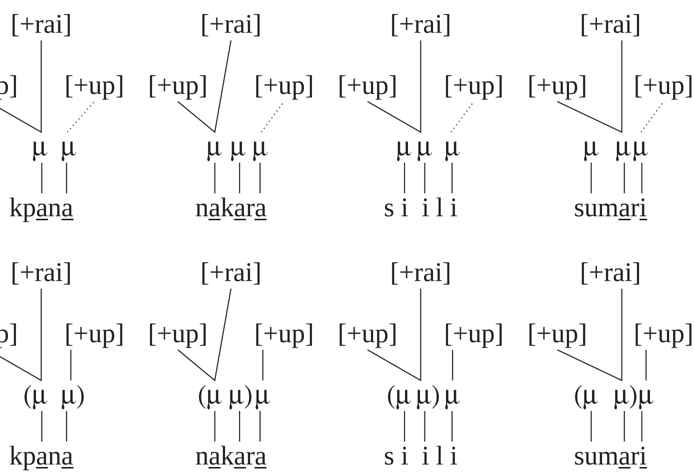


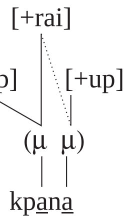
The steps in the derivation of nouns with lexical tone patterns in which a final toneless mora is preceded by a H tone, given in (39), involve the processes of Upper Docking, Foot Construction, Raised Spread and Upper Doubling.

(39) a Upper Docking (17): Dock the Associative Morpheme.
 b. Foot Construction (20): Construct a rightheaded binary foot on L edge.
 c. Raised Spread (37): Spread [+raised] rightwards once to the foothead.
 d. Upper Doubling (18): Spread [+upper] leftwards once.

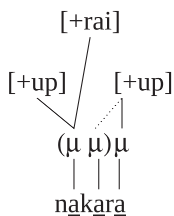
The derivations in (40) show long and short nouns which are HL, HLL, and LHL in isolation. In (a), the associative morpheme docks. In (b), a foot is constructed at the left edge of the noun. In (c), Raised Spread occurs only in the bimoraic noun because that is the only noun in which a [+upper] feature is associated to an adjacent TBU which forms the head of a foot. This underscores the importance of the prosodic foot in this analysis. [+raised] fails to spread to a TBU with no tone features, such as in *ná̱kā̱rā̱*. In (d), Upper Doubling occurs in the noun which is HLL in isolation. The attested tone patterns are derived in all cases.

(40) a.
b.



c. 

NA NA NA

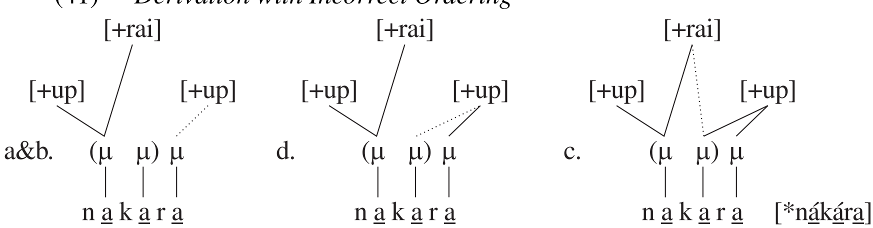
d. NA 

NA NA

SR:	kpá̱ná̱	ná̱kā̱rā̱	sì ílī	sùmá̱rī
	'pot'	'shoe'	'insectivore'	'secret society'

The crucial orderings involved in this derivation are that Upper Docking and Foot Construction must precede Raised Spread, since they provide the structural description of the process. Furthermore, Raised Spread must precede Upper Doubling in order to account for the failure of Raised Spread in HLL nouns like *ná̱kā̱rā̱*. This is the counterfeeding order which must be dealt with. On the surface, a mora with a [+raised] specification is followed by one with only a [+upper] specification, which conforms to the structural description of Raised Spread. But Raised Spread fails to occur. If Raised Spread (40c) were not ordered before Upper Doubling (40d), the attested pattern would not be derived for *ná̱kā̱rā̱*. Instead of the attested pattern HMM, the tone pattern HHM would be derived, as shown in (41). The incorrect ordering in (41) involves Upper Doubling in a feeding relationship with Raised Spread.

(41) *Derivation with Incorrect Ordering*



In the derivations given so far, the need for foot construction has been demonstrated. Justification for constructing the foot on the left edge rather than the right edge is provided in (42). The first and second rows differ according to whether the foot is constructed on the right or left edge. For both rows, this is the point in the derivation represented in (40b) where the foot has been constructed and the associative morpheme is docked, but doubling and spreading have not yet taken place. Since all three tone patterns represented by the words in (42) surface with a different tone pattern in the associative construction, it is expected that at some point in the derivation, they will have a different pattern. In the first row, the tone patterns within the foot are all distinctive which provides a basis for the distinctive surface representations. But when the foot is constructed on the right edge, as in the second row, the tone patterns of the first and second nouns are the same within the foot, and there is no principled way to get the contrastive surface tones from this intermediate stage.

(42) Left edge (kpá̱nā̱) (sumá̱)rī̱ (ná̱ka̱)rā̱
Right edge (kpá̱nā̱) su(má̱rī̱) ná̱(ka̱rā̱)

In this section, a derivational analysis has been presented which can account for the attested surface tones in a straightforward manner. This analysis requires the construction of a foot on the left edge of the noun and the crucial ordering of Raised Spread and Upper Doubling. The next section will cover the same data from a ranked constraint perspective.

4.2 *Ranked Constraint Analysis of [+raised] Spreading*

A ranked constraint analysis of tonal alternations in the associative construction involving [+raised] will be attempted in this section. The imposition of foot structure needed in the derivational account is also necessary in a ranked constraint analysis. The constraints already proposed to account for [+upper] spreading will be useful in an account of [+raised] spreading, but a new constraint is also needed in order to capture the generalization that in bimoraic HL nouns, [+raised] spreads to the rightmost TBU in the associative construction. ALIGN[R]-R in (43) will capture this generalization by aligning the right edge of [+raised] to the right edge of a foot.

(43) ALIGN[R]-R
ALIGN([+raised],R,Foot,R)[4]: Align the right edge of [+raised] to the right edge of a foot.

The Tableau in (44) shows a violation of this constraint by the nonoptimal candidates in which [+raised] is not linked to the TBU at the right edge of the foot. PARSE[F] is also violated by one of the other reasonable candidates. Thus, the constraints ALIGN[R]-R and PARSE[F] are the only constraints needed (other than those necessary for the foot itself) to select the candidate with the optimal tone pattern for a bimoraic noun with the tone pattern HL in isolation.

(44)

H | kpa̲na̲ + M	Align[R]-R	PARSE[F]
☞(kpá̲ná̲)		
(kpá̲nā̲)	*!	
(kpá̲na̲)	*!	*

No crucial rankings are revealed by the tableau in (44). Consideration of a noun with the tone pattern LHM in isolation in (45) indicates that MTA is violated by the winning candidate. We already know that ALIGN[U]-R and PARSE[F] are ranked above MTA, but it is still not clear where ALIGN[R]-R would be ranked.

(45)

H | suma̲ri̲ + M	ALIGN[U]-R	ALIGN[R]-R	PARSE[F]	MTA
☞(sumá̲)rī̲				*
(sumá̲)rí̲		*!		*
(sumá̲)ri̲			*!	
(sūmá̲)ri̲	*!*			*

4. This causes an apparent problem for the isolation form of HL nouns, but this problem is handled by the constraint in (48), ALIGN[RAISED],[UPPER].

In considering whether to rank ALIGN[R]-R very highly, the question of whether or not it is undominated is pertinent. A look at HLL nouns reveals that it can be violated in output forms. For example, *ná̠kà̠rà̠* ‘shoe’ becomes *ná̠kā̠rā̠* in the associative construction. In order to satisfy ALIGN[R]-R, the incorrect output **ná̠ká̠rā̠* would be expected.

A summary of the constraints relevant for [+raised] spreading and their crucial rankings in accounting for the data considered so far is provided in (46).

(46) ALIGN[U]-R
|
PARSE[F], ALIGN[R]-R
|
MTA

There is a problem, however, when the constraints that have served us so far are applied to HLL nouns, where it will be recalled, there is a counterfeeding order between Raised Spread and Upper Doubling in the derivational analysis. The Tableau in (47) shows that the incorrect tone pattern HHM, marked with a reversed pointing finger ☜, is technically preferred over the attested pattern HMM. The attested pattern violates ALIGN[R]-R while the technically optimal pattern violates nothing. It is not possible to eliminate the constraint ALIGN[R]-R since it constitutes the only motivation for the rightward spread of [+raised] in HL nouns and is necessary in the selection of the optimal candidate in these nouns (see *kpá̠ná̠* ‘pot’ in (44)), as well as in LHL nouns (see *sùmá̠rī̠* ‘secret society’ in (45)). Thus, the constraints that have been established for other patterns involving [+raised] in the associative construction do not result in the selection of the attested form as optimal for HLL nouns.

(47)

H | na̱ka̱ra̱ + M	ALIGN[U]-R	ALIGN[R]-R	PARSE[F]	MTA
☞(ná̱kā̱)rā̱		*!		
☜(ná̱ká̱)rā̱				
(ná̱ka̱)rā̱		*!		*
(ná̱kā̱)ra̱	*!	*		*
(ná̱ká̱)rá̱		*!		
(ná̱ká̱)ra̱	*!			*
(ná̱ka̱)ra̱		*!	*	

Alternative foot structure cannot solve the problem. If the foot is constructed on the right edge, incorrect forms of LHL nouns, such as **sùmá̱rí̱,* will be selected, as indicated in the previous section.

We saw at the outset of this paper that in Chuckchee reference to the contrast between derived and underlying elements can be used within an OT-type analysis to deal with a case that the derivational approach handles in terms of the counterfeeding ordering of rules. One might, therefore, think to account for the counterfeeding relationship between Raised Spread and Upper Doubling by reference to derived vs. underlying tonal associations. In this case, such an analysis does not provide a solution. Recall that in HL nouns in the associative construction [+raised] spreads, while in HLL nouns [+raised] fails to spread. In both cases, the presence of [+upper] on the mora that serves as the potential target of [+raised] spreading is derived. In the HL nouns, where [+raised] spreads, the presence of [+upper] on the vowel following [+raised] is derived by Upper Docking. In the HLL nouns, where [+raised] fails to spread, the presence of [+upper] on the vowel following [+raised] is derived by Upper Doubling. In terms of constraints, the [+upper] in question is positioned on the potential target of [+raised] spreading because of PARSE[F] in HL nouns and because of MTA in HLL nouns. Thus, the crucial distinction is not between derived and underived. Rather it is a distinction based on the source of the [+upper] specification — the process or constraint that resulted in the positioning of [+upper] adjacent to [+raised]. The derivational approach, because of extrinsic ordering, allows us to use information about the origin of [+upper]. There is no way to distinguish between positioning motivated by different constraints in the ranked constraint analysis, and thus the origin of [+upper] is unrecoverable.

As it happens, there is a way to get closer to solving the problem in this specific case by adding another constraint. But we will see that ultimately this solution will fail. The constraint to be added, ALIGN[R],[U], given in (48), calls for the right edge of [+raised] to be aligned to the right edge of some [+upper].

(48) Align[R],[U]
ALIGN([+raised],R,[+upper],R): Align the right edge of [+raised] with the right edge of some [+upper].

This constraint can be satisfied by a failure to spread [+raised], since the underlying [+raised] is aligned to an underlying [+upper]. Failure to spread, however, can result in a violation of ALIGN[R]-R, which we still need to motivate spreading. Because it is necessary to allow satisfaction of Align[R],[U] by failure to spread, the [+upper] referred to in this constraint must not be restricted to either grammatical or lexical [+upper]. Any kind of [+upper] will suffice.

ALIGN[R],[U] is active in the tableau in (49). There it can be seen that ALIGN[R],[U] eliminates the HHM candidate which was problematic for the earlier ranked constraint analysis. Notice, though, that it does no eliminate the HHH candidate (marked **), which shares the same violations as the attested HMM form. The HHH candidate is eliminated by *STRUC, not shown in the tableau. The piece of structure that eliminates it is an extra association line between [+raised] and the final TBU.

In the tableau in (49), the first four candidates are given with explicit feature specifications. The first three represent a surface form with the attested HMM tone pattern. The feature specifications for the other candidates can all be assumed to have the underlyingly specified [+raised, +upper] on the leftmost mora. Except for the last candidate, they can also be assumed to have the floating [+upper] morpheme associated to at least one of the other moras.

(49)

H | na̲ka̲ra̲ + M	Align[U]-R	Align[R][U]	Align[R]R	Parse[F]	MTA
☞ (ná̲kā̲) rā̲ +up +up +rai			*		
(ná̲kā̲) rā̲ +up +up +rai		*!	*	*	
(ná̲kā̲) rā̲ +up +up +rai		*!	*		*
(ná̲kā̲) rā̲ +up +up +rai		*!			
(ná̲ka̲)rā̲			*		*!
(ná̲kā̲)ra̲	*!		*		*
**(ná̲ká̲)rá̲			*		
(ná̲ká̲)ra̲	*!				*
(ná̲ka̲)ra̲			*	*!	

The tableau in (49) establishes the crucial ranking of Align[R],[U] over Align[R]-R. We also know that Align[U]-R is undominated. It is not actually clear whether Align[R],[U] is undominated or not, as will be indicated below with reference to HM nouns. Thus, at least three constraint rankings are possible, but the choice among them does not affect the outcome. The three possibilities are given in (50).

(50) a.

ALIGN[U]-R
|
ALIGN[R],[U]
|
PARSE[F], ALIGN[R]-R
|
MTA

b.

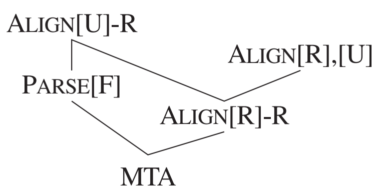
ALIGN[U]-R, ALIGN[R],[U]
|
PARSE[F], ALIGN[R]-R
|
MTA

c.

ALIGN[U]-R
|
ALIGN[R],[U], PARSE[F]
|
ALIGN[R]-R
|
MTA

The indeterminacy of where to rank ALIGN[R],[U] can be expressed as in (51), but this is difficult to capture in a tableau where all of the relevant constraints are considered.

(51)



Keeping this in mind, it is still possible to demonstrate that the constraints and constraint rankings given so far produce the desired result in HL and LHL nouns, as in (52) and (53). In fact, the added constraint has no effect and the tableaux are essentially the same as those previously given.

(52)

H | kpa̱na̱ + M	ALIGN[U]-R	ALIGN[R],[U]	ALIGN[R]-R	PARSE[F]	MTA
☞ (kpá̱ná̱) +up +up +rai					*
(kpá̱nā̱)			*!		*
(kpá̱na̱)			*!	*	

(53)

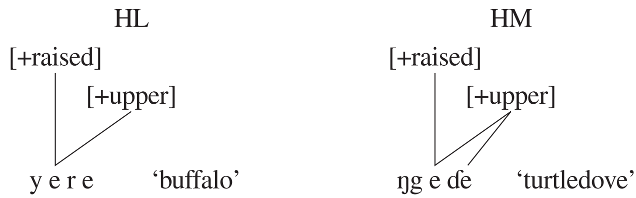
H | suma̱ri + M	ALIGN[U]-R	ALIGN[R],[U]	ALIGN[R]-R	PARSE[F]	MTA
☞ (sumá̱)rī̱ +up +up +rai					*
(sumá̱)rí̱			*!		*
(sumá̱)ri̱				*!	
(sūmá̱)ri̱	*!*				*

In this section, it has been shown that there is a way to handle [+raised] spreading in a ranked constraint analysis. However, this does not mean that there are no problems with such an analysis. In the next section, it will be shown that the constraints and constraint rankings used here pose a problem for lexical tone patterns in Suma. The derivational analysis does not run into this problem.

5. Lexical Tone Patterns: A Problem for the Ranked Constraint Analysis

The inadequacy of the ranked constraint solution to [+raised] spreading is evident when lexical tone patterns are taken into consideration. The lexical tone patterns that are problematic in this regard are underlying HL and HM patterns. The constraint Align[R]-R (43) gives preference to a candidate with a HH tone pattern when the underlying tone pattern is HM or HL. Yet HM and HL nouns are attested in Suma. Align[R][U] (48) poses problems for HM nouns if the underlying [+upper] is shared by both TBUs. The featural structure of HL and HM nouns is assumed to be as in (54).

(54) *Lexical Tone Patterns: HL, HM*



A look at lexical HM nouns in a ranked constraint analysis shows the problem posed by the constraints needed for [+raised] spreading in the associative construction. In the Tableau in (55), both the alternative candidates satisfy the constraints better than the attested HM candidate, which fatally violates both Align[R],[U] and Align[R]-R.

(55)

+rai +up ŋg e ɗ e	ALIGN[U]-R	ALIGN[R],[U]	ALIGN[R]-R	PARSE[F]	MTA
☞ (ŋgédē)		*!	*		
☜ (ŋgédé)					
☜ (ŋgēdé)					

In tableau (55), the optimal candidate would be either HH or MH in tone. The HH candidate corresponds to a HL noun in the associative construction and incurs no violations. If we assume that the underlying tones are linked in some fashion to specific TBU's and must remain linked to these TBU's with the possibility of extending their linkages elsewhere, then the MH candidate can be rejected. The notion of a sponsor, and the use of alignment-to-sponsor constraints, following Cassimjee (1995), assumes that lexical tones have specific TBU's which act as their sponsors. Once a sponsor is identified, it is possible to force the realization of the tone on its sponsor by alignment-to-sponsor constraints which align one of the edges of the tone feature to one of the edges of the sponsor. In this way, the MH candidate can be rejected.

More specifically, there are two constraints, ALIGN([+raised],L,Sponsor,L) and ALIGN([+raised],R,Sponsor,R). These constraints require that a [+raised] feature be aligned to the right and left edge of its sponsor, the mora to which it is underlyingly affiliated. In Suma, the first constraint is undominated and captures the generalization that lexical H's do not spread to the left. The second constraint is dominated by other constraints in Suma, since lexical H's do spread rightwards. This constraint will cause the rejection of a MH candidate, as with the third candidate in the tableau in (56). Since the sponsor of the [+raised] feature is the first mora and [+raised] is realized only on the second mora in the third candidate, the left edge of the tone feature is not aligned to the left edge of its sponsor and it fatally violates the undominated alignment-to-sponsor constraint.

(56)

+rai +up ŋgeɗe	ALIGN[R],S-L	ALIGN[R],[U]	ALIGN[R]-R	PARSE[F]
☞ŋgéɗē		*!	*	
☜ŋgéɗé				
ŋgēɗé	*!			

As the Tableau shows, the HH candidate remains a problem. It does not violate the undominated alignment-to-sponsor constraint. It does violate a lower ranked constraint, which aligns the right edge of [+raised] to the right edge of its sponsor. We will refer to this constraint as ALIGN[R],S-R. If we make this constraint active by ranking it above ALIGN[R],[U] and ALIGN[R]-R, then no spreading of lexical [+raised] features would occur. In order to account for the spread of [+raised] in the associative, ALIGN[R]-R must be ranked above a constraint aligning [+raised] to the right edge of its sponsor. The tableau in (57) shows how ranking ALIGN[R],S-R above the constraints inducing spreading will result in the incorrect output for nouns like *kpá̱ná̱* 'pot' in the associative.

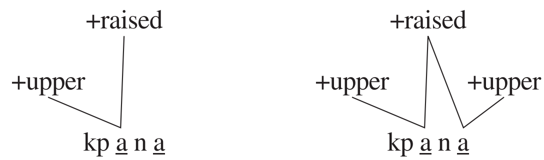
(57)

kpá̱na̱ + [+upper]	ALIGN[R], S-L	ALIGN[R], S-R	ALIGN [R],[U]	ALIGN [R]-R	PARSE[F]
☞ kpá̱ná̱		*!			
☜ kpá̱nā̱				*	
kpá̱na̱				*	*

Another way to attempt to eliminate a lexical HH candidate (**kpáná*) is by making reference to lexical vs. grammatical tones in the statement of constraints. This, however, will not handle the present case. We will define a lexical tone as a tone that is affiliated to a morpheme with segmental content. Conversely, a grammatical tone is one that is not affiliated to a morpheme with segmental content. The problem for HL nouns in Suma is that [+raised] must spread in the associative construction, and it must not spread otherwise. The [+raised] feature in question is lexical in both cases — those in which it spreads and those in which it does not. The [+upper] feature to which [+raised] aligns to avoid violation of ALIGN[R],[U] must remain unspecified as to whether it is lexical or grammatical. In the case of *ná̱kā̱rā̱* 'shoe', ALIGN[R],[U] would be fatally violated if it were to specify alignment to a grammatical [+upper]. In the case of *kpá̱ná̱* 'pot', ALIGN[R],[U] would be fatally violated if it were to specify alignment to a lexical [+upper]. This is shown in (58) where grammatical [+upper] is in italics. In (58a), [+raised] is right-aligned to a lexical [+upper] and in (58b), [+raised] is right-aligned to a grammatical [+upper].

(58) a. *UR* *SR in associative*

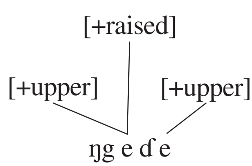
b. *UR* *SR in associative*



If we make ALIGN[R],[U] sensitive to the lexical/grammatical distinction such that [+raised] must align to a grammatical [+upper], then instead of the attested form *ná̲kā̲rā̲*, the incorrect form **ná̲ká̲rá̲* would be optimally selected. If this constraint specifies that [+raised] must align to a lexical [+upper], then instead of the attested form *kpá̲ná̲*, the incorrect form **kpá̲nā̲* would be optimally selected.

The HH candidate could be eliminated by *STRUC if the HM candidate did not violate any of the other constraints. One way that a HM noun could avoid violations might be to postulate a different underlying structure, one in which each TBU has a separate [+upper] feature, as in (59).

(59)



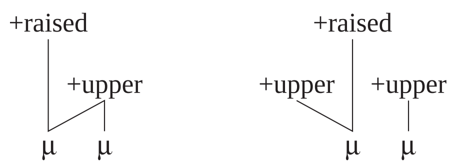
If this underlying form were to be assumed, ALIGN[R],[U] would not be violated. However, ALIGN[R]-R is still fatally violated.

An attempt to invoke different levels whereby nouns in their underlying form are subject to a different ranking of constraints than nouns in the associative construction will not save the ranked constraint analysis. HM nouns in the associative construction remain HM in tone, rather than surfacing as HH as expected from the constraints. An example is given in (60).

(60) ŋgéɗē zāŋ zɔ̲̌ ‘turtledove of the bush’

In the derivational analysis, the correct output results from the nature of the structural description of Raised Spread, where an autonomous specification of [+upper] is required. In (61), the lexical HM pattern is contrasted with the environment for Raised Spread. In the structural description of Raised Spread, [+raised] spreads when a separate specification of [+upper] is on the adjacent TBU. This structural description is not met in the case of the lexical HM pattern because there is only one specification for [+upper].

(61) *Lexical HM* *Raised Spread*



Thus, a derivational analysis can handle the tonal data from Suma and a ranked constraint analysis cannot. If we have a choice between a derivational and ranked constraint analysis of [+raised] spreading, the derivational analysis is superior.

6. Conclusions

In this paper, tone alternations on head nouns in the associative construction in Suma have been analyzed in a derivational and in a ranked constraint framework. In a derivational framework, extrinsic ordering involving a counterfeeding relationship between 2 processes is found to be crucial in accounting for the rightward spreading of [+raised]. The counterfeeding order captures a distinction between a [+upper] derived by the process of Upper Docking and a [+upper] derived by the process of Upper Doubling. A ranked constraint analysis cannot distinguish between [+upper]'s motivated by different constraints, and ultimately fails to provide a satisfactory analysis of the tonal alternations. Therefore, a serious problem is posed for an approach based on ranked constraints by data which is effectively analyzed in a derivational framework in terms of counter-feeding orders which make a distinction between properties derived by different processes.

Acknowledgments

I would like to thank Namgbei, Joel for information on the Suma language and Dave Odden for helpful comments and discussion.

Uniformity in Extended Paradigms

Eugene Buckley

1. Introduction

A number of researchers have explored recently the advantages of replacing traditional cyclic analyses of phonological patterns with constraints on identity among related words (e.g. Benua 1995; Burzio 1994; Flemming 1995; Kenstowicz 1996; McCarthy 1995; McCarthy and Prince 1994, 1995; Orgun 1994; this volume). This reformulation of cyclicity, essentially as analogy, is necessary in a parallel model of phonology, and is motivated to the extent that strict parallelism is motivated (e.g. Prince and Smolensky 1993). There are also more specific reasons to prefer correspondence among outputs over cyclicity, since it can account for phenomena which elude a cyclic framework, particularly in the case of paradigm uniformity (e.g. Flemming 1995; Kenstowicz 1996) where the word which exerts influence is not a morphological subset of the target word.

In this paper I show that in Kashaya (a Pomoan language of northern California: Oswalt 1961), paradigm uniformity makes possible a superior account of relations between underlying and surface vowel length, and its effect on the location of stress. In brief, if both phenomena are attributed to stages in a derivation, the rules must be complex and ad hoc; but if stress is influenced by uniformity, vowel length can be accounted for by means of simple and well motivated surface constraints. However, the fact that vowel length, which interacts with stress, can also be derived from morphological concatenation, leads to the conclusion that the object on which uniformity is defined is not always just the root but some larger complex constituent, which includes at least the root and the first suffix.

Much of this paper is concerned with demonstrating the need to separate the analysis of vowel length from the treatment of stress shift, a necessary background to the basic conclusions which follow. In Section 2 I give the basic foot structure of the language, contrasting analyses in ordered lexical phonology and surface-oriented Optimality Theory. In Section 3 I present and account for 'Foot Flipping', whereby the sequence CvvCv becomes CvCvv. This is accompanied by a shift in stress placement without a triggering environment in the surface form, and so motivates an appeal to paradigm uniformity, as discussed in Section 4. I then turn in Section 5 to shortening of long vowels in closed syllables , which also shows stress shift in opaque environments. In Section 6 we see the important case of Elision, whereby long vowels are derived by morphological concatenation, creating the necessary trigger for stress shift. In Section 7 I develop a notion of 'extended root' to account for the stress shift in root + suffix combinations that undergo Elision: because a mora of the suffix is incorporated into the root syllable, the suffix itself is incorporated into the root paradigm. In Section 8 I consider a special case of vowel-length alternation which might also require an extended root. A brief conclusion is given in Section 9.

2. Metrical Structure

Kashaya builds iambs from left to right. There are no secondary stresses, but the need for iterative footing is shown by Iambic Lengthening (=IL). The main (only) stress is normally on the first foot (the important exception is treated below). Note the vowel length alternations in the suffixes due to varying foot structure.

(1) a. kel-mul-ad-uced--u
→ (kél)(mula:)(duce:)(du) 'keep peering around'
b. mo-mul-ad-uced--u
→ (momú:)(ladu:)(cedu) 'keep running around'

A word-final vowel (1b) never undergoes IL; final long vowels are avoided, but more to the point all verb-final suffixes belong to a class of suffixes which permit no IL of their vowels. Non-lengthening suffixes uniformly occur to the right of those which permit IL, so the domain of IL is a substring at the left side of the word. The double hyphen (--) indicates the beginning of this non-lengthening domain.

(2) a. mo-mac-ed--ela
→ (momá:)(cede̲)(la) 'I keep running in there'

b. s'i--p^{h}ila
→ (sip^{h}í̲)(la) 'if [it] happens'

c. hoṭ^h-ala--s'uw-em
→ (hoṭ^{h}á:)(las'u̲)(wem) 'it would warm [us] up'

d. mo-mac-ed--eti
→ (momá:)(cede̲)(ti) 'although he kept running in there'

In lexical phonology (e.g. Kiparsky 1982a), the derivation of these forms requires several stages; first the lengthening suffixes are added to the root; then IL applies; and then the non-lengthening suffixes are added (Buckley 1994a). After the second round of morphology, IL does not reapply.

(3)	a.	*MORPHOLOGY 1*	mo + mac	mo + mac + ed
	b.	*Phonology (with IL)*	(moma:) ⟨c⟩	(moma:)(ce) ⟨d⟩
	c.	*MORPHOLOGY 2*	(moma:) c + eti	(moma:)(ce) d + ela
	d.	*Phonology (no IL)*	(moma:)(ceti)	(moma:)(cede)(la)

This analysis crucially refers to an intermediate representation (3b), but research in Optimality Theory has placed that approach in doubt; there is much to be gained if ordered rules are replaced with constraints on surface representations (cf. Prince and Smolensky 1991, 1993 and much subsequent work). An example is the need for provisional final-consonant extrametricality in (3b), to permit IL in intermediate *momac* (see also Buckley 1995a, b).

If we cannot appeal to *momaced* as an intermediate representation, we must refer to it as a substring of the surface representation within which IL occurs. Below, within $\{\ldots\}_1$ IL occurs, within $\{\ldots\}_2$ it does not.

(4)	a.	*Input with domains*	$\{$momaced$\}_1\{$ela$\}_2$
	b.	*Output with feet*	(momá̲:̲)(cede)(la)

This difference can be attributed to the interaction of constraints on (i) the weight of the strong branch of an iamb, and (ii) the maintenance of underlying vowel length.

First, we must generate the foot structure on which IL is based. In a surface analysis, there is no notion of directionality; instead, we must refer to the alignment of feet (McCarthy and Prince 1993a). As Crowhurst and Hewitt (1995) show, the precise manner in which a directionally based generalization such as 'left-to-right foot construction' translates into the alignment framework depends on whether degenerate feet are permitted. Below in (22) I show that it is right alignment that must be used in Kashaya; this means that degenerate feet must be permitted, as supported by the existence of monomoraic words (e.g. *cá* 'stay!').

While I give here only single-word examples, Kashaya stress is assigned to the phrase (see Oswalt 1961; Buckley 1997). The following two constraints generate the basic foot structure.

(5) ALIGN-R
Align the Right edge of every Foot with the Right edge of a Phrase (Ft,R; Phrase,R)
PARSE-SYL
Every syllable must be parsed by a foot.

In order to get the effect of iterative footing it is necessary to rank PARSE-SYL over ALIGN-R (McCarthy and Prince 1993a). I assume the undominated constraint FTFORM(Iamb).

(6)

keladucedu	PARSE-SYL	ALIGN-R
a. ☞(kelá:)(duce:)(du)		*,***
b. ke(ladú:)(cedu)	*!	**
c. keladu (cedú)	*!**	

As mentioned, using ALIGN-R to achieve the effect of left-to-right footing requires that we permit degenerate feet (but see also McCarthy and Prince 1993a: 91). In languages that avoid such feet, FTBIN is responsible (Prince and Smolensky 1993). I assume that universally no foot is larger than two syllables.

(7) FTBIN
A foot is binary under moraic or syllabic analysis.

Since, however, degenerate feet are necessary for ALIGN-R in Kashaya, PARSE-SYL » FTBIN.

(8)

keladucedu	PARSE-SYL	FTBIN
a. ☞(kelá:)(duce:)(du)		*
b. (kelá:)(duce:) du	*!	

The basic effect of IL is to achieve a perfect or canonical iamb, which consists of a light (and unstressed) syllable followed by a heavy (and possibly stressed) syllable (cf. Hayes 1985, 1995a).

(9) ASYM
In a branching iamb, the strong branch must be heavy.

ASYM is ranked lower than ALIGN-R, as will be shown by the candidates (14b) and (14d). Notice in (10) that the location of feet is determined by PARSE-SYL and ALIGN-R, while ASYM secondarily determines the internal composition of those feet.

(10)

keladucedu	PARSE-SYL	ALIGN-R	ASYM
a. (kelá)(duce)(du)		*,***	*!*
b. (kelá:)(duce)(du)		*,***	*!
c. ☞(kelá:)(duce:)(du)		*,***	
d. (kelá:)(du)(cedu)		**,***!	*
e. (kelá:) ducedu	*!**	***	

IL must be prevented in the non-lengthening domain by a constraint of the following type (cf. Urbanczyk 1995: 512, McCarthy 1995a: 43).

(11) Q-IDENT
The quantity of each input segment must be identical to its output quantity.

The difference between lengthening and non-lengthening suffixes is quite simply a matter of which constraint wins: ASYM or Q-IDENT. Since the winner differs across the two domains, there must be a different constraint ranking in those domains.

Following Buckley (1995a, b), I assume the existence of C[onstraint]-domains to which constraints can be particularized.

(12) $\{$root + lengthening suffixes$\}_1$ $\{$non-lengthening suffixes$\}_2$

Two C-domains, C1 and C2, require two domain-specific constraints. Q-IDENT$^{\{1\}}$, which evaluates only segments in the lengthening C1, is ranked below ASYM; while Q-IDENT$^{\{2\}}$, for the non-lengthening C2, dominates ASYM to prevent IL.

(13) Q-IDENT$^{\{2\}}$ » ASYM » Q-IDENT$^{\{1\}}$

The UR, with domains labeled, is shown in the upper left corner of the tableau.

(14)

$\{$kelala$\}_1$ $\{$ p^hila$\}_2$	Q-IDENT$^{\{2\}}$	ALIGN-R	ASYM	Q-IDENT$^{\{1\}}$
a. (kelá)(lap^hi)(la)		*,***	**!	
b. ☞(kelá:)(lap^hi)(la)		*,***	*	*
c. (kelá:)(lap^hi:)(la)	*!	*,***		*
d. (kelá:)(la)(p^hi)(la)		*,**,**!*		*

Although in (14c) iambic structure is perfectly satisfied, it happens at the expense of preservation of underlying vowel length in the suffix *-p^hila*, subject to high-ranking Q-IDENT$^{\{2\}}$. In (b), iambic asymmetry is met only within the domain where low-ranked Q-IDENT$^{\{1\}}$ is violated, making it optimal.[1]

It is a basic fact of Kashaya that the first syllable of the word is extrametrical when the root is at least two syllables in length, or when a root of any length is preceded by a prefix. This can be seen in the displacement of both stress and IL.

(15) a. libut-ad--u
→ ⟨li⟩ (butá:)(du) 'keep whistling'
b. bimucid-uced--u
→ ⟨bi⟩ (mucí:)(duce:)(du) 'used to eat'
c. du-k'il-ic' — i
→ ⟨du⟩ (k'ilí:)(c'i) 'point at yourself'
d. do-hqoṭol-ic'-ed--a-em
→ ⟨doh⟩ (qoṭó:)(lic'e:)(dam) 'couldn't get around'

The examples given earlier all have monosyllabic roots; syllable extrametricality is introduced here because it figures in many examples below, but it is orthogonal to the main discussion. See Buckley (1995c) for discussion and analysis; for present purposes I simply assume satisfaction of constraints, in particular

1. Not only is Q-IDENT$^{\{1\}}$ low-ranked relative to ASYM, it in fact never plays any role in choosing candidates. Any form that Q-IDENT$^{\{1\}}$ might favor is ruled out by ALIGN-R, which dominates Asym and therefore necessarily Q-IDENT$^{\{1\}}$. (Every long vowel leads to a new foot and adds violations of ALIGN-R). As noted by Buckley (1995a), an alternative to the view that a constraint such as Q-Ident exists in two domain-specific forms is that there is only one constraint, but (in this case) it is ignored in C1. The important point is that violations within C1 never matter, whether this is treated as low ranking of a domain-particularized constraint, or by completely ignoring the violations.

NONINITIAL, which ensure that the initial syllable is excluded from foot structure in these cases.

3. Foot Flipping

Of central interest in this paper is the process that Buckley (1994a, b) calls Foot Flipping. When the leftmost footed sequence of the word is CvvCv, the vowel lengths in the two syllables are 'flipped' or reversed, resulting in the perfect iamb CvCvv.

(16) a. di̲:̲c'-aq^{w}-ic'--i
→ (dic'a̲:̲)(qoc'í) 'take a message out!'
b. q'a̲:̲-cid--u
→ (q'aci̲:̲)(dú) 'keep leaving'
c. miku̲:̲ṭ-ad--e:
→ ⟨mi⟩ (kuṭa̲:̲)(dé:) 'keep humming'
d. mu-bo̲:̲k'-ibic--ʔ
→ ⟨mo⟩ (bok'i̲:̲)(bíʔ) 'start to rise'

In addition to the flipping of vowel lengths, notice that the stress falls on the second foot, rather than on the first one as is normally the case in Kashaya. Below in Section 4 this stress shift receives an analysis in terms of paradigm uniformity.

Cvv followed by a heavy syllable does not undergo Flipping. The reason: the maximal syllable in Kashaya is CvC, and Flipping would result in *CvvC or *Cvvv (see Section 3.1).

(17) a. di:c'--iʔba
→ (di̲:̲)(c'íʔ)(ba) 'could tell'
b. q'a:-muc'--ba
→ (q'a:)(múc')(ba) 'after leaving each other'
c. miku:ṭ--e:
→ ⟨mi⟩ (ku̲:̲)(ṭé:) 'hums'
d. kilu:ca-:qa--w
→ ⟨ki⟩ (lu̲:̲)(cá:)(qaw) 'a lock'

In this case the stress also falls on the second foot in the word. What both cases of stress shift have in common is a long vowel which seems to start out at the beginning of the foot that is skipped.

Buckley (1994a, b) proposes a serial analysis whereby a rule of Foot

Extrametricality applies to any foot beginning with Cvv, thereby uniting Cvv and (underlying) CvvCv. This requires temporary creation of the ill-formed ‘anti-iamb’ CvvCv, which persists until Foot Extrametricality applies, after which a literal rule of Foot Flipping simply reverses the vowel lengths to create a true iamb. Henceforth I use « » for an extrametrical foot.

(18)	i.	*Foot Construction*	(dí:)(c’ah)(qaw)	(dí: c’a) (qo c’i)
	ii.	*Foot Extrametricality*	«di:»(c’áh)(qaw)	«di: c’a»(qo c’í)
	iii.	*Foot Flipping*	---	«di c’a: (qo c’í)

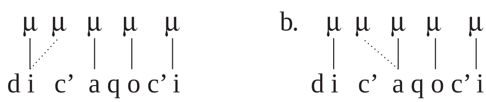
In addition to the ad hoc nature of Foot Flipping and the temporary creation of ill-formed structures, Foot Extrametricality also requires the dubious generalization “begins with Cvv”, to cover Cvv (17) and CvvCv (16). (This generalization also applies to CvvC in Section 5.)

A more principled analysis is possible using constraints. The change does not need to be analyzed as ‘flipping’ per se, whereby the mora moves from one syllable to another. Rather, it can be seen as underlying indeterminacy in the association of the mora, which is resolved by metrical and syllabic well-formedness; that is, the association of the second mora is underspecified (cf. Kiparsky 1993).

(19) a. *root morpheme* b. *with suffixes*

There are two basic surface realizations possible: leftward or rightward linking to a vowel.[2]

(20) a. b.



The choice between these two forms is made by ALIGN-R, which prefers branching feet at the left edge. (A raised period [·] indicates a floating mora in the UR; a colon [:] is a linked mora.)

2. While it appears to be a fact about Kashaya that verb roots normally do not have true underlying long vowels, a few are non-alternating and fail to undergo Foot Flipping (e.g. *c’a:hac-* ‘get married’). Such roots simply have two underlyingly linked moras (Buckley 1997).

(21)

$\{\text{di·c'aqoc'}\}_1\{\text{i}\}_2$	Q-IDENT$^{\{2\}}$	ALIGN-R	ASYM
a. (di:)(c'aqo:)(c'i)		*,**!*	
b. ☞(dic'a:)(qoc'i)		**	*
c. (dic'a:)(qoc'i:)	*!	**	

The next example shows clearly that rightward alignment is necessary in Kashaya, since (22a, b) are identical in all respects except for the location of vowel length and foot boundaries.

(22)

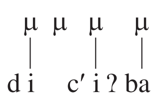
$\{\text{miku·ṭad}\}_1\{\text{e}\}_2$	NONINITIAL	PARSE-SYL	ALIGN-R
a. mi (ku:)(t̥adé:)		*	**!
b. ☞mi (kut̥a:)(dé:)		*	*
c. (miku:)(t̥adé:)	*!		**

The major success of the constraint-based analysis is that the same constraint needed to determine foot structure in simple cases — namely, ALIGN-R — serves as the motivation for Foot Flipping. The fact that the lexical phonology analysis requires the ad hoc rule of Foot Flipping to accomplish the same task constitutes a strong argument in favor of the constraint-based approach.

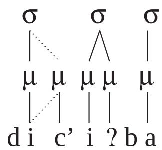
3.1 *Phonologically Blocked Flipping*

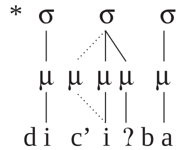
As seen in (17), vowel length surfaces in the root, i.e. there is no Flipping, when the rightward potential docking site of the floating mora is a heavy syllable. Below is a relevant input representation.

(23)



The first consonant in the cluster /ʔb/ syllabifies as a coda, where it must bear a mora itself. That makes a total of two moras in the syllable, with no room for the floating mora (by the constraint BIMORA). This forces leftward linking.

(24) a. 

b. 

In (24b), to avoid a bimoraic syllable, but maintain the long vowel, the coda would have to be eliminated from the syllable, either by deleting the consonant (*dic’i:ba*) in violation of MAX; or by inserting a vowel so the consonant can syllabify as an onset (*dic’i:ʔ<u>i</u>ba*), violating DEP (see McCarthy and Prince 1995).

(25)

{di·c’iʔ}$_1${ba}$_2$	BIMORA	MAX	DEP	ALIGN-R
a. ☞(di:)(c’iʔ)(ba)				*,**
b. (dic’i:ʔ)(ba)	*!			*
c. (dic’i:)(ba)		*!		*
d. (dic’i:)(ʔiba)			*!	**

See Buckley (1995c) for discussion of constraints which prevent other rearrangements of the moraic structure, e.g. MORA-IDENT, MAX(μ), and LINEARITY.

3.2 *Morphologically Blocked Flipping*

In addition to its ad hoc nature, the serial lexical phonology analysis in (18) treats IL and Foot Flipping as independent rules, yet the domains of the two rules are identical. For example, notice that the suffix *-mela* resists IL (26a) as well as Foot Flipping (b).

(26) a. baṭi--mela → ⟨ba⟩ (ṭimé)(la) ‘[we] camped’
*⟨ba⟩ (ṭimé:)(la)
b. q’a·--mela → «q’a:»(melá) ‘I left’
*«q’ame:»(lá)

Formally, this shared restriction on IL and Foot Flipping is easy enough to state, by assigning both rules to the same lexical level (cf. Buckley 1994a). But this move provides no explanation as to why this correlation should obtain, and it is predicted that a similar language might have the same rules in different levels. This prediction is dubious, since both processes result in the same perfect iamb.

In the OT analysis, the high-ranking status of Q-IDENT$^{\{2\}}$ accounts for both facts: IL and Foot Flipping both introduce a long vowel, and Q-IDENT$^{\{2\}}$ ensures that this does not occur in C2.

(27)

$\{$q’a·$\}_1\{$mela$\}_2$	Q-IDENT$^{\{2\}}$	ALIGN-R	ASYM
a. ☞(q’a:)(melá)		**	*
b. (q’ame:)(lá)	*!	*	

The special status of C2 is stipulated for a single constraint, which by itself accounts for the lack of both processes. Such an explanation is not possible in the ordered-rule approach, and this fact is a powerful argument against such a derivation. See also Buckley (1996).

4. Stress Uniformity

The constraint-based analysis elegantly unifies the accounts of IL and Foot Flipping. Now we must deal with the similarity in stress patterns between the flipped and non-flipped words. Recall the general pattern according to which an initial Cvv foot is skipped in choosing the main stress of the word — i.e. it is extrametrical (28). The initial foot dominating the same root is extrametrical even when it is not of the shape Cvv, i.e. when it is flipped or shortened (29).

(28)	a.	di·c’--iʔba	→	«di:» (c’íʔ)(ba)	‘could tell’
	b.	di·c’--ela	→	«di:» (c’elá)	‘I tell’
	c.	di·c’--i	→	«di:» (c’í)	‘tell!’
(29)	a.	di·c’-aq^{w}-ic’--i	→	«dic’a:» (qoc’í)	‘take a message out!’
	b.	di·c’-id--a-em	→	«dic’i:» (dám)	‘told about’

Contrast this the situation of a root with no (underlying) long vowel in the first syllable, and therefore no skipping of the first foot.

(30)	a.	kel--iʔba	→	(kelíʔ)(ba)	‘could peer’
	b.	kel--ela	→	(kelé)(la)	‘I peer’
	c.	kel--i	→	(kelí)	‘peer!’
(31)	a.	kel-adad--u	→	(kelá:)(dadu)	‘look at while riding’
	b.	kel-ma--w	→	(kél)(maw)	‘peer down at’

The forms in (31a) and (29a) have identical syllable structures, but different stresses. In (29) underlying vowel length has been shifted, but the resulting foot is skipped just like Cvv in (28).

A framework tied to surface constraints cannot refer to intermediate levels as was done in the lexical phonology analysis illustrated in (18), but as we saw that analysis has numerous problems anyway. First let is account for the case of a simple Cvv foot. Such a foot, when the first one in the domain, is skipped for stress. For present purposes, this can be accomplished by the following constraint.

(32) Skip-Ft
Do not stress an initial CVV foot.

The formal consequence is exclusion of the foot from line 2 constituency in the metrical grid. It is not possible to have a constraint which says not to stress **any** Cvv foot, since the effect is not iterative; cf. ⟨*ki*⟩*lu:cá:qaw* in (17d).

Skip-Ft must dominate two basic constraints on metrical structure.

(33) Parse-Ft
Incorporate a line 1 constituent (a foot) into a line 2 constituent.
Align-Hd
Align the head of the phrase with the left edge of the phrase.

Under this analysis, the location of stress in a phrase (here, just one word) beginning with surface Cvv (modula syllable extrametricality) is due to Skip-Ft and the floating mora linking leftward (28). Every word that undergoes Flipping is related to a word where none occurs, and where shift is motivated transparently on the surface by Cvv. This fact is illustrated for the root *di·c'* in (28) and (29). Whereas in many words shift is directly due to Skip-Ft, in words that lack initial Cvv (thanks to Flipping), the shift follows not from the surface form but from uniformity with words that do have Cvv.

(34)	a.		(x)	b.			(x)
		(x)	(. x)		(.	x)	(. x)
		di:	c'e lá		di	c'a:	qo c' í
		by Skip-Ft			*by uniformity with (a)*		

Specifically, the location of the metrical head of the word must remain consistent across instantiations of a root or stem.

(35) Uniformity
If the first foot is stressed in one instantiation of a root, then it must be stressed in all instantiations of that root.

In the following tableau, related output forms are evaluated together, so that the winning candidate is actually a set of forms, rather than a single form.

	Candidate Sets	UNIFORMITY	SKIP-FT	PARSE-FT
(36)	a. ☞(kelá:)(dadu) (kelíʔ)(ba)			
	b. «kela:» (dadú) «keliʔ» (bá)			*! *
	c. (kelá:)(dadu) «keliʔ» (bá)	*!		 *
(37)	a. (dic'á:)(qoc'i) (dí:)(c'iʔ)(ba)		 *!	
	b. ☞«dic'a:»(qoc'í) «di:» (c'íʔ)(ba)			* *
	c. (dic'á:)(qoc'i) «di:» (c'íʔ)(ba)	*!		 *

In (36), perfect satisfaction of all three constraints is possible. In (37), the two highest ranked (UNIFORMITY and SKIP-FT) can be satisfied, so there is no necessary ranking between them, only that both dominate PARSE-FT. The winning candidate set obeys SKIP-FT in the word where that constraint matters, and patterns the second word after the first (in order to obey UNIFORMITY).

5. Closed-Syllable Shortening

Foot Flipping is not the only source of surface-opaque stress shift, and therefore not the only motivation for paradigm uniformity. Notice in (38) that the first foot is again skipped for stress, even though on the surface it does not contain a long vowel.

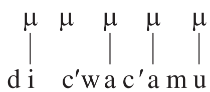
(38)
a. di·c'-wac'--a-emu
→ «dic'»(wac'á)(mu) 'what they say (is)'
b. di·c'-maq--o
→ «dic'»(maqó) 'bring the message in!'
c. di·c'-mul-ic'--i
→ «dic'»(mulí:)(c'i) 'bring the message around!'
d. miku·ṭ--qa-e:
→ ⟨mi⟩«ku·$ṭ^h$»(qá:) 'must have hummed'

e. ca-hwe·n-muc'--ʔli
→ ⟨cah⟩«wen»(múʔ)(li) 'a seesaw'
f. da-li·t-qa--w
→ ⟨da⟩«lit^h»(qáw) 'let wave with the hand'

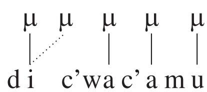
As with Foot Flipping, in a serial analysis an ill-formed structure — in this case superheavy CvvC — must be temporarily permitted until Foot Extrametricality applies, after which it undergoes Shortening (cf. Buckley 1991).

(39)
i. *Foot Construction*
(di:c')(wac'a)(mu) ⟨da⟩ (li:t^h)(qaw)
ii. *Foot Extrametricality*
«di:c'»(wac'a)(mu) ⟨da⟩ «li:t^h»(qaw)
iii. *Shortening*
«dic'»(wac'á)(mu) ⟨da⟩ «lit^h»(qáw)

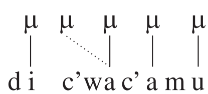
The OT analysis developed so far actually accounts quite easily for Closed-Syllable Shortening. The following input has a cluster of two consonants, /c'w/, following the root vowel.

(40)


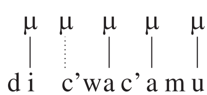
In this context, the floating mora has not just two places to link, in the ways seen above, but it can also, as shown in (41c), serve as the mora for the coda consonant /c'/.

(41) a.


= *(di:c')(wac'a)(mu)

b.


= *(dic')(wa:)(c'amu)

c.


= *(dic')(wac'a)(mu)

The form in (41a) is ruled out by BIMORA, as in (25). The linking to the following vowel in (41b), by contrast, is well-formed syllabically. But it is not as well aligned as (41c), where the floating mora links to the coda consonant and pre-empts Weight-by-Position (e.g. Hayes 1989).

(42)

{di·c'wac'}$_1${amu}$_2$	BIMORA	ALIGN-R	ASYM
a. (di:c')(wac'a)(mu)	*!	*,***	*
b. (dic')(wa:)(c'amu)		**,***!	*
c. ☞(dic')(wac'a)(mu)		*,***	*

This derivation captures a reasonable intuitive interpretation of Closed-Syllable Shortening, that the coda consonant 'steals' the second mora of a long vowel. The only quirk in Kashaya is that the mora in question was never actually linked to the vowel. (Stress shift with Shortening is treated below in Section 7.)

6. Elision

Within C1, adjacent vowels /V_iV_j/ become long [V_i:] by Elision.[3] This occurs where Flipping is blocked, whether phonologically (43) or morphologically (44).

(43) a. mo-ibic--ʔ
→ «mo:»(bíʔ) 'run away'
b. ca-ad--uʔba
→ «ca:»(dúʔ)(ba) 'could fly'
c. puhṭi-aqac--ʔ
→ ⟨puh⟩«ṭi:»(qáʔ) 'go up alone'

(44) a. mo-aq--ela
→ «mo:»(qalá) 'I'm running'
b. mo-ad--eti
→ «mo:»(detí) 'even though [it] was running'
c. c^hi-de-ad--u
→ ⟨c^hi ⟩«de:»(dú) 'carry along'

The two adjacent vowels are also a common source of Foot Flipping when the next syllable would otherwise be light Cv.

---

3. In C2 the result is short [V_i], by Q-IDENT, which prevents compensatory lengthening; cf. (38a), where /ae/ surfaces as short /a/.

(45) a. mo-aloq^{w}-ic’--i
→ «molo:»(qoc’ í) ‘run up out here!’
b. do-ibic--i
→ «dobi:»(cí) ‘raise your hand!’
c. yehe-ala-mec’--t^{h}u-ʔ
→ ⟨ye⟩«hela:»(méʔ)(t^{h}uʔ) ‘don’t drag yourself down!’

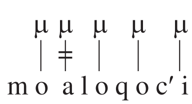
In an ordered-rule framework, Buckley (1994a, b) has to assume that even in cases of Flipping, the intermediate step exists in which the two vowels are syllabified together, and then the length is flipped.

(46) *Underlying form*
moaloqoc’i cahnoaduc’i
i. *Elision and Footing*
(mo:lo)(qoc’i) ⟨cah⟩(no:du)(c’i)
ii. *Foot Extrametricality*
«mo:lo»(qoc’i) ⟨cah⟩«no:du»(c’i)
iii. *Foot Flipping*
«molo:»(qoc’ í) ⟨cah⟩«nodu:»(c’ í)

This intermediate CvvCv foot is necessary to trigger both Foot Extrametricality and Foot Flipping. Similarly, intermediate superheavy CvvC is required for forms with Closed-Syllable Shortening (cf. (38)), though in this case the only need is to trigger Foot Extrametricality.

(47) a. mo-aq--mela
→ «mo:h»(melá)
→ «moh»(melá) ‘I ran through there’
b. p^{h}ila-ac’--me-ʔ
→ ⟨p^{h}i⟩«la:c’»(méʔ)
→ ⟨p^{h}i⟩«lac’»(méʔ) ‘come up here! (pl.)’

The floating-mora analysis extends easily to these cases, and obviates the intermediate steps. Assume that loss of the second set of vowel features is accomplished by constraints such as ONSET (e.g. Prince and Smolensky 1993) and NODIPHTHONG (e.g. Rosenthall 1994).

(48)


The second mora, prohibited from remaining linked to its own features, behaves identically to an underlyingly floating mora, as in (19b). The two output possibilities are precisely those outlined in (20).

(49)

$\{moala\}_1\{qoc'i\}_2$	ALIGN-R	ASYM
a. (mo:)(laqo:)(c'i)	*,**!*	
b. ☞(mola:)(qoc'i)	**	*

In a sense, a root like *di·c'*- resembles *mo-aq*-, rather than vice versa; there is no long vowel in the underlying form of *di·c'*-, simply a short vowel and a floating mora, just as in *ca-ad*- there is a short root vowel plus a mora provided by the suffix. Once again the constraint-based analysis makes possible a simpler and more elegant account of the alternations.

7. Extended Roots

For consonant-final roots, the property of skipping the first foot is shared for all instantiations of that root. For any root ending in a long vowel, this is also true. For example, in the case of Closed-Syllable Shortening (a), the root is skipped under the influence of forms with the same root but which do not have shortening (b-c).

(50) a. di·c'-wac'--a-emu
→ «dic'»(wac'á)(mu) 'what they say (is)'
b. di·c'--iʔba
→ «di:»(c'íʔ)(ba) 'could tell'
c. di·c'--ela
→ «di:»(c'elá) 'I tell'

(51) a. hi-s'a· -hqa--w
→ ⟨hi⟩«s'ah»(qáw) 'cause to break'
b. hi-s'a·--ti
→ ⟨hi⟩«s'a:»(tí) 'in order to break'
c. hi-s'a·--mela
→ ⟨hi⟩«s'a:»(melá) 'I broke'

In fact, with Closed Syllable Shortening, it is crucial that paradigm uniformity apply only to the root itself. The reason for this is that the consonant cluster following the root vowel ensures that it will shorten in every case of that suffix.

For example, when *-hqa* follows *hi-s'a:-*, the root vowel is always short. Only by looking at the root in cases where another suffix follows, e.g. *hi-s'a:-tí*, is it possible to detect a surface-motivated example of stress shift. So far, then, the paradigm which must be uniform is defined by the root morpheme alone.[4]

The situation is more complicated when it comes to a root ending in a short vowel, however. Because of Elision, such roots may or may not have a long vowel to undergo Flipping, depending on which suffix immediately follows. Take, for example, *mo-* 'run'. If a consonant follows, it patterns like *kel-* in (36), with no foot skipping.

(52) a. mo-mul--i
→ (momú:)(li) 'run around!'
b. mo-ht-mul--ʔ
→ (móh)(timul') 'ran around (pl.)'

If a vowel follows, however, Elision yields an extra mora, and the complex stem patterns like *di·c'-* (37), i.e. the first foot is uniformly skipped, with or without Flipping.

(53) a. mo-ibic--ʔ
→ «mo:»(bíʔ) 'to run away'
b. mo-ibic--ba
→ «mo:»(bíc')(ba) 'after running away'
c. mo-ibic--i
→ «mobi:»(cí) 'run away!'
d. mo-ibic-ed--u
→ «mobi:»(cedú) 'to run away'
e. mo-ibic-ed-uced--u
→ «mobi:»(cedú:)(cedu) 'to keep running away'

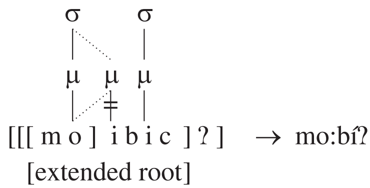
In determining the uniform paradigm, then, more than just 'root' is relevant: we must be able to take account of the first suffix as well. It is this morphologically complex constituent for which paradigm uniformity is enforced.

This exception to the 'root only' generalization is motivated prosodically by the fact that, in cases such as (53a–b), a suffixal mora joins with the mora of the root to create the long vowel in the root syllable. The apparent consequence of this fact is that the combined root + suffix constituent is what serves as the basis of paradigm uniformity. In other words, when a suffixal mora is incorporated

4. Prefixes such as *hi-* in (51) are not relevant since they are always extrametrical.

into the same syllable as the root, then the suffix itself joins into an ‘extended’ root constituent.

(54)



In (54), the morphological affiliation of the second mora is with the suffix *-ibic*. But its phonological fate becomes tied to that of the first mora, which belongs to the root. Because the root nucleus extends into the suffix, the root paradigm does as well. The result is an extended root which is the new basis for paradigm uniformity. An alternative way to think of it is that the root, for the purposes of defining the paradigm, cannot end in the middle of a nucleus — the latter constituent is simply too coherent an entity.[5] In this regard, the operative notion is related to the idea of alignment (McCarthy and Prince 1993a): just as some morphological constituents are prohibited from spanning a syllable, so in Kashaya the root paradigm cannot span a nucleus.

This principle defines combinations like [[mo]ibic] as complex paradigms. While prosodically motivated, the paradigm that results is fundamentally morphological in nature (i.e. it consists of the root and suffix, not the root syllable alone). Therefore this extended root serves as a complex paradigm in all instantiations of [[mo]ibic], including those where the root syllable does not in fact include a mora from the suffix — as is the case, of course, in all examples with Foot Flipping, such as [[[mo]ibic]i] → *mobi:cí*.

An interesting parallel for the Kashaya extended root comes from the well-known operation of analogy in certain roots taking a nasal infix in Latin (e.g. Ernout 1953). This infix functioned originally to mark the present stem and preserves this function in many cases.

(55)

	‘break’	‘split’	‘conquer’
pres. act. indic. 1sg.	rump-ō	find-ō	vinc-ō
pres. infinitive	rump-ere	find-ere	vinc-ere
perf. pass. part. m. sg.	rup-tus	fis-sus	vic-tus
perf. act. indic. 1sg.	rūp-ī	fid-ī	vīc-ī

---

5. It is necessary to speak of the nucleus, and not just the syllable, since a coda consonant from a suffix is not sufficient to induce a similar extension of the root paradigm (as in *hi-s’a-h.qá-w*).

However, in a number of other roots the infix has been generalized to some or all other forms of the verb, not just the present.

(56)		'fashion'	'joint'	'lick'
	pres. act. indic. 1sg.	fing-ō	jung-ō	ling-ō
	pres. infinitive	fing-ere	jung-ere	ling-ere
	perf. pass. part. m. sg.	fic-tus	jūnc-tus	līnc-tus
	perf. act. indic. 1sg.	fīnx-ī	jūnx-ī	līnx-ī

These verbs historically have no nasal in the root (cf. the English cognates *yoke*, *lick* with Latin *jung*-, *ling*-), but synchronically the nasal is fully incorporated into the root, just like in those cases where the nasal is an original part of the root (e.g. *tingō* 'moisten'). This analogical change is an entirely natural development, given the fact that the nasal morpheme intrudes phonologically into the root. In Kashaya the incorporation of a suffixal mora into the root syllable creates a similar intrusion that motivates a parallel extension of the root paradigm.

Another, quite distinct phenomenon which supports the notion of extended paradigm proposed here comes from Bantu reduplication. As shown by Sibanda (1997), in Ndebele material from an inflectional suffix (such as final vowel -*e*) is not normally included in reduplication; instead a default /a/ is inserted where necessary to satisfy the disyllabic template.

(57)	a.	-bon-e	-bon-a + bon-e	'see'
			*-bone + bone	
	b.	-bon-il-e	-bon-a + bon-il-e	'saw'
			*-boni + bonil-e	

The relevant data arise when the perfective *-il-e* undergoes *imbrication*, i.e. fusion with a root ending in /l/ (and sometimes other consonants), as illustrated in (58). When such a form is reduplicated, not only is default /a/ possible (59a), but optionally the suffix material can be included instead (59b).

(58)	a.	-libal-		'forget'
	b.	-libal-il-e	→ -libel-e	'forgot'
(59)	a.	-lib-a + lib-el-e		
	b.	-libe + libel-e		

In the case where default /a/ is used, it is likely that the overwhelming Bantu pattern in favor of CVC roots and VC suffixes leads to the reanalysis of the output of imbrication — [el] in (59a) — as suffixal rather than root material, in which case it follows the pattern in (57b). The alternative, in (59b), is an extended root that incorporates the suffixal material.[6]

Significantly, if we look at imbrication with a shorter root we see that the final suffix -*e* cannot be included here, just as it cannot be included in (57b).

(60)	a.	-thath-			'take'
	b.	-thath-il-e	→	-theth-e	'took'
(61)	a.	-theth-a + theth-e			
	b.	*-thethe + thethe			

The generalization is that only a suffix which has been fused with the root can be included in reduplication; a concatenatively distinct suffix such as -*e* cannot, even when an adjacent suffix has undergone imbrication. Similar data exist elsewhere in Bantu, such as Kinande (Hyman and Mutaka 1990) and Kikerewe (Odden 1996). Just as in Kashaya and Latin, as phonological fusion of the root and suffix creates a single 'extended' constituent to which other processes are sensitive.

In his discussion of paradigmatic analogy, Kenstowicz (1996) states that uniformity functions to 'minimize the differences in the realization of a lexical item (morpheme, stem, affix, word)'. For example, uniformity in the realization of /s/ applies specifically to the Spanish prefix *des*- but not to other morphemes; while uniformity in correspondence of /r/ applies to Latin noun roots such as *hono:s*, which is subject to the influence of *hono:r-em* etc. In Kashaya, the basic lexical item in question is the root; but with the special case of the extended root in the case outlined above. A clear and important point that we can draw from the Kashaya data is that paradigm uniformity cannot in all cases be restricted to single morphemes. If we could refer only to, say, the root *mo*- 'run', we would be trapped in the contradiction between the forms with this root in (52), stressed on the first foot; and those in (53), stressed on the second foot. The extended root resolves this conflict.

8. A Special Case

A phenomenon distinct from Elision which, however, also leads to the same sort of dual behavior for a root is termed Morphological Shortening by Buckley (1994a). Rather than lengthening an underlying short vowel, this process —

6. It is worth noting that the reanalysis of a CVC root plus a VC suffix as simple root of the shape CVCVC is the diachronic origin of most of the longer roots in modern Bantu languages.

triggered by certain null or contentful suffixes such as the ‘Plural Act’ (Oswalt 1961) — eliminates an underlying floating mora from the root.

(62) a. di-cʰi·ṭ- → dicʰi· ṭ- ‘fall out (sg.)’
b. di-cʰi·ṭ-Ø- → dicʰiṭ- ‘fall out (pl.)’

(63) a. duqa·c- → duqa·c- ‘get lost (sg.)’
b. duqa·c-ta- → duqatac- ‘get lost (pl.)’

In (63b), the *-ta* allomorph is infixed; cf. (66d) below.

As expected, only the singular form with the floating mora shows stress shift, but it does so uniformly.

(64) a. di-cʰi·ṭ--eti
→ ⟨di⟩«cʰi:»(ṭetí) ‘despite falling out (sg.)’
b. di-cʰi·ṭ-ad-uced--u
→ ⟨di⟩«cʰiṭa:»(ducé:)du ‘habitually fall out (sg.)’

The plural, with a shortened vowel, uniformly shows **no** stress shift.

(65) a. di-cʰiṭt̥--eti
→ ⟨di⟩ (cʰiṭé)ti ‘despite falling out (pl.)’
b. di-cʰit̥-ad-uced--u
→ ⟨di⟩ (cʰiṭá:)(duce:)du ‘habitually fall out (pl.)’

These facts demonstrate that whether or not the Plural Act suffix is present must be taken into account in determining paradigm uniformity — that is, this suffix must also be included in an extended root. But unlike in the case of Elision, there is not the same prosodic motivation; the suffix results in the loss of a mora in the root syllable, not the addition of one. If we generalize the notion illustrated in (54) so that it appeals not only to the inclusion of a suffixal mora in the root syllable, but rather to any suffix which results in the addition or loss of a root-syllable mora, the desired effect follows. The intuitive appeal of the explanation is perhaps then weakened, but since the nature of pseudo-phonological rules which serve as the exponence of a morphological category is not well understood, it is difficult to say to what degree the extended root in this case is well motivated.

An alternative analysis is to say that Morphological Shortening is not actually a productive process; that is, pairs like *cʰi·t̥-* and *cʰit̥-* are not derived but are listed lexically. There is good evidence for this position, since Plural Act morphology in general is highly idiosyncratic. For example, there are a number of suffixes and infixes attested, many of which appear in similar environments (Oswalt 1961).

(66)		SINGULAR	PLURAL	
	a.	diʔk'ol-	diʔk'ol-aq	'prune (branch)'
	b.	baʔt'il-	baʔt'il-m-	'be too noisy'
	c.	dahqoṭol-	dahqoṭol-ta-	'fail (to do)'
	d.	bilaq^{h}am-	bilaq^{h}a-ta-m	'feed'
	e.	p^{h}anem-	p^{h}ane-t-m-	'hit with the fist'
	f.	kel-	ke-h-l	'peer'

The choice of affix here must be lexically listed. Further, some patterns are unique, and must be listed as suppletive roots.

(67)		SINGULAR	PLURAL	
	a.	-q^{h}o·c-	-q^{h}oṭ-	'take out by the roots'
	b.	-moʔon-	moʔoʔta-	'strike'
	c.	s'oʔo·m-	s'oʔohwim-	'excrete'
	d.	ʔaš-	ʔac^{h}ulaq-	'miss (a target)'

Notice that some of these verbs involve shortening as well as segmental changes (67a,c). One might treat the pure shortening examples such as (62) as similarly listed forms, which happen not to have any segmental changes. If this is the case, then the plural is morphologically a simple root, without any internal structure, and it is hardly surprising that each constitutes a distinct paradigm from the independently listed singular root. But since other forms with shortening, illustrated by (61), do occur with a relatively productive suffix, the lexicalization analysis seems to miss an important generalization. At any rate, the available data regarding Morphological Shortening are rather limited, and I leave the resolution of this question open.

9. Conclusion

I have argued that an analysis with surface constraints captures the formal similarities between phenomena in Kashaya such as Iambic Lengthening and Foot Flipping in ways not available to an analysis reliant on intermediate steps — namely, the interaction of constraints such as ALIGN, ASYM, and Q-IDENT. To provide a full accounting of the facts, two enrichments to the theory are necessary; constraint domains, which permit substrings to be subject to different constraint rankings; and paradigm uniformity, which permits the optimal form of one word to be determined in part by reference to the output form of another word. Further, uniformity must be able to refer to paradigms defined by a root plus a suffix in particular circumstances where the phonological bond between the two morphemes is sufficiently strong.

Acknowledgments

I would like to thank participants at the Tilburg conference and other forums for their comments, in particular Laura Benua, Orhan Orgun, and Paul Smolensky, as well as Rolf Noyer and Don Ringe for later discussions.

Directionality Constraints on Derivation?

Matthew Y. Chen

1. Introduction

Classical generative phonology conventionally prescribes an ordered set of operations that produce the desired output, often requiring long derivations with many intermediate representations. This 'processual' model has been faulted variously on the grounds of descriptive adequacy, psychological plausibility and computational tractability (cf. McCarthy 1993; Fukui 1996; and contributions to Goldsmith (ed.) 1993). In a radical break from this tradition, Optimality Theory (OT) "shifts the burden from the theory of operations (Gen) to the theory of well-formedness (H-eval)" (Prince and Smolensky 1993: 4). In other words, OT cares about the end result or *what* is (the output), but not *how* it came to be what it is (derivational steps leading to the output). While OT has succeeded spectacularly in not merely describing but explaining an impressive range of linguistic phenomena without resorting to serial derivation, the editors of this volume pose the intriguing question as to whether or not there remains a residue of linguistic facts that can be described only — or at least, most insightfully — in derivational terms.

I will attempt to answer this question in the affirmative by detailing certain tone sandhi facts of the Chinese dialect of Tianjin. I will first present and motivate an account which crucially relies on constraints that hold not on outputs per se, but on the derivations that lead to the outputs (sections 2–3). Inasmuch as these constraints are processual in nature, they do not readily translate into classical, monostratal OT terms (Section 4). Technically speaking, it is possible to restate the facts in a two-level OT, extending somewhat the "correspondence theory of faithfulness" (McCarthy and Prince 1995), but in doing so one runs into some conceptual problems and obscures certain insights implicit in the derivational account (Section 5).

It is worth noting that Prince and Smolensky (1993: 79) made allowance for a "serial" version of OT, and McCarthy (1996) pointed out that it would be a mistake to think of OT as inherently non-derivational. In this light, constraints on competing derivations are not necessarily incompatible with the spirit, if not the current practice, of OT: if we construe OT as a proposal about evaluating some set of candidates with respect to some set of soft constraints, where candidates could be either representations or derivations.[1]

2. A Derivational Account

The northern Mandarin dialect spoken in Tianjin (a major metropolis, located about 100 km southeast of the capital city of Beijing) has the familiar 4-tone system shown in (1).[2] In connected speech, tonal strings must obey a number of constraints. Specifically, OCP (2a) bars the juxtaposition of identical tones such as LL, RR or FF (H, L, R, F = high, low, rising, falling). I follow Yip (1989) in assuming that H is unmarked, and underlyingly unspecified; a HH sequence, therefore, does not incur OCP violation. OCP′ (2b) goes one step further, and bans certain adjacent *partially* identical tones. Since a falling tone is decomposable into a sequence of H and L, a F+L sequence consists in effect of HL+L, with two abutting L's, a tonal string that is disfavored by the more restrictive OCP′. OCP and OCP′ are enforced by the rules of Dissimilation and Tonal Absorption stated in (3). Tonal Absorption, restated as (4) has analogs in a number of African languages, including Bamileke, Mende, Kikuyu, Hausa, Ngizim (cf. Hyman and Schuh 1974).

(1) *Tianjin Tones*:

11	55	24	53	(1 = low, 5 = high)[3]
L	H	R	F	(symbols used here, for *low*, *high*, *rising*, *falling* tones)

(2) a. OCP
a sequence of adjacent like tones is not allowed (except HH)
b. OCP′ (partial OCP)
the sequence FL (= H<u>L.L</u>) is not allowed

---

1. Cf. Paul Smolensky (p.c.) and David Pesetsky's posting on OT-net, September 21, 1995.

2. I take Li and Liu (1985) as the primary source of the Tianjin data discussed below. For additional data and discussions, see Chen (1986, 1987), Hung (1987) and Tan (1987).

3. Consistent with the notational convention introduced by Chao (1930), numbers indicate values on a 5-point pitch scale (5 = high, 1 = low)

(3) DISSIMILATION

a. LL → RL

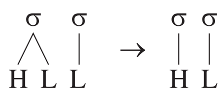
b. RR → HR

c. FF → LF

TONAL ABSORPTION

FL → HL

(4)



The effects of Dissimilation and Tonal Absorption are illustrated by the examples of (5)

(5)

	base form	*sandhi form*		
a.	LL	RL	guan xin	‘to be concerned’
b.	RR	HR	hen hao	‘very good’
c.	FF	LF	lang fei	‘wasteful’
d.	FL	HL	kan shu	‘to read a book’

The above sketch sums up in a nutshell all the relevant facts about disyllabic tone sandhi. However, longer tonal strings pose a non-trivial problem regarding the mode of rule implementation. Limiting ourselves to trisyllabic strings, we have 64 (4^3) logically possible 3-tone combinations. Of these 7 patterns (P1–7 of Table 1) constitute the focus of our interest, since they alone are potentially subject to more than one instance of rule application.

Table 1

	Input	Output	[x x] x	x [x x]	[x x x]
P1.	FFL	LHL	[*si-ji*] *qing* 'evergreen'	*zuo* [*dian-che*] 'take a tram'	
P2.	RRR	HHR	[*li-fa*] *suo* 'barber shop'	*mu* [*lao-hu*] 'tigress'	[*ma-zu-ka*] 'mazurka'
P3.	FFF	HLF	[*su-liao*] *bu* 'plastic cloth'	*ya* [*re-dai*] 'sub-tropical'	[*yi-da-li*] 'Italy'
P4.	LLL	LRL	[*tuo-la*] *ji* 'tractor'	*kai* [*fei ji*] 'pilot a plane'	
P5.	RLL	HRL	[*bao-wen*] *bei* 'thermo-cup'	*da* [*guan-qiang*] 'speak in a bureacratic tone'	
P6.	LFF	RLF	[*wen-du*] *ji* 'thermometer'	*tong* [*dian-hua*] 'make a phone call'	
P7.	FLL	FRL	[*lu-yin*] *ji* 'cassette recorder'	*shang* [*fei-ji*] 'board an airplane'	

[…] indicate morphosyntactic constituency

Take P1 of Table 1. In (6a), scanning the tonal string from left to right, Dissimilation-c first changes FFX to LFX (where X = any arbitrary string); moving further to the right, Tonal Absorption converts XFL to XHL.[4] Applied from right to left Tonal Absorption changes XFL into XHL, thereby preempting or, to use the conventional terminology, bleeding Dissimilation-c (= the derivation in (6b)).

(6)	P1	a.	FFL	(left to right)
			LFL	Dissimilation-c
			LHL	Tonal Absorption (right output)
		b.	FFL	(right to left)
			FHL	Tonal Absorption
			FHL	n/a (wrong output)

It should be clear that the rule implementation problem is not reducible to rule ordering. Take P2 (7). The three-tone sequence involves only one single rule,

4. Both Dissimilation and Tonal Absorption scan a two-tone substring for possible rule application. I use underlining to highlight this two-tone window of scansion.

that of Dissimilation-b; hence the question of rule ordering cannot in principle arise. Yet, depending on the direction in which Dissimilation-b iterates, the process yields different outputs.

(7)	P2	a.	RRR	(left to right)
			HRR	Dissimilation-b
			HHR	Dissimilation-b (correct)
		b.	RRR	(right to left)
			RHR	Dissimilation-b
			RHR	n/a (wrong)

The problem would be trivial, if the directionality of rule application followed from the cyclic principle, which would predict that rules operate in a left-to-right direction on a left branching construction, and in the opposite direction on a right-branching construction. But the prediction is patently counterfactual: only a left-to-right application of phonological rule(s) produces the right sandhi forms for both P1 and P2, regardless of morphosyntactic structure. In fact, in the case of polysyllabic words of foreign origin like *ma.zu.ka* ‘mazurka’, there is no internal structure to speak of. Nor can we simply stipulate that rules iterate in a consistently left-to-right or right-to-left direction. The fact of the matter is that rules must apply from left to right in P1–2, but in the opposite direction in P3–7. Consider P3. Only a right-to-left application generates the correct output (8b). Morphosyntactic structure or, for that matter, the lack thereof plays no role in determining the sandhi form, at least as far as trisyllabic constructions, whether lexical or phrasal, are concerned.

(8)	P3	a.	FFF	(left to right)
			LFF	Dissimilation-c
			LLF	Dissimilation-c (wrong)
		b.	FFF	(right to left)
			FLF	Dissimilation-c
			HLF	Tonal Absorption (correct)

What then determines the directionality of rule application? The most obvious observational generalization can be stated as (9).

(9) By default rules apply from left to right — unless such a mode of application produces an illformed output, in which case reverse the direction of operation.

Unpacked in OT terms, (9) consists of two hierarchically ranked constraints: WFC » Temporal, as stated in (10).

(10) WFC
Output must obey OCP and OCP′

TEMPORAL
Apply rules left to right
WFC » TEMPORAL

Notice that for both P1 and P2 (derivations (6a) and (7a) respectively), a left to right rule application yields perfectly well-formed outputs, i.e. sandhi forms that satisfy both OCP and OCP′. On the other hand, the same directional application as dictated by TEMPORAL would yield an ill-formed output *[LLF] in (8a), where the LL substring instantiates an OCP violation. In precisely cases like this, the high ranking WFC imposes a right to left direction — even at the cost of violating TEMPORAL sequencing. Consequently, the output of (8b) rather than (8a) emerges as the attested sandhi form for P3. The alternative modes of rule application and their outputs are summarized in the tableau in (11), where violations of WFC or TEMPORAL are marked with an asterisk (*) in the usual manner. Notice that the ranking of WFC » TEMPORAL alone picks out the correct, attested, sandhi forms for each of P1–6 in the tableau in (11).

(11)

					WFC	PREEMPT	TEMPORAL
P1	☞	a.	⇒	FFL → LFL → LHL			
		b.	⇐	FFL → FHL		*!	*
P2	☞	a.	⇒	RRR → HRR → HHR			
		b.	⇐	RRR → RHR			*!
P3		a.	⇒	FFF → LFF → LLF	*!		
	☞	b.	⇐	FFF → FLF → HLF			*
P4		a.	⇒	LLL → RLL → RRL	*!		
	☞	b.	⇐	LLL → LRL			*
P5		a.	⇒	RLL → RRL	*!		
	☞	b.	⇐	RLL → RRL → HRL			*
P6		a.	⇒	LFF → LLF	*!		
	☞	b.	⇐	LFF → LLF → RLF			*
P7		a.	⇒	FLL → HLL → HRL		*!	
	☞	b.	⇐	FLL → FRL			*

Key: ⇒, ⇐ indicate directionality of rule application

P7 requires a further refinement. As demonstrated in (12), by applying the relevant rules in either direction, we would produce a perfectly wellformed sandhi form — that is to say, an output that respects both OCP and OCP′. WFC, therefore, is neutral between (12a) and (12b). Consequently, we expect TEMPORAL to pick a default left to right mode of rule application resulting in *[HRL] as the expected sandhi form. Surprisingly, it is [FRL] (12b) that is the attested reading for a P7 expression like [lu yin] ji 'cassette recorder' and shang [fei ji] 'to board an airplane'.

(12) P7 a. FLL (left to right)
HLL Tonal Absorption
HRL Dissimilation-a (wrong)

b. FLL		(right to left)
	FRL	Dissimilation-a
	FRL	Tonal Absorption does not apply (correct)

One might suspect some sort of *Derivational Economy* at work: both (12a) and (12b) produce a well-formed target, but (12b) achieves the goal with greater economy, deriving the output [FRL] in one single step. In other words, one could posit the ranking {WFC, ECONOMY} » TEMPORAL. But such a move would lead to a ranking paradox: in order to insure that (7a = P2a) is picked over (7b = P2b), we must stipulate exactly the reverse constraint hierarchy, since neither output violates either OCP or OCP′, and yet it is (7a) that prevails. Notice that (7a) satisfies TEMPORAL (left to right application) at the expense of Derivational Economy (two derivational steps instead of one). I will simply posit the preemptive clause formulated as (13),[5] to be ranked above TEMPORAL.

(13) PREEMPT
Dissimilation takes precedence over all other rules.
{WFC, PREEMPT} » TEMPORAL

A highly ranked PREEMPT correctly eliminates P7a of the tableau in (11), leaving P7b as the winner candidate, despite its contrarian directionality of serial rule application.

But the left to right or right to left modes of rule application do not exhaust the logical possibilities. There is a third option: rules apply consistently left to right, backtracking where necessary to produce a wellformed string — as illustrated in the tableau given in (14).

5. See the next section for a discussion on the intuitive content of PREEMPT.

(14)

					WFC	NoBt	Temporal
P3		a.	⇒	FFF → LFF → LLF	*!		
	☞	b.	⇐	FFF → FLF			*
		c.	BT	FFF → LFF → LLF → RLF		*!	
P4		a.	⇒	LLL → RLL → RRL	*!		
	☞	b.	⇐	LLL → LRL			*
		c.	BT	LLL → RLL → RRL → HRL		*!	

Key: BT = backtracking

To exclude candidates (c) of P3 and P4 in the tableau in (14), one needs to explicitly posit a ban on backtracking (15). The overall ranking stated in (16) encapsulates the derivational account.

(15) NoBacktracking (NoBt)
Do not backtrack

(16) {WFC, Preempt, NoBt} » Temporal

3. Constraints on Derivation?

The derivational account sketched in Section 2 makes crucial use of three constraints that are distinctly derivational or processual in character: 1. Extrinsic order (stated as Preempt, (13)); 2. Left to right directionality (formulated as Temporal (sequence), (10)); 3. NoBacktracking, (15). Before rejecting them out of hand, consider their a priori plausibility and empirical support.

Extrinsic order is something of a last resort in any account. If pushed, one might speculate that Preempt reflects a gradient robustness of OCP effects: when a tonal string simultaneously violates OCP (total identity) and OCP′ (partial identity), the tendency is to undo the more egregious OCP offense (by Dissimilation) before worrying about the minor OCP′ violation (that can be fixed by Tonal Absorption). The parochial nature of OCP′ is illustrated in (17), which shows that partial identities are freely tolerated elsewhere. (Partial identities are underlined.)

(17)	a.	FL = HL.L	*, → [H.L] by Tonal Absorption
	b.	FR = HL.LH	ok
	c.	LR = L.LH	ok
	d.	RH = LH.H	ok

The bias for the left to right directionality accords with common sense. For instance, the parsing of syllables into feet is predominantly left to right (cf. Hayes 1995a). This is certainly so in Chinese, where footing is diagnosed by means of tonal distribution (18). A right-to-left processing would require buffering of long stretches of speech (cf. Levelt 1989). For psycholinguistic evidence showing a left-to-right bias in speech organization (phonological encoding), see Meyers (1990, 1991).

(18) *Shanghai tone/stress domain (Duanmu 1993, 1995)*

a.	HL	LH	HL	LH	HL	
	gao	er	ba	qiao	fu	'Gorbachev'
	(HL	.LH)	(HL	.LH)	HL	left to right footing (left-headed); no degenerate foot
	(HL	. o)	(HL.	o)	o	tone deletion (affecting stressless syllables)
	(H L)		(H	L)	o	tone association (left to right), ok

b.	HL	LH	HL	LH.	HL	
	gao	er	ba	qiao	fu	
	HL(LH.HL)(LH				.HL)	right to left footing
	o (LH. o)(LH				. o)	tone deletion
	o (L H)(L H)					tone association, *

As for NoBacktracking, virtually all sentence processing models (from both production and comprehension end) assume a left-to-right, incremental parsing of materials as soon as they are heard (rather than waiting until the end of the sentence). This occasionally gives rise to the classic garden-path phenomenon, whereby the hearer is misled into committing him/herself to a default analysis until the surprise ending, at which point s/he has to *backtrack* and reparse (for recent surveys, cf. Pritchett 1992; Clifton, Frazier and Rayner 1994; Tanenhaus and Trueswell 1995). Backtracking represents a particularly complex processing task. A phonological analog (outside of the Tianjin case) would be the English Rhythm Rule (19), which requires backtracking at step (c).

(19) thirteen abstract paintings
- a. thirTEEN abSTRACT
- b. thirTEEN ABstract PAINtings
- c. THIRteen ABstract PAINtings (xx = current window)

(Stressed syllables are capitalized.) However, empirical evidence for the speaker's ability to perform stress retraction on-line proves to be elusive (cf. Cooper and Eady 1986; Kelly and Bock 1988; Levelt 1989; Beckman et al. 1990; for a possible explanation, see Hayes 1995a). On the other hand, MalakMalak stress clash resolution clearly shows the avoidance of backtracking. Goldsmith's analysis of the relevant facts is paraphrased in (20).

(20) *MalakMalak (Birk 1976; Goldsmith 1990: 173–7)*
- a. Group syllables into left-headed feet, from right to left (weak prohibition on degenerate feet)
- b. Word-level prominence falls on the leftmost stressed syllable
- c. Stress clash resolution:
  - i. *either* apply 'trochaic reversal' (restricted to *one step* in the repair)
  - ii. *or* delete the degenerate foot

A right-to-left parsing of syllables into trochees of necessity gives rise to stress clash in polysyllabic words containing an odd number of syllables, as in [(x)(x.)] and [(x)(x.)(x.)] (where x marks a stressed syllable). Stress-clash resolution employs two alternative strategies: (i) 'trochaic reversal' (turning a trochee into an iamb) clearly violates the otherwise general left-headedness constraint; on the other hand, (ii) destressing or deletion of a degenerate foot leaves a syllable unfooted, in violation of PARSE-σ. Which strategy prevails depends on the 'one step' repair principle, which is basically a derivational constraint. Cast in OT terms, example 2 of the tableau in (21) demonstrates the ranking PARSE-σ » LEFT-HEADEDNESS; but example 3 calls for the reverse ranking. Hence the ranking paradox observed in the tableau given in (21). Note in particular that multiple violation of LEFT-HEADEDNESS is irrelevant under the strict domination hypothesis.

(21)

					*CLASH	PARSE-σ	LFT-HD
1	/mu.nan.ka.ra/	☞		(″MU.nan)(′KA.ra)			
2	/mɛl.pa.pu/		a.	(″MɛL)(′PA.pu)	*!		
			b.	mɛl.(″PA.pu)		*!	
		☞	c.	(″MɛL)(pa.′PU)			*
3	/ar.ki.ni.yaŋ.ka/		a.	(″AR)(′KI.ni)(′YAŋ.ka)	*!		
			b.	(″AR)(ki.′NI)(′YAŋ.ka)	*!		*
		☜	c.	(″AR)(ki.′NI)(yaŋ.′KA)			**
		☞	d.	ar.(″KI.ni)(′YAŋ.ka)		*!	

Constraints:
PARSE-σ: Parse syllables into feet
LFT-HD: Form left-headed feet

Symbols:
″σ primary stress
′σ secondary stress
☜ expected winner

Restated in the derivational terms of (22), candidate b (= tableau 21, 3c) is eliminated because it entails backtracking symbolized by the current window (underlined xx) first moving leftwards (direction of footing), then rightwards (direction of clash resolution).

(22) a. /mɛl.pa.pu/
→ mɛl.(′PA.pu) R-to-L footing
→ (″MɛL).(′PA.pu) R-to-L footing, *CLASH
(″MɛL).(pa.′PU) Trochaic reversal (1)
b. /ar.ki.ni.yaŋ.ka/
→ ar.ki.ni.(′YAŋ.ka) R-to-L footing
→ ar.(′KI.ni)(′YAŋ.ka) R-to-L footing
→ (″AR)(′KI.ni)(′YAŋ.ka) R-to-L footing, *CLASH
→ (″AR)(ki.′NI)(′YAŋ.ka) Trochaic reversal (1), *CLASH
→ (″AR)(ki.′NI)(yaŋ.′KA) Trochaic reversal (2)

In summary, there is at least prima facie plausibility to our account that appeals to constraints of a derivational nature, namely TEMPORAL, NOBACKTRACKING and PREEMPT.

4. An Output-Driven Account

Can we reconceptualize and recast the derivational analysis given above in declarative, output-driven terms? Dissimilation and Absorption can be reformulated as output constraints, if we extend the notion of I/O-correspondence not only to prescribe identity (perfect match), but also to circumscribe the range of permissible deviations or alternations (23).[6] (23) says that an input L may alternate with (equivalently, correspond to) R, and F with either L or H, etc. ALTERNATION (23) in conjunction with OCP and OCP″ (an extended version of OCP′ that penalizes *all* partial identities)[7] correctly picks the winner candidate in two-tone combinations, as shown in the tableau in (24). Note in particular that while it is possible to replace the OCP-offending input LL with either HL (candidate 3b), or RL (candidate 3c), only the latter is consistent with the ALTERNATION constraint (L can alternate with R but not with H). Hence, RL prevails.

(23) ALTERNATION: L ~ R ('~' = alternates with)
R ~ H
F ~ L,H

---

6. (23) is kindred in spirit to the notion of "feasible pairs" in two-level phonology (Koskenniemi 1983, cf. Orhan Orgun, lecture notes Winter 1996, UCSD). Thus L may pair with or correspond to R, but not to H etc. For a more elaborate reformulation of Dissimilation and Tonal Absorption employing exclusively output constraints, see Chen (in press).

7. Thus a string like FR (= HL.LH) is consistent with both OCP and OCP′, but barred by OCP″, since it contains a substring HL.LH.

(24)

					OCP	OCP′	OCP″	ALTERN
1	FL		a.	FL (= HL.L)		* !	*	
			b.	LL	*!			
		☞	c.	HL				
2	FF		a.	FF	*!			
		☞	b.	LF (= L.HL)				
			c.	HF (= H.HL)			*!	
3	LL		a.	LL	*!			
			b.	HL				*!
		☞	c.	RL				
4	RR		a.	RR	*!			
		☞	b.	HR				
			c.	LR (= L.LH)			*!	*
			d.	FR (= HL.LH)			*!	*

While it is reasonably simple to replicate the effects of Dissimilation and Tonal Absorption rules, it is not as straightforward to capture the derivational constraints by non-derivational means. We will consider four possibilities. First, OT can mimic directionality effects via *Alignment* (McCarthy and Prince 1993a). Thus, a left-to-right (Pintupi) or right-to-left grouping of syllables into feet (Yakan) can be recast in terms of aligning the metrical unit to the left or the right edge of a prosodic word (Crowhurst and Hewitt 1995; cf. Hayes 1995a).[8] Likewise, directional syllabification readily translates into alignment between the syllable boundary and one or the other of the word edges (see Mester and Padgett 1993; Davis 1995). However, unlike footing and syllabification, tone sandhi does not create structures with constituent edges to align with some reference point.

Second, output conditions, i.e. WFC, underdetermine the choice of winner

8. Assuming no degenerate feet. If degenerate feet are constructed at the edges, then the direction of alignment would have to be reversed.

candidates: for every input there are at least two possible wellformed outputs, depending on the directionality of rule application, as illustrated in the tableau in (25). In particular P7a [HRL] is perfectly wellformed, in fact the attested output corresponding to three different inputs (26).

(25)

			WFC	ALTERN	FAITH
P7		/FLL/	*!*		
	a.	HRL			**!
☞	b.	FRL			*
P4		/LLL/	*!*		
	a.	RRL	*!		**
☞	b.	LRL			*
	c.	HRL		*!	**
P1		/FFL/	*!*		
☞	a.	LHL			**!
☜	b.	FHL			*
P2		/RRR/	*!*		
☞	a.	HHR			**!
☜	b.	RHR			*

(26) /HRL/ = [HRL] [niu jiao] jian 'splitting hairs' (lit. tip of a horn)
/HLL/ → [HRL] pi [shu bao] 'leather briefcase'
/RRL/ → [HRL] lao [mu ji] 'old hen'

Third, *Faithfulness*: pick the most faithful candidate consistent with WFC — faithfulness being measured by any tonal substitution. This works for P4 and P7 of the tableau in (25), but makes counterfactual predictions for P1 and P2. Finally, we may plausibly hypothesize that Tianjin picks whichever direction of rule application that results in the least marked tonal sequences — assuming, quite reasonably, that contour tones (R and F) are more complex, and therefore more marked than level tones (H and L). Markedness correctly picks P3b, P1a, P2a, but makes the wrong choice in P7, as illustrated in the tableau given in (27).

In short, there is no obvious way to recast derivational constraints in terms of optimal output configurations or faithfulness to the input.

(27)

			WFC	ALTERN	MARK	FAITH
P3		/FFF/	*!*			
	a.	LLF	*!			
☞	b.	HLF			*	**
	c.	RLF		*!	**	**
P1		/FFL/	*!*			
☞	a.	LHL				**
	b.	FHL			*!	
P2		/RRR/	*!*			
☞	a.	HHR			*	**
	b.	RHR			**!	*
P7		/FLL/	*!*			
☜	a.	HRL			*	**
☞	b.	FRL			**!	*

5. Cross-level Constraints

Let us explore another approach. Technically, it is possible to derive the desired outputs by means of declarative statements of correspondences along the lines of two-level rules developed by Koskenniemi (1983), Karttunen (1993), Kaplan and Kay (1994), Orgun (1995), and McCarthy and Prince (1995), *inter alia*. The two-level rules of (28) are interpreted as follows.[9] For clarity, input and output strings are represented on the top and the bottom lines respectively. Vertical lines symbolize legitimate correspondences. Contextual conditions are specified either at the input (top) and/or the output (bottom). Thus, R1 reads: input L corre-

9. I owe the formulation of (28) to Orhan Orgun.

sponds to output R, if it is followed by another L on the right at the surface. R3b states, on the other hand, that R corresponds to H if it is followed by another R at the input or underlying representation. The net effect of R3b taken together with R3a is that RR is paired with HR, whether this RR string originates at lexical representation or arises from other rules/correspondences. Finally, R4b imposes a conjunctive condition: F alternates with L iff it is followed by an F at input *and* a H at output.

(28) *Two-level Rules*

```
R1. L        R2. F        R3a.   R      R4a.   F
    |            |               |             |
    R L          H L             H R           L F

                          R3b.   R R    R4b.   F F
                                 |             |
                                 H             L H
```

Table 2 recasts the three alternative directional applications of tone sandhi rules in terms of two-level rules. The legitimate and illegitimate input/output pairings in accordance with the two-level rules formulated in (28) are symbolized by ‘|’ and ‘ ’ respectively. The three columns headed by ⇒, ⇐ and BT correspond to the outputs generated by the three traditional modes of rule application in the derivational account: left to right, right to left, and left to right with backtracking.

Table 2

<table>
<tr><th></th><th></th><th>⇒</th><th></th><th>⇐</th><th></th><th>⇒, BT</th></tr>
<tr><td>P1</td><td>a.
☞</td><td>FFL
| |
LHL</td><td>b.</td><td>FFL
|
FHL</td><td></td><td></td></tr>
<tr><td>P2</td><td>a.
☞</td><td>RRR
| |
HHR</td><td>b.</td><td>RRR
|
RHR</td><td></td><td></td></tr>
<tr><td>P3</td><td>a.</td><td>FFF
|
LLF</td><td>b.
☞</td><td>FFF
| |
HLF</td><td>c.</td><td>FFF
|
RLF</td></tr>
<tr><td>P4</td><td>a.</td><td>LLL
|
RRL</td><td>b.
☞</td><td>LLL
|
LRL</td><td>c.</td><td>LLL
|
HRL</td></tr>
<tr><td>P5</td><td>a.</td><td>RLL
|
RRL</td><td>b.
☞</td><td>RLL
| |
HRL</td><td>c. = b.</td><td></td></tr>
<tr><td>P6</td><td>a.</td><td>LFF
|
LLF</td><td>b.
☞</td><td>LFF
| |
RL</td><td>c. = b.</td><td></td></tr>
<tr><td>P7</td><td>a.</td><td>FLL
|
HRL</td><td>b.
☞</td><td>FLL
|
FRL</td><td></td><td></td></tr>
</table>

|, symbolize legitimate and illegitimate pairings respectively

As Table 2 demonstrates, the two-level rules of (28) consistently pick the correct output, while rejecting the alternatives as containing one of more illegitimate input/output pairings. Given the input string FFL (P1), an output like FHL (P1b generated by the right to left application of Dissimilation and Tonal Absorption) is rejected because it contains an illegitimate pairing of F:F, in contravention to R4b, which requires an F:L correspondence in that particular context. On the other hand, an output like LHL (P1a) generated by a left to right rule application contains all and only pairings licensed by (28), and is therefore the correct, attested, sandhi form.

While the two-level rules approach succeeds in eliminating the need for directional rule application and derivational steps, it poses serious conceptual and theoretical problems. First, given the nature of the sandhi processes (with context on the right), one can mimic directionality by imposing the condition on the context-sensitive rules either on the input or on the output: (a) a condition on the *input* forces a left-to-right rule application: as long as the requisite condition is met at the input level, rules apply regardless of what happens (‘subsequently’) to the context; (b) conversely, by imposing the condition on the output, one forces a right-to-left mode of implementation: since one needs to ‘anticipate’ what eventually happens to the right context before deciding whether it licenses a particular correspondence. This is illustrated in (29) and (30), with conditions highlighted in boldface.

(29)					⇒			⇐	
	P2	RRR	☞	a.	HHR		b.	RHR	Rule 3b
	P4	LLL		a.	RRL	☞	b.	LRL	Rule 1

(30)	R3b	R **R**	R1	L
		ǀ		ǀ
		H		R **L**

Since the hallmark of two-level rules is their ability to stipulate at will conditions on the input and/or output, the implication is that directionality is a rule-specific idiosyncracy. By contrast, in a derivational account directionality follows from the ranking WFC » TEMPORAL SEQUENCE, as shown in (31):

(31)	P2	⇒	RRR	→	HRR	→	HHR	= ok
	P4	⇒	LLL	→	RLL	→	RRL	= * (output = ill-formed)
		⇐	LLL	→	LRL			= ok

Second, rule ordering and opacity effects: in a derivational account, Dissimilation precedes Absorption by virtue of PREEMPT (13), as illustrated in (32a); in other words, Absorption ‘counterbleeds’ Dissimilation.

(32)	a.	FFL		b.	FFL	
		LFL	Dissimilation		FHL	Absorption
		LHL	Absorption		n/a	Dissimilation

In general, two-level rules handle opaque relations by stipulating conditions on the input (to signal the fact that a correspondence/rule is not ‘surface-true’). But one cannot simply restate R4b as R4c (as in (33), simply because it wrongly predicts P3 /FFF/ will emerge as [LLF] (34). Instead, R4b must define the condition jointly on the input *and* the output. Clearly, R4b is a notational variant of R4d (35), since the only context in which a tone is F at input but H at output is where this F is followed by a L (by virtue of R2 formulated in (28)). R4d brings into focus three types of conceptual problems: (a) One forfeits the locality condition, requiring a non-contiguous L as the environment.[10] (b) R4d redundantly repeats R2–the classic argument for rule ordering in conventional rule-based phonology. (c) R4d telescopes two separate, elementary processes R2 and R4a into one complex correspondence. Given R2 and R4a, there is no need for R4b — if R2 and R4a are allowed to interact in some principled fashion (in this case, consistent with TEMPORAL). In short, R4b, functionally equivalent to R4d, forces into one single complex and totally opaque correspondence two elementary processes each of which is transparent and motivated taken separately.[11] In this sense, two-level rules are curiously anti-analytical and non-explanatory.

```
(33)   R4c.   F    F
              |
              L

(34)  F    F  F
      |    |  |
      L    L  F    = *

(35)   R4b.   F  F      =    R4d.  F  F  L
              |                    |
              L  H                 L
```

Needless to say, one can dispense with the ungainly R4b/d of (35) in a three-level model (Goldsmith 1993; Lakoff 1993) by stipulating that R2 is a W:P rule (while all other rules function as M:W constraints),[12] in effect mimicking rule ordering (Dissimilation precedes Absorption). This approach is illustrated in (36). The obvious objection here is that there is no independent motivation for this level separation.[13] Specifically, both R4 and R2 apply indifferently at lexical

10. For a recent discussion on the locality condition, see Archangeli and Pulleyblank (1994: 26f and *passim*)

11. The argument here is the flipside of McCarthy's (1993) objection against breaking up functionally related chain shifts into formally distinct rules as illustrated by the following example from Hijazi Bedouin Arabic:

R1 i → Ø
R2 a → i

For further discussion of similar chain shifts, see Orgun (1995a) and Kirchner (1995).

12. M, W, P refer to morphophonemic (underlying), word-level and phonetic representations (see Goldsmith 1993a).

13. This is precisely the same objection raised by Padgett (1995).

and phrasal levels as demonstrated by (37).

(36) *Three-level rules*

	P7		P1		P3	
M:	F L L		F F L		F F F	
	ǀ	R1	ǀ	R4	ǀ	R4
W:	F R L		L F L		F L F	
				R2	ǀ	R2
P:	F R L		L H L		H L F	

(37)

	R4 FF → LF	R2 FL → HL
Lexical	*lang fei* 'wasteful'	*jiao shi* 'teacher'
Phrasal	*song xin* 'to deliver a letter'	*diao gu* 'to catch a turtle'

6. Concluding Remarks

It appears that the tone sandhi facts of Tianjin resist both a two-level rule style account as well as a standard monostratal OT analysis. A natural and explanatory description needs to appeal to such derivational constraints as TEMPORAL, NOBACKTRACKING and PREEMPT (which amounts to extrinsic order). It seems unlikely that the Tianjin case merely constitutes an oddity rather than pointing to certain systematic aspects of language that are best described in dynamic, processual terms.

I conclude with a remark on one of the major arguments against serialist derivation, namely the dubious status of intermediate representations such a (38b) given below. In (38) and (39), parentheses enclose the prosodic entities p-word and p-phrase, successively larger domains on which tone sandhi rules operate (p-word level and p-phrase level operations are separated by a broken line in (38–39)). Points of interest (current window of scansion) are underlined. In a derivational account, Dissimilation and Absorption operating on the two separate p-words produce (38a) as the output, which is attested. (38a), however, contains a substring of two L's. Therefore, Dissimilation operating on the larger domain of p-phrase generates the intermediate form (38b), which is less than ideal, containing as it does two abutting R's, thereby triggering another round of

Dissimilation, with (38c) as the eventual output. Unfortunately the intermediate form (38b) is not attested, at least not in the published sources (Li and Liu 1985; Chen 1986, 1987; Hung 1987 and Tan 1987). This raises the question whether intermediate forms such as (38b) are not merely an artefact of the serial, rule-based analysis. Without addressing the validity of this argument in its generalized form, I whish to point out that in the particular case at hand the absence of intermediate forms like (38b) may well be an accidental gap. Similar theoretically postulated intermediate forms do occasionally surface. One such instance is shown in (39b), which survives with an offending [...RR...] substring, presumably because the ideally expected (39c) involves two successive steps of backtracking. Attested intermediate forms like (39b) give us a 'candid camera' snapshot of an on-line processing task that is aborted by excessive backtracking. This lends a certain sense of reality to the processual metaphor.

(38) 'older-brother smokes 'Zhangdou' (brand) cigarettes'

```
        dage chou    zhangdou yan
        (F L L)w     (F F L)w
           |            |
        (F R L)      (L F L)
                       |
   a.   (F R L)      (L H L)   attested
  ----------------------------------
        (F R L       L H L)IP
             |
   b.   (F R R       L H L)    not attested
           |
   c.   (F H R       L H L)    attested
```

w = phonological word

IP = intonational phrase

(39) ‘telephone emits strange noise’

	dianhua	*fachu*	*guai jiao*	
	(F F)w	(L L)w	(F F)w	
	|	|	|	
a.	(L F)	(R L)	(L F)	attested
	(L F	R L	L F)$_{IP}$	
		|		
b.	(L F	R R	L F)	attested
		|		
c.	(L F	H R	L F)	not attested

Acknowledgments

I have benefited greatly from discussing various aspects of this paper with Chris Barker, Gary S. Dell, Jeff Elman, John Moore, Orhan Orgun, Jaye Padgett, David Perlmutter, Paul Smolensky, David Swinney and Moira Yip. My thanks also go to the audiences at UC San Diego, UC Irvine and Tilburg University. Usual disclaimers apply.

Alignment and the Cycle are Different

San Duanmu

1. Introduction

This paper discusses whether *Alignment constraints* (McCarthy and Prince 1993a) can replace the cycle and achieve a one-step analysis of what used to require multiple steps. I argue that the answer is no. The reason is that the cycle has two properties, (a) using information on morphosyntactic bracketing and (b) preserving phonological structures built on a previous cycle, whereas a one-step analysis with Alignment can achieve (a) but not always (b). I also show that given a cyclic analysis the preservation of previously built structures is obtainable without assuming additional constraints. Finally, I will discuss a way in which cyclicity can be seen to be compatible with a one-step analysis.

For the reason of space I will not compare the present approach to cyclicity with others presented in this volume. In addition, subsequent development on the topic, such as Kenstowicz (1996), will be discussed either very briefly or not at all (see Duanmu 1997a for more discussion). Still, the claim of this paper remains valid, I believe, namely, Alignment cannot replace the cycle. Another purpose of this paper is to provide most cyclic data in Shanghai Chinese, which can be used by those who want to examine them in a different framework.

It is well-known that syntactic boundaries can affect phonology. This phenomenon is traditionally captured by the cycle (e.g. Chomsky and Halle 1968). Recent works in Optimality Theory (Prince and Smolensky 1993, hereafter OT) suggest that some cyclic effects can be handled by Alignment, according to which edges of syntactic units must coincide with those of phonological units (McCarthy and Prince 1993a). As a result, what used to be analyzed in several steps with the cycle can now be analyzed in one step (e.g. Cohn and McCarthy 1994 and Kenstowicz 1994). The question I address here is: Can ALIGN always do the work of the cycle? I will use data from compound stress in

Shanghai Chinese (hereafter Shanghai), a dialect spoken by the majority of people in Shanghai City. For the analysis of cyclicity, Shanghai is more interesting to look at than many other languages, such as Indonesian. For example, in Indonesian we know the stress locations but not the direction of the foot (see Halle and Idsardi 1994 vs. Cohn and McCarthy 1994 and Kenstowicz 1994), but in Shanghai both the location of stress and the direction of the foot are unambiguous. In addition, in Indonesian only the stem and suffixes affect stress, but a stem can only take at most two suffixes. In Shanghai, on the other hand, the levels and the kinds of embedding are very rich, and theoretically unlimited. I will show that the analysis with ALIGN cannot always predict the correct result. Instead, stress must be analyzed cyclically. In other words, ALIGN cannot do the work of the cycle. I will also discuss how the preservation of previously built structures is achieved in a cyclic analysis within OT. Finally, I will offer an interpretation of the cycle within a one step analysis.

2. Compound Stress in Shanghai

My data come from Xu et al. (1988), Duanmu (1995), and the native speakers I consulted. Metrical and tonal properties of Shanghai compounds are discussed in Duanmu (1995). Here I will only examine nominal compounds, which constitute the majority of all compounds. First, I offer some background on Chinese compounds and how stress is determined in Shanghai.

2.1 *Nominal compounds in Chinese*

Chinese has two nominal structures, [M N] and [M *de* N], where N is a noun, M a modifier, and *de* a particle. An example of each from Shanghai is given in (1), where [ge′] is the Shanghai pronunciation of *de* (transcribed in phonetic symbols without tones; ignoring number, which is not marked in Chinese; [′] indicates a glottalized vowel)

(1) a. ko se
tall mountain
'tall mountain'
b. ko ge′ se
tall DE mountain
'tall mountain'

All researchers agree that [M *de* N] is a phrase (e.g. Fan 1958; Chao 1968).

However, the analysis of [M N] is somewhat controversial. The [M N] in (1a) is semantically compositional, as its English counterpart, which is usually considered a noun phrase. In addition, to many speakers (1a) is synonymous to (1b). Moreover, although stress can sometimes distinguish a compound from a phrase in English, the same is not obvious in Chinese. Thus, some linguists consider both (1a) and (1b) phrases (e.g. Zhang 1992), with an optional deletion of the particle *de*. There is, however, compelling evidence that [M N] and [M *de* N] are syntactically different. In particular, [M N] is not a phrase but a compound. In addition, recursive derivatives of [M N], such as [M [M N]] and [[M N] N], are also compounds. The arguments can be found in Dai (1992) and Duanmu (1997b), whose view I will adopt here.

2.2 *Tone and Stress in Shanghai*

Shanghai has five phonetic pitch patterns on isolated syllables. The five patterns correlate with onset voicing and rime glottalization. It is possible, therefore, to posit just two underlying patterns, LH and HL. In a polysyllabic domain the tonal pattern is solely determined by the initial syllable: when the initial syllable is LH, the domain pattern is [L H L…L], and when the initial syllable is HL, the domain pattern is [H L…L] (a third domain pattern will be discussed shortly). In other words, the surface tones of the first two syllables are the same as the underlying tones of the initial syllable, and all other syllables get L. (2) gives some examples, where underlying syllable tones are shown above surface tones; [z] is a syllabic fricative, [~] indicates nasalization, and syllables in a polysyllabic word are separated by a dash for visual clarity.

(2)	HL LH LH	LH LH LH	HL HL LH	LH HL LH
	H L L	L H L	H L L	L H L
	ko vã- tsz	wã vã- tsz	çĩ kõ- ts^{h}ã	lo kõ- ts^{h}ã
	‘tall house’	‘yellow house’	‘new factory’	‘old factory’

Although native speakers find it hard to tell stress in Shanghai (Selkirk and Shen 1990: 315), Duanmu (1995) has argued that Shanghai has left-headed feet. First, we have seen that the underlying tones of the initial syllable are preserved, and those of noninitial syllables are lost. Since it is common in Chinese languages for unstressed syllables to lose their underlying tones, the preservation of tones from the initial syllable supports left-headed stress in Shanghai.

Second, polysyllabic foreign words form disyllabic tonal domains, indicated by parentheses in (3), which suggests binary foot construction.

(3) LH LH HL LH LH LH
(L H) (H L) (L H)
dʒe′- k^{h}a′- sz- lu- va′- k^{h}a′
‘Czechoslovakia’

The underlying tones of a foreign word are those of the characters that are used to represent it. In extremely careful, character-by-character speech, the underlying tones can surface on each syllable. However, in normal speech only tones from foot initial syllables are preserved.

Third, there is an asymmetry between [1 2] and [2 1] compounds (digits indicate the number of syllables in a word). [1 2] always forms one domain, as in (4), but [2 1] can form either one or two domains, as in (5).

(4) HL LH LH
(H L L) *(HL) (L H)
tçi ts^{h}z- pã tçi ts^{h}z- pã
‘chicken wing’

(5) LH HL HL
(L H L) or (L H)(HL)
lu- sõ t^{h}ã lu- sõ t^{h}ã
‘Russian soup’

Their difference predicted if stress is left-headed, so that in [1 2] there is stress clash, but in [2 1] there is not. This is shown in (6).

(6) x x x x
(x) (x x) (x x)(x)
[1 2] [2 1]

The optionality of the second foot in [2 1] will be discussed below.

Having considered the locations of stress, we turn to the locations of metrical boundaries. Since stress is left-headed, the left boundary must be before the initial syllable. But what about the right boundary? For the example in (7), there are three possibilities.

(7) LH LH LH
L H L x x x
wã vã- tsz (x)x x (x x) x (x x x)
‘yellow house’ a b c

It is easy to reject (7a): it has a monosyllabic foot, and it raises the question of how the second syllable can get its surface tone from the first syllable across a foot boundary. Between (7b) and (7c) the choice is less obvious. One may suggest that, other things being equal, (7b) is preferred, since it has a binary foot, which is more common than a ternary foot. On the other hand, there is evidence for (7c). Besides the two tonal patterns discussed so far, there is a third, which occurs when the initial syllable is underlyingly LH, with a voiced onset and a glottalized vowel (shown by [′]). In this case the first syllable is L, the last syllable H, and the intermediate syllable(s) L (when the domain is four or more syllables, this pattern becomes optional). (8) shows some examples, where (8d) is a made-up compound.

(8) a. LH LH LH
(L L H)
lo′ vã- tsz
'green house'

b. LH HL HL
(L L H)
lo′- se- tçi
'Los Angeles'

c. LH LH LH
(L L H)
[ba′ [bi ɦa]]
'white leather shoe'

d. LH LH LH HL LH LH LH
(L L H) (H L) (L H)
ba′ dʒe′- k^ha′- sz- lu- va′- k^ha
'White Czechoslovakia'

How tonal domains are determined will be discussed below. The point of interest here is that in the special pattern the H moves all the way to the end of the domain: it can land at the end of the current word, as in (8b), or in the middle of another word, as in (8d), or travel through one (or more) words, as in (8c). If we assume that tone movement takes place within a foot, then the special pattern suggests that the size of a tonal domain is the same as a foot.

2.3 *Single Words*

We now consider domain formation in Shanghai. Single words with even numbered syllables are shown in (9) and those with odd numbered syllables are shown in (10).

(9)

HL	LH				
(H	L)				
pa-	li				
‘Paris’					
HL	LH	HL	LH		
(H	L)	(H	L)		
ya-	lu-	sa-	lã		
‘Jerusalem’					
LH	LH	HL	LH	LH	LH
(L	H)	(H	L)	(L	H)
dʒe′-	k^ha′-	sz-	lu-	va′-	k^ha′
‘Czechoslovakia’					

(10)

LH				
(LH)				
mo				
‘horse’				
LH	HL	HL		
(L	L	H)		
lo′-	se-	tçi		
‘Los Angeles’				
HL	LH	LH	LH	LH
(H	L)	(L	H	L)
ka-	li-	fo′-	ɲi-	ya
‘California’				

It can be seen that Shanghai cannot be analyzed as a pitch accent language where H falls on the head of the foot, or we would get irregular foot patterns as well as stress clashes. In traditional terms, single words show left-to-right construction of left-headed binary feet. In OT terms, (9) shows (at least) two constraints, LEFT-HEADEDFOOT and FOOTBINARITY (FTBIN). LEFT-HEADEDFOOT is never violated in Shanghai and so will be ignored for now. I will return to FTBIN below.

Now consider (10). First consider the monosyllabic case. Although there is just one syllable, it still forms a foot. In OT terms there is a constraint PARSE-SYL, which is ranked above FTBIN and which ensures that every word is metrified. Next consider 3-syllable and 5-syllable words. It is clear that the final syllable does not form a foot by itself, but it is not obvious whether (a) the final syllable is unfooted or (b) it is inside a trisyllabic foot. There is no evidence for (a) beyond the fact that Shanghai has binary feet. On the other hand, there is some evidence for (b). As seen in (8), in the special tone pattern the H goes all the way to the third syllable in a trisyllabic unit, instead of stopping at the second syllable. Since tonal domains are determined by stress in Shanghai, it will be natural to assume that a tonal domain both starts and ends at a foot boundary. The choice between (a) and (b) will obviously affect one's analysis in some ways. However, for the present purpose, namely, whether the cycle is needed in OT, the choice is not consequential. I will therefore assume (b) in the discussion below. Following Kenstowicz (1994), I posit the constraints in (11).

(11) PARSE-SYL
All syllables should be metrified.

FTBIN
A foot should contain exactly two elements.

ALIGN-FT-L
Align the left side of a foot with the left side of a lexical word.

Since every syllable is footed, PARSE-SYL is ranked highest. Some researchers have proposed that FTBIN should be split into two constraints, such as 'At Least Two' and 'At Most Two' (e.g. Hewitt 1994; Green 1995). Since that proposal does not affect our discussion, I use the simpler definition of FTBIN in (11). ALIGN-FT-L reflects left-to-right foot construction, exemplified in (12), where S = syllable and ( ) = foot boundaries.

(12)

	ALIGN-FT-L
a. (SSS)(SS)	#, #**!*
b. ☞(SS)(SSS)	#, #**

ALIGN-FT-L tallies for each foot how many syllables away it is from the left edge of a word (#) (see Kenstowicz 1994, and references cited there). The first foot in both (12a) and (12b) is zero syllables away from the left word boundary; the second foot is three syllables away in (12a) but two syllables away in (12b). Thus (12) correctly predicts the domain pattern for 5-syllable words.

Next consider the ranking between FTBIN and ALIGN-FT-L. Clearly FTBIN should rank above ALIGN-FT-L. If not, all single words should form just one foot, since additional feet will violate ALIGN-FT-L. This is shown in (13) and (14), where ☜ indicates a wrongly predicted best candidate.

(13)

/SSSS/	ALIGN-FT-L	FTBIN
☜(SSSS)	#	**
(SS)(SS)	#, #*!*	

(14)

/SSSS/	FTBIN	ALIGN-FT-L
(SSSS)	*!*	#
☞(SS)(SS)		#, #**

Following a suggestion by Michael Kenstowicz (p.c.), I assume that FTBIN is gradiently evaluated. It gives an asterisk for each extra syllable in a foot; it also gives an asterisk for a monosyllabic foot for being a syllable short (see below). In (SSSS) there are two extra syllables, giving two asterisks. Ranking FTBIN above ALIGN-FT-L gives the correct result in (14). I summarize the ranking so far in (15), along with the analyses in (16)–(20).

(15) PARSE-SYL » FTBIN » ALIGN-FT-L

(16)

/S/	PARSE-SYL	FTBIN	ALIGN-FT-L
☞(S)		*	#
S	*!		

(17)

/SS/	PARSE-SYL	FTBIN	ALIGN-FT-L
☞(SS)			#
(S)(S)		*!*	#,#*

(18)

/SSS/	PARSE-SYL	FTBIN	ALIGN-FT-L
☞(SSS)		*	#
(SS)(S)		*	#,#*!*
(S)(SS)		*	#,#*!

(19)

/SSSS/	Parse-Syl	FtBin	Align-Ft-L
(SSSS)		*!*	#
☞(SS)(SS)			#,#*!*

(20)

/SSSSS/	Parse-Syl	FtBin	Align-Ft-L
(SSSSS)		**!*	#
☞(SS)(SSS)		*	#,#**
(S)(SS)(SS)		*	#,#*,#**!*

In (16) there are two candidates, a footed syllable and an unfooted one. The footed syllable has one asterisk under FtBin, since it is a syllable short; the unfooted syllable has blanks under FtBin and Align because the two constraints are met vacuously. The other analyses are straightforward. They all give the correct results.

2.4 *Two-word Compounds*

Let us now look at two-word compounds. First, consider [2 3] and [3 2]. If compounds behave like single words, both [2 3] and [3 2] should form (SS)(SSS), as a 5-syllable morpheme does. On the other hand, if compound internal word boundaries play a role, then [2 3] and [3 2] can differ from a 5-syllable word, and/or from each other. The facts in (21) and (22) show that while [2 3] forms (SS)(SSS), [3 2] forms (SSS)(SS) (in the examples below, word for word gloss will be omitted when it is obvious).

(21) LH LH LH HL LH
(L H) (L H L)
[ze′ -pẽ mu- se- k^ha′]
‘Japanese mosaic’

(22) HL LH LH HL HL
(H L L) (H L) *(H L) (L H L)
[pa- na- ma çã- tço] [pa-na- ma çã- tço]
‘Panama banana’

Evidently, the intermediate word boundary plays a role. In the traditional analysis (21) and (22) are handled by the cycle, whereby each word is analyzed separately first, before the entire compound is analyzed. Crucially, feet built on the first cycle are preserved on the second (unless a serious offense occurs, such as stress clash, to be discussed later). In the OT analysis the same effect can be achieved by applying ALIGN-FT-L to each component word. This is shown in (23) and (24), where the internal word boundary is indicated by #.

(23)

/SS#SSS/	PARSE-SYL	FTBIN	ALIGN-FT-L
☞(SS)#(SSS)		*	#
(SS#S)(SS)		*	#,#*!

(24)

/SSS#SS/	PARSE-SYL	FTBIN	ALIGN-FT-L
(SS)(S#SS)		*	#,#*!*
☞(SSS)#(SS)		*	#,#

Both analyses are correct. Next consider [2 1], which can form either (SS)#(S) or (SS#S), the former being used in careful speech. This is shown in (25) (again, [z] can be syllabic).

(25)
```
LH LH   LH
(L   H  ) (LH)      or    (L   H  L)
[zã- he    z  ]           [zã- he z]
'Shanghai City'
```

Both patterns are correctly predicted in the present analysis, as shown in (26).

(26)

/SS#S/	PARSE-SYL	FTBIN	ALIGN-FT-L
☞(SS)#(S)		*	#,#
☞(SS#S)		*	#

The fact that in careful speech (SS)#(S) is preferred to (SS#S) reflects a general tendency for there to be more domains in slower speech than in faster speech (see Cheng 1973, who notes a similar phenomenon in Mandarin). A further example is shown in (27) (the first pattern of (27b) is the special spreading pattern).

(27) a. *Careful speech*
LH LH LH LH
(L H) (L H)
[voʹ- de [da ɦoʹ]]
[Fudan [big school]]
'Fudan University'
b. *Casual/fast speech*
(L L L H) (L H L L)
[voʹ-de [da ɦoʹ]] or [voʹ-de [da ɦó]]

In careful speech [2 [1 1]] form (SS)#(S#S), but in casual speech (SS)#(S#S) and (SS#S#S) can both be used. In our analysis (SS)#(S#S) is a better form, as shown in (28).

(28)

/SS#S#S/	PARSE-SYL	FTBIN	ALIGN-FT-L
☞(SS)#(S#S)			#,#
(SS#S#S)		*!*	#

The fact that (SS#S#S) can be used in faster speech means that speed or casualness can affect constraint rankings, such as demoting FTBIN in to a lesser role.

Let us now consider [1 2]. Unlike [2 1], which has two good output patterns, (S#SS) is the only one for [1 2], shown in (29).

(29) LH LH HL
(L H L) *(LH)(L H)
[nø zo- çi] [nø zo- çi]
'South Korea'

The lack of (S)#(SS) in [1 2] is unexpected, since according to the analysis developed so far both (S#SS) and (S)#(SS) should be good, as shown in (30).

(30)

/S#SS/	PARSE-SYL	FTBIN	ALIGN-FT-L
(S#S)(S)		*	#,#*!
☞(S)#(SS)		*	#,#
☞(S#SS)		*	#

It will be noted, though, that (S)#(SS) contains a stress clash but (SS)#(S) does not. To account for the difference between [1 2] and [2 1], we can propose a constraint *CLASH, given in (31), which will favor (S#SS) over (S)#(SS) for [1 2], as shown in (32) (′ = stress).

(31) *CLASH
Avoid stresses on adjacent syllables.

(32)

/S#SS/	*CLASH
(′S)#(′SS)	*!
☞(′S#SS)	

Since *CLASH is never violated in Shanghai compounds, it must rank above FTBIN. Now consider [1 1], which only has (S#S). This is expected, as shown in (33) and (34).

(33)

HL HL	
(H L)	*(HL)(HL)
[çi ko]	[çi ko]
west melon	
'watermelon'	

(34)

/S#S/	PARSE-SYL	*CLASH	FTBIN	ALIGN-FT-L
(′S)#(′S)		*!	**	#,#
☞(′S#S)				#

The candidate (S)#(S) violates both FTBIN and *CLASH, so it is bad. (S#S) violates no constraint and is good.

2.5 *Cyclic Cases*

We have analyzed some compounds in one step. I now discuss cases that cannot be analyzed in one step. First, consider [1 4] and [1 5]. In normal speech, the only good pattern for [1 4] is (S#SS)(SS) and that for [1 5] is (S#SS)(SSS). This is shown in (35) and (36).

(35) LH HL LH HL LH
(L H L) (H L)
[nø ya- lu- sa- lã]
'South Jerusalem'
*(L H) (L H L)
[nø ya- lu- sa- lã]

(36) LH LH LH LH HL LH
(L H L) (L H L)
[nø yĩ- du- ɲi- çi- ya]
'Southern Indonesia'
*(L H) (L H) (H L)
[nø yĩ- du- ɲi- çi- ya]

However, in our analysis so far the predictions are different. For [1 4] (S#S)(SSS) should be the best, as shown in (37), and for [1 5] (S#S)(SS)(SS) should be the best, as shown in (38).

(37)

	/S#SSSS/	FtBin	Align-Ft-L
☞	(S#S)(SSS)	*	#,#*
	(S#SS)(SS)	*	#, #*!*

(38)

	/S#SSSSS/	FtBin	Align-Ft-L
☞	(S#S)(SS)(SS)		#,#*,#***
	(S#SS)(SSS)	*!*	#, #**

In these examples the cyclic analysis and the Alignment analysis differ. Consider [1 4] first. In the cyclic analysis [1] and [4] are analyzed separately on the first cycle, giving (S) and (SS)(SS). On the second cycle, all feet built on the first cycle are preserved unless there is a serious offense, which there is, namely, the stress clash between the first two syllables. The repair, however, will be limited to the removal of the stress from the second syllable, which causes the first two feet to merge into one. The third foot will be left intact. The same applies to [1 5] in (38), and [1 6] in (8d), which forms (S#SS)(SS)(SS)) where clash on the second cycle leads to the merger of the first two feet without altering other feet.

In the noncyclic analysis so far Align is the only tool for achieving the cyclic effect. However, Align is successful when it is satisfied at the inter-word boundary, such as in [3 2]. When it is not, as in [1 4] and [1 5], where it is overridden by *Clash, Align has no effect in preserving the foot patterns of the component words. As a result, [1 4] and [1 5] should pattern like a single word and form (SS)(SSS) and (SS)(SS)(SS) respectively, but neither does. The analysis in terms of Align runs into problems with [1 3] as well. The only good pattern for [1 3] is (S#SSS), as shown in (39), but the Align analysis expects (S#S)(SS) to be better, as shown in (40).

(39) LH LH LH HL
(L L L H)
[ba′ ŋu- lu- sz]
'white Russia'
*(LH) (L H L)
[ba′ ŋu- lu- sz]
*(L H) (L H)
[ba′ ŋu- lu- sz]

(40)

/S#SSS/	FtBin	Align-Ft-L
☞ (S#S)(SS)		#,#*
(S#SSS)	*!*	#

The problem with a one-step Align analysis is noted by Kenstowicz (1994), who suggests that [1 n] requires a two-step analysis. On the first step the second word is metrified first. On the second step the entire compound is analyzed. In addition, there is a constraint Overwrite to preserve feet built on the first step.

Kenstowicz's proposal raises two questions. First, why should the second word be analyzed before the first word? Second, will more complicated compounds need more steps? An alternative to the two-step analysis is to assume the cycle, where the number of steps is not limited to two. An example that bears on this issue is [1 [1 2]]. In the cyclic analysis, the inner [1 2] is analyzed before the entire compound, giving (SSS). The input to the final cycle will then be [1 3], which gives (SSSS). [1 [1 2]] compounds are not common. One reason is that when a compound gets long, one tends to break it up by inserting the particle [ge′], the Shanghai equivalent of Mandarin *de*. This tendency is especially strong when monosyllables are repeatedly added on the left. Nevertheless, for the examples that are available, such as (41), the only pattern for [1 [1 2]] is (SSSS), as the cyclic analysis predicts.

(41) LH LH LH LH
(L H L L)
[ço [ɦõ lo- bo′]]
[small [red turnip]]
'small carrot'
*(L H) (L H)
[ço [ɦõ lo- bo′]]

For 'small carrot' one would prefer to use the phrase [ço ge′ ɦõ lo-bo′] , with the particle [ge′], which forms two domains (ço ge′)(ɦõ lo-bo′) (for some linguists, such as Sproat and Shih 1991, this means 'carrots that are small'). But [ço ɦõ lo-bo′] can appear in places like a price board at a market (along with [du ɦõ lo-bo′] 'large carrot' for example), and in such cases one must use the pattern (SSSS); the pattern (SS)(SS) is bad for [1 [1 2]]. Similarly, consider [1 [1 [1 1]]]. In the cyclic analysis the innermost [1 1] forms (SS). On the next cycle [1 [1 1]] is like [1 2], which forms (SSS). This in turn makes the final cycle similar to [1 3], which forms (SSSS). Thus, (SSSS) is the only pattern for [1 [1 [1 1]]]. The cyclic prediction is correct, as shown in (42).

(42) LH LH HL LH
(L L L H)
[ba′ [ya [t^hi ŋu]]]
[white [wild [sky goose]]]
'white wild swan'
*(L H) (H L)
[ba′ [ya [t^hi ŋu]]]

Again, in the normal case one would prefer to use the phrasal form [ba′ ge′ ya t^hi ŋu] 'wild swans that are white', which forms (ba′ ge′)(ya t^hi ŋu). But if one has to use the compound form without the particle [ge′], such as when reading a bird's name in a zoo, then the only good pattern is (SSSS). Like [1 n], [1 [1 2]] and [1 [1 [1 1]]] are problems for the one-step analysis, which wrongly predicts (SS)(SS) to be the best form for both, as shown in (43) and (44).

(43)

	[1 [1 2]]	FtBin	Align-Ft-L
☞	(S#S)#(SS)		#,#
	(S#S#SS)	*!*	#

(44)

[1 [1[1 1]]]	FtBin	Align-Ft-L
☞ (S#S)#(S#S)		#,#
(S#S#S#S)	*!*	#

It should be obvious that such examples will be very difficult to analyze in a two-step analysis as well. For example, in the two-step analysis, it was assumed that in a two-word compound the second word was metrified first. In [1 [1 2]] the second word is the middle [1]. There is no reason why it should be analyzed first. Besides, analyzing the middle [1] first has no use. The reason is that, although on the second step the first [1] will be forced to merge with the second [1], the third word [2] is free to form a foot by itself. Thus, one wrongly predicts (SS)(SS) to be the best pattern. What is needed, instead, is for [1 2] to form (SSS) on the first step, so that on the second step the input is similar to [1 3], which will form (SSSS) (see below). This can be achieved by assuming that it is the right-hand branch of a compound, instead of the second word, that is analyzed first.

Leaving certain questions aside, such as why it is not the left-hand branch that is analyzed first, or why the [1 [1 1]] in [1 [1 [1 1]]] forms (SSS) instead of (S)(SS), let us see whether this interpretation of the two-step analysis will work. Consider first [2 [1 [1 [1 1]]]], where [1 [1 [1 1]]] is the right-hand branch. In the two-step analysis the right-hand branch is analyzed in one step, so [1 [1 [1 1]]] should form (SS)(SS). On the second step the first word forms a binary foot. Since there are no violations, the output should be (SS)(SS)(SS). This prediction, however, is incorrect. The actual pattern is (SS)(SSSS), as shown in (45).

(45)
```
 LH  LH  LH  LH  HL  LH
(L   H)  (L   L   L   H  )
[ze′- pẽ  [ba′ [ya  [tʰi  ŋu]]]]
[Japan [white [wild [sky goose]]]]
'Japanese white wild swan'
*(L   H) (L   H) (H  L)
[ze′- pẽ [ba′ [ya [tʰi ŋu]]]]
```

In order for the right-hand branch to form (SSSS), it is necessary for it to be analyzed in two steps, as discussed earlier. But then there will be three steps for the entire compound, instead of two. Similarly, the compound [2 [1 [1 2]]] is predicted to form (SS)(SS)(SS) in the two-step analysis, whereas it in fact forms (SS)(SSSS), as shown in (46).

(46) LH LH LH LH LH LH
(L H) (L H L L)
[ze′- pẽ [ço [ɦõ lo- bo′]]]
[Japan [small [red turnip]]]
'Japanese small carrot'
*(L H) (L H) (L H)
[ze′- pẽ [ço [ɦõ lo- bo′]]]

To get the correct result there need to be two steps just for the right-hand branch. Next consider [[1 [1 [1 1]]] [1 1]]. In this compound the right-hand branch is [1 1], which will form (SS). If the left-hand branch is not analyzed on the first step, then on the second step it should form (SS)(SS), and the entire compound should be (SS)(SS)(SS). But this again is incorrect. The correct pattern is (SSSS)(SS), as shown in (47).

(47) LH LH HL LH LH LH
(L L L H) (L H)
[[ba′ [ya [t^hi ŋu]]] [ve ti]]
[[white [wild [sky goose]]] [meal store]]
'White Wild Swan Restaurant'
*(L H) (H L) (L H)
[[ba′ [ya [t^hi ŋu]]] [ve ti]]

In order for the left-hand branch to form (SSSS), it should itself require two steps before analyzing the entire compound. Similarly, in [[1 [1 2]] [1 1]] there need to be two steps just for the left-hand branch, which forms (SSSS) and not (SS)(SS), as shown in (48).

(48) (LH LH LH LH LH LH
(L H L L) (L H)
[[ço [ɦõ lo -bo′]] [ve ti]]
[[small [red turnip]] [meal store]]
'Small Carrot Restaurant'
* (L H) (L H) (L H)
[[ço [ɦõ lo- bo′]] [ve ti]]

Clearly, then, both branches must be analyzed before the entire compound, and each branch may need more than one step. This is exactly what the cycle assumes. Further examples, such as [[1 1] [1 5]] and [[1 5][1 1]], illustrate the same point, as shown in (49) (a made-up geographic location) and (50).

(49) (LH LH LH LH LH LH HL LH
(L H) (L H L) (L H L)
[[ɲe′- ta] [nø yĩ- du- ɲi- çi- ya]]
[[warm belt] [south Indonesia]]
'Tropical Southern Indonesia'
* (L H) (L H) (L H) (H L)
[[ɲe′- ta] [nø yĩ- du- ɲi- çi- ya]]

(50) LH LH LH LH HL LH LH LH
(L H L) (L H L) (L H)
[[nø yĩ- du- ɲi- çi- ya] [ve ti]]
[[south Indonesia] [meal store]]
'Southern Indonesia Restaurant'
* (L H) (L H) (H L) (L H)
[[nø yĩ- du- ɲi- çi- ya] [ve ti]]

In (49) the right-hand branch [1 5] needs two steps. In (50) the left-hand branch needs two steps. This again shows that each branch of a compound needs to be analyzed in advance and each branch may need more than one step. Neither the one-step analysis nor the two-step analysis can account for such cases in a straightforward way.

2.6 *Identity Constraints*

The main problem with the ALIGN analysis is the preservation of previously built feet. In particular, preservation is available if ALIGN is observed, but not if ALIGN is violated. This problem is avoided with the use of *identity constraints* (McCarthy and Prince 1995; McCarthy 1996b), which preserves the structural similarity between two representations. For example, [1 n] compounds can be analyzed as in (51).

(51)	a.	/Wd/ word input	→ ID-A	[Wd] word output
	b.	[Wd1] [Wd2] compound input (word output)	→ ID-B	[Wd1 Wd2] compound output

There are two steps. First, each word is analyzed under certain identity constraints (ID-A) (along with other constraints, such as LEFT-HEADEDFOOT and FTBIN). Second, the output of single words forms the input to a compound under certain identity constraints (ID-B). For the sake of argument, let us ignore ID-A and focus on ID-B. In addition, for ID-B let us focus on just one constraint, STRESS-ID, stated in (52).

(52) STRESS-ID
The locations of stress in the input and the output must be identical.

STRESS-ID is a gradient constraint that tallies the number of syllables on which stress has changed. STRESS-ID must be ranked below *CLASH, otherwise the [1] in [1 n] will remain a monosyllabic foot. Let us now consider the analysis of [1 5], shown in (53), where S = stressed syllable, s = unstressed syllable.

(53)

(S), (Ss)(Sss)	*CLASH	STRESS-ID	FTBIN
a. (S)(Ss)(Sss)	*!		**
b. (Ss)(Ss)(Ss)		**!**	
c. ☞(Sss)(Sss)		*	**

The input (to the compound) has two separately metrified (surface) words, (S) and (Ss)(Sss). In (53a) all the input stresses are preserved, which leads to a violation of *CLASH between the first two syllables. In (53b) there are four violations of STRESS-ID: the second and the fourth syllable lost stress, and the third and the fifth syllables gained stress. In (53c) only the second syllable lost stress, incurring one violation of STRESS-ID. Even though (53c) incurs two violations of FTBIN, it is still the best output. This shows that STRESS-ID must be ranked above FTBIN.

(53) reminds one of the two-step analysis discussed earlier. One must ask whether more steps are needed for more complicated compounds. Consider a two-step analysis of [1 [1 2]], using STRESS-ID, as shown in (54).

(54)

(S), (S), (Ss)	*CLASH	STRESS-ID	FTBIN
a. ☞(Ss)(Ss)		*	
b. (Ssss)		**!	**

In (54a) there is one violation of STRESS-ID and no violation of FTBIN. In (54b) there are two violations of STRESS-ID and two violations of FTBIN. Thus, (54a) is predicted to be better, yet (54b) is the real output. Thus, with STRESS-ID, the two-step analysis still makes the wrong prediction. The problem lies in the fact that the inner [1 2] forms a closer unit, which gives (Sss). If the inner unit can be analyzed before the entire compound, the correct result can be obtained. The same applies to other compounds involving three or more words.

2.7 *A Cyclic OT analysis*

In this section I show that a cyclic analysis using STRESS-ID will give correct results. Let us first consider [1 [1 2]]. On the first cycle, each word is analyzed separately, giving (S), (S), and (Ss). Next, [1 2] is analyzed, shown in (55).

(55)

(S), (Ss)	*CLASH	STRESS-ID	FTBIN
a. (S)(Ss)	*!		*
b. ☞(Sss)		*	*

Since *CLASH is ranked above STRESS-ID, [1 2] gives (Sss). Finally, the entire compound is analyzed, which is shown in (56).

(56)

(S), (Sss)	*CLASH	STRESS-ID	FTBIN
a. (Ss)(Ss)		**!	
b. ☞(Ssss)		*	**

In (56a) there are two violations of STRESS-ID: the second syllable lost stress and the third gained stress. In (56b) there is one violation of STRESS-ID. Since STRESS-ID is ranked above FTBIN, (56b) is predicted to be better, which is correct. The same analysis will give correct results to all two-word compounds. As examples, [1 1] and [2 1] are shown in (57) and (58).

(57)

(S), (S)	*CLASH	STRESS-ID	FTBIN
a. (S)(S)	*!		**
b. ☞(Ss)		*	

(58)

(Ss), (S)	*Clash	Stress-Id	FtBin
a. (Sss)		*!	*
b. ☞(Ss)(S)			*

Next consider the three-word compounds [1 [1 1]] and [[1 1] 1]. Since the inner [1 1] forms (Ss), [1 [1 1]] is the same as [1 2], which forms (Sss), and [[1 1] 1] is the same as [2 1], which forms (Ss)(S). Both results are correct. Next, consider [1 [1 [1 1]]]. Since the inner [1 [1 1]] forms (Sss), [1 [1 [1 1]]] is like [1 3], which like (56) forms (Ssss). As final examples, we consider [[1 1] [1 5]] and [[1 5] [1 1]] in (59)–(60).

(59) [[1 1] [1 5]]
a. Inner brackets
[1 1] → (Ss)
[1 5] → (Sss)(Sss) same as (53)
b. Outer brackets

(Ss), (Sss)(Sss)	*Clash	Stress-Id	FtBin
a. ☞(Ss)(Sss)(Sss)			**
b. (Ss)(Ss)(Ss)(Ss)		*!**	

(60) [[1 5] [1 1]]
a. Inner brackets
[1 5] → (Sss)(Sss) same as (53)
[1 1] → (Ss)
b. Outer brackets

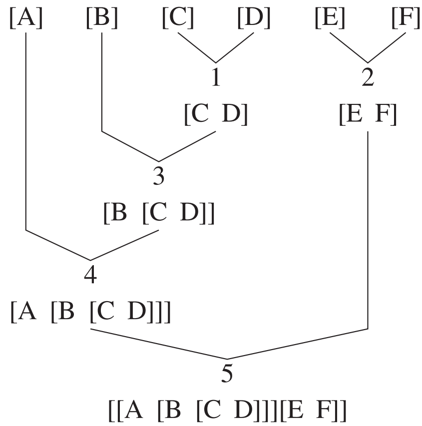
(Sss)(Sss), (Ss)	*Clash	Stress-Id	FtBin
a. ☞(Sss)(Ssss)(Ss)			**
b. (Ss)(Ss)(Ss)(Ss)		*!**	

In both cases the correct result is predicted.

3. Cyclicity, Base-Identity, and Parallelism

We have seen that in order to get the correct result, the relation of STRESS-ID (along with other constraints such as *CLASH and FTBIN) must hold cyclically. The more complex a compound is, the more cyclic steps there are. In the compound [[A [B [C D]]] [E F]], of which the structure [[1 [1 [1 1]]][1 1]] given in (47) is an example, there will be five STRESS-ID relations. They are numbered 1 through 5 in (61).

(61)



Relation 1 holds between the input [C] and [D] on the one hand and the output [C D] on the other. Relation 2 holds between the input [E] and [F] on the one hand and the output [E F] on the other. Relation 3 holds between the input [B] and [C D] on the one hand and the output [B [C D]] on the other. Relation 4 holds between the input [A] and [B [C D]] on the one hand and the output [A [B [C D]]] on the other. And the relation 5 holds between the input [A [B [C D]]] and [E F]] on the one hand and the output [[A [B [C D]]] [E F]] on the other. These five STRESS-ID relations correspond to five cyclic steps. This result appears to contradict *parallelism*, which, according to McCarthy and Prince (1995), is a fundamental idea of OT. Parallelism assumes that given an input, all output candidates are evaluated in one step. In the present analysis, however, a compound can require many cyclic steps — theoretically unlimited.

A solution to the conflict between cyclic steps and parallelism, I suggest, is to assume a panoramic view of all the steps in the evaluation of a compound with one snapshot. In particular, for the example in (61), one can assume that all the five steps are considered together, without assuming any sequential ordering among them. This proposal is similar to the idea of BASE–IDENTITY, discussed in Kenstowicz (1996), which Kenstowicz attributes to Edward Flemming.

(62) 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.

Both the present analysis and BASE-IDENTITY assume that syntactic structures are strictly binary branching (Kayne 1984, 1994). What (62) says, when applied to our case, is that STRESS-ID should hold between a compound and its two immediate constituents (i.e. not between a compound and the constituents of its immediate constituents). Although BASE-IDENTITY is given as a single relation, it is obvious that it must hold recursively (along with other constraints). For example, in [1 [1 2]], STRESS-ID must hold between [1 [1 2]] and its immediate constituents [1] and [1 2]. In the inner compound [1 2], STRESS-ID must hold between [1 2] and its constituents [1] and [2]. In other words, a single STRESS-ID relation holding directly between [1 [1 2]] on the one hand and [1], [1], and [2] on the other will not work, as we have seen in (54). Instead, we need two STRESS-ID relations for [1 [1 2]]. Similarly, for the compound in (61), we need five STRESS-ID relations, as required in the present analysis.

The idea of considering several cyclic steps simultaneously raises some technical questions. For example, in a cyclic analysis of (61) STRESS-ID is violated once in four of the five steps (1, 2, 3, and 4). Will the overall result be better, or worse, if all the STRESS-ID violations be concentrated in step 5, or if step 4 has 2 STRESS-ID violations and step 5 has another 2? And how do we decide this? However, such questions are beyond the scope of this paper.

4. Conclusions

I have argued that Alignment constraints cannot replace the cycle. Instead, the cycle should be retained in OT. It can be seen, too, that since the cycle already makes use of morphosyntactic boundaries, the need for Alignment constraints will be reduced. If the cycle is available in Optimality Theory anyway, there is little motivation to force a one-step analysis even when it seems possible, and especially when it comes at a cost. For example, Cohn and McCarthy (1994) and Kenstowicz (1994) showed that Indonesian word stress can be analyzed in one step, but only at the cost of admitting mixed foot types. A cyclic analysis can avoid positing mixed foot types with no additional cost. I have also suggested that if all cyclic steps can be considered simultaneously, cyclicity is compatible with parallelism.

Acknowledgments

I thank Michael Kenstowicz for discussions throughout the development of this paper, and the organizers and the participants at the Tilburg conference. I also thank Ben Herman and Marc van Oostendorp for their meticulous editorial work. An earlier version of this paper was presented at the ICCL-4/NACCL-7 conference in June 1996 at the University of Wisconsin Madison and I benefited from comments from the the audience. This paper was writeen in 1996. A revised and considerably expanded version appeared as Duanmu (1997a).

Stricture is Structure

Chris Golston Harry van der Hulst

1. Introduction

Most models of phonology assume that lexical representations involve a linear sequence of segments. Segments consist of a featural organization and some of the features (i.e. major class and certain stricture features) determine how a set of syllabification rules assign a syllabic organization to the string. In this article we propose a different view. We propose that lexical representations are syllabified in the lexicon and that there is no segmental representation per se. We present various types of linguistic and psycholinguistic evidence for this proposal. A consequence of our proposal is that syllable structure can be used to encode most major class[1] and manner properties. Hence, stricture is structure. We remove all stricture features from the 'segmental structure', reducing the latter to place features. Laryngeal features are represented as modifiers of the syllabic constituents onset, nucleus and coda, following Golston and Kehrein (1998, 1999).[2] Thus, if we view syllabic structure as the phonological counterpart of

---

1. Some major class features like [consonantal] may not be needed in phonology at all. See Hume and Odden (1996).

2. Jensen (1994) proposes a version of Government Phonology element theory that does not have stricture elements. Hence, in order to express lexical contrast between, for example, stops and fricatives, other means must be used. Jensen proposes to encode such contrasts in terms of extra syllabic structure. This approach thus identifies stricture with structure. Our approach is less radical since we, strictly speaking, maintain a set of stricture features, be it as labels of syllabic terminal nodes, rather than as features that associate to these and non-terminal nodes. By representing these as labels of syllabic terminal nodes we do, therefore, represent stricture features as direct extensions of the syllabic structure.

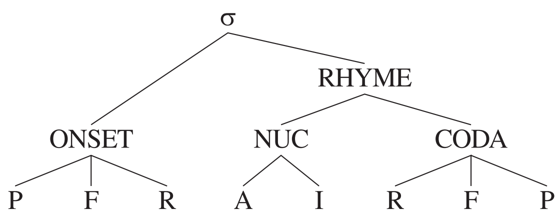
syntactic phrase structure and the place features as the phonological 'lexical items', the latter are essentially 'category-free', i.e. whether [labial] represents a consonantal or vocalic segment will depend on where it associates in the syllabic phrase structure.

Summarizing, we seek to demonstrate the following:

a. Stricture is part of syllable geometry.
b. Syllable structure is underlying.

A preview of the stricture model is given in (1). The general syllable template is meant to express that each syllabic constituent has one anchor for place features:

(1)

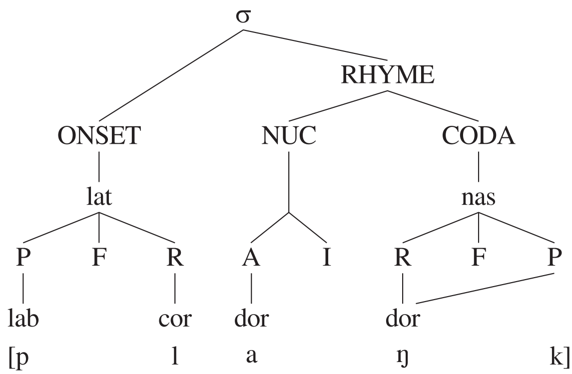


The capitals represent the labels of the syllabic terminal positions. In (1) P stands for *plosive*, F for *fricative*, R for *sonorant,* A for *low* or *open*, and I for *high* or *closed*. Thus, we regard aperture properties as nuclear stricture features.[3] Actual syllables result from attaching place, laryngeal and nasality features to a specific syllabic structure. A preliminary example makes clearer how this is to be done. Consider a representation for [plaŋk]:[4]

3. See Van der Hulst (1995, 1996, 1999, forthc.) for a different notational expression of the claims that are made in this article.

4. We overlook a number of other possibilities here for the sake of expediency. The final /k/ of this form, for instance, might be part of degenerate syllable, it might be extrametrical, and so on. These are issues we will not touch upon here because they do not affect our main point: that stricture is part of syllable structure.

(2)



The relevance of our proposal for the present volume is this: given claims A and B, one might argue that there is no need for a *process of syllabification*. We treat syllabification as a state of affairs, not as a derivational process. This rids generative phonology of what is perhaps its central derivational tenet: the derivation of surface representation from underlying representation through syllabification. If there is a derivational residue in phonology, syllabification need not be a part of it. Once distinctive laryngeal, place and manner features are linked into syllable structure, we know linear order, just as we know the linear order of a noun and a verb in a given language once we know that the former is the subject of the latter.

This article is organized as follows. Section 2 gives a general overview of lexical representation in generative phonology; it explains how we got where we are today and argues that we do not want to stay there. In Section 3 we develop the proposal that major class and major stricture features are part of syllable geometry and how this obviates the need for segments in phonological representation. In Section 4 we provide psycholinguistic and grammatical evidence for underlying syllabic organization. Section 5 concludes.

2. Background

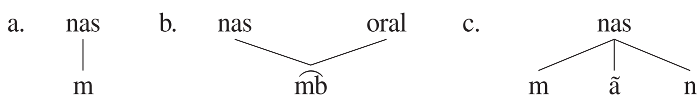
From the mid-sixties to the mid-seventies generative views on phonology were largely based on the theory behind Chomsky and Halle's *Sound Pattern of English* (SPE, Chomsky and Halle 1968). SPE did not recognize a syllabic organization in addition to the linear string of segments. In later developments of the theory, however, it was proposed that syllabic structure was necessary in order to capture significant phonological generalizations. This never led to the view that syllabic organization was present in the lexical representation. The assumption was — and continues to be — that syllabic organization is predictable on the basis of the segmental make-up and is therefore not part of underlying representations; syllabic structure was assigned by (early) rules.

Phonological theory has undergone a number of important changes since the mid-seventies. These changes involved structure both below and above the level of the segment. Sub-segmental changes led to the development of autosegmental phonology, supra-segmental changes led to the development of metrical phonology.

Autosegmental phonology differs from the SPE-conception in that phonological segments are not represented in terms of a feature bundle, but rather in terms of a skeletal point (i.e. essentially a syllabic position) to which a number of features associate. The features, as it has sometimes been put, form segments on their own ('autosegments'). Segmentation is done separately for each dimension of speech that is captured in a phonological feature category. Despite this broader use of the notion segment, the closest equivalent to the traditional notion is the skeletal point.

A general principle of licensing requires that all features be associated to a skeletal point in order to be realized phonetically. Within this conception, features and skeletal points may be related in a one-to-one, many-to-one, or one-to-many relation. In the unmarked case each point bears exactly one feature for tone, one for place, one for laryngeal specification and so on (cf. 3a). Complex segments have more than one such specification, (3b), per segment. Contours or prosodies have more than one segment per such specification, (3c):

(3)



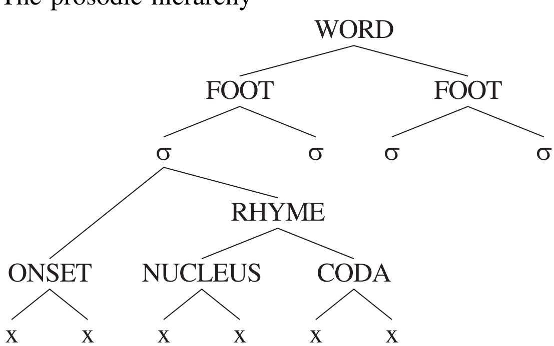
Note that this autosegmental treatment of complex segments such as affricates makes crucial use of linear ordering of specified features within the scope of single segments. We will argue that features can also associate to non-terminal syllabic nodes (such as Onset). Thus, we deprive the skeletal level of its special status, moving toward a non-segmental phonology. This move will imply other ways of representing complex segments, not involving linear information.

In certain cases the association of specified features to skeletal points is predictable, for example simply because *all* segments (of some type and within a certain domain) show up with this specified feature. In such cases the association can be removed from the lexical or underlying representations and be added by rule, usually in a left-to-right fashion. This approach has been especially succesful in the analysis of tone (Leben 1973; Goldsmith 1976; Pulleyblank 1983) but has also been used for consonantal and vocalic 'melodies' (McCarthy 1979), and individual vocalic features (Clements 1976; Pulleyblank 1988; Archangeli and Pulleyblank 1994). This marks a major departure from SPE conceptions of phonology because it is now allowed that the information contained in lexical entries is only partially linearized. We will argue that linear order can be suppressed more rigorously.

A further enrichment of phonological representations claimed that features are hierarchically organized, such that it is possible to refer to groups of features, as well as to individual features (Clements 1985, 1991; Sagey 1986; Den Dikken and Van der Hulst 1988; McCarthy 1990; Padgett 1991). This in turn led to a general recognition that stricture features (which define sonority) are in some sense more central than those which define place, oral/nasal or laryngeal specifications. Van der Hulst (1990, 1995, 1996, 1999, forthc.) argues that stricture (incl. major class) features are *head* features on which all other features are dependent. With respect to major class features this has also been argued by McCarthy (1988) whereas Steriade (1993, 1994a, 1995) suggests that many features are dependent on stricture. Our proposal below incorporates this particular dependency between manner and place in a principled way.

Metrical phonology is a theory of how skeletal points are hierarchically organized into prosodic structure: segments into syllables, syllables into feet, feet into phonological words, and so on:

(4) The prosodic hierarchy



Most metrical theories claim that prosodic constituents are headed, i.e. that within every constituent one unit is the head and the others are dependents. The most sonorant segment heads the nucleus, the nucleus heads the rhyme, the rhyme heads the syllable; one syllable heads each foot, one foot heads each word, and so on.

Despite all of these changes above and below the segment since 1968 the central role of the segment (X-slot, etc.) has gone unchanged, especially in underlying representation. This has led to what we might call complete prosodic underspecification:

(5) *Complete Prosodic Underspecification*
Underlying forms are not syllabified; surface forms are fully syllabified.

This results in an increased disparity between what types of object underlying and surface forms are. In SPE, both URs and SRs were the same types of object: strings of segments (no syllable structure at either level). With the advent of metrical phonology, a schism has developed between UR and SR such that they are completely different types of object. Golston (1996a) calls this the theorem of *impossibility*:

(6) *Impossibility*
Every underlying form is an impossible surface form and vice versa.

Impossibility is the result of having surface forms from the 1990s and underlying forms from the 1960s.

Impossibility should immediately call into question the psychological plausibility of Complete Prosodic Underspecification since, if understood as a psychologically real model, *impossibility* entails the claim in (7):

(7) Speakers cannot store what they can say and vice versa.

How or why a species would develop a communication system as inefficient as this is not clear to us. Why should something as fundamental as prosodic structure be left out of underlying form? The traditional answer is of course that prosodic affiliation is predictable given a linear string of segments. This is no doubt true and in any case not a point we would want to contest. Maintaining the idea that predictable properties are derived by the grammar, we wish to make the opposite claim: that the linear order of autosegments is predictable given a prosodic affiliation.

The point is hardly a new one and rests on two well-known universals of phonology, cf. especially Anderson (1987). The first is that the order of onset, nucleus and coda is fixed for all languages:

(8) *Universal left-right order of syllable constituents*
Syllable: Onset > Nucleus > Coda

The second is the familiar Sonority Sequencing Principle (Sievers 1881, Jespersen 1904; Fujimura and Lovins 1978; see Clements 1990 for discussion):

(9) *Sonority sequencing*

Onset:	less sonorant > more sonorant
Nucleus and Coda:	more sonorant > less sonorant

Given the universality of (8) and (9), we may safely factor the generalizations that these statements embody out of underlying representation, along with the linear order they entail. That is, we know the answer to (10) because we know that the principles of linearization in (8) and (9) hold in English:

(10) *Think of an English word in which*
[æ] is in the nucleus
[p] is in the coda
[m] is in the coda
[l] is in the onset
[k] is in the onset

Given the syllabic affiliation of a set of sounds, their linear order is completely predictable. This paves the way for a theory of representation in which there is no distinctive linear ordering. The importance of this should not be overlooked: a central observation about linguistic structure in syntax, morphology and semantics is that it is hierarchically structured, not linearly ordered; the same claim is fairly obvious for surface representation in phonology. The odd man out is underlying phonological representation, the last stronghold of linearity in linguistic theory. Our proposal is that linearity plays no role here either.

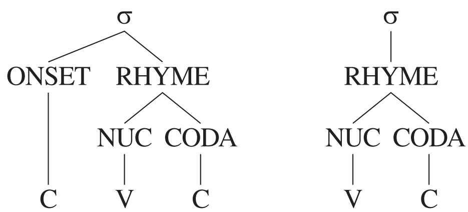
Our idea to make underlying and surface form isotypical is related to the view within Government Phonology that phonological representations are fully interpretable at all levels of representation (Kaye, Lowenstamm and Vergnaud 1990). Indeed, within Goverment Phonology it is also assumed that syllable structure is omnipresent. This model, however, does not address the issue of linearization.

Before we turn to a sketch of the kinds of representations that we use for various types of segments and (complex) syllabic constituents, we must address a potentially damaging empirical problem for our proposal.[5] It is a widely observed fact that a sequence of a closed syllable followed by a syllable that starts with a vowel, is empirically unattested. The traditional view that assumes

5. We are grateful to Marc van Oostendorp for pointing out this problem to us.

that linearly organized strings form the input for syllabification explains this by saying that a string VCV is universally syllabified as V.CV. This is due to constraints that Prince and Smolensky (1993) call Onset and NoCoda. Onset penalizes vowel-initial syllables and NoCoda penalizes consonant-final syllables. As a consequence, any possible contrast between V.CV and VC.V will be neutralized. It would seem, at first sight, that our model predicts that such contrast could exist since nothing seems to prevent us from having a sequence of an (unordered) syllable rhyme (V/C) that is directly followed by a rhyme:

(11)



A further aspect to the same problem is that in languages that allow complex onsets like /pr/, no lexical contrast can occur between V.prV and Vp.rV (and Vpr.V). A further related problem involves what has been called the 'Syllable Contact Law' (Vennemann and Murray 1987). According to this law, in many languages codas are less sonorant than following onsets. How can we express such a generalization if linear order is not part of the lexical representation?

We believe that these effects are best captured by assuming that phonological output is subject to constraints like Onset, NoCoda and the Syllable Contact Law, just like in segment-based phonology. The linearization principles in (8) and (9) are responsible for linearization *within* the syllable and notions like Onset, NoCoda and the Syllable Contact Law are responsible for well-formedness across syllables.

If this seems like a weaker position than having no syllable structure underlyingly, think again. If no language has syllable structure underlyingly, why (we must ask) does *every* language have syllable structure on the surface? The fact that every language makes use of syllables on the surface and the very reasonable assumption that speakers store things that are similar to what they say strongly suggests that syllable-free underlying representations are unlikely. Our position and the standard position differ only in the magnitude of the problems they face. We must account for why no language contrasts CV.CV with CVC.V and we do so with principles that regulate syllable structure. Standard segment-based phonology must account for the same fact and it does so with the same principles *plus* a stipulation that underlying forms are universally unpronouncable and must be syllabified to be spoken. Occam's razor places the burden of proof on theories that use *both* syllabified (SR) and unsyllabified (UR) material.

3. The Representation of Stricture

Stops and fricatives are articulated with a degree of stricture that involves full contact or close approximation leading to turbulence; vowels are made with a degree of stricture that never involves contact or close approximation; and sonorants fall somewhere in between, with no turbulense but some contact or close approximation.

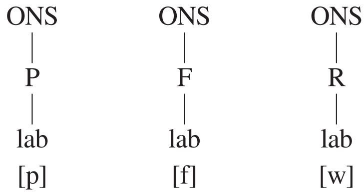
It is a well-founded observation that stricture features do not spread. Alongside processes like nasalization, place-assimilation and voicing assimilation we do not commonly find occlusivization (C → stop / __stop), fricativization (C → cont / __cont) or sonorization (C → son / __son). We understand this as a corollary of our claim that stricture is part of syllable structure: stricture does not spread because prosodic structure does not.

This section will be aimed at refining this claim and at treating the difference between simplex and complex onsets, nuclei and codas. The particulars of our proposals are somewhat tentative and will surely need to be refined further.

3.1 *Simple Margins*

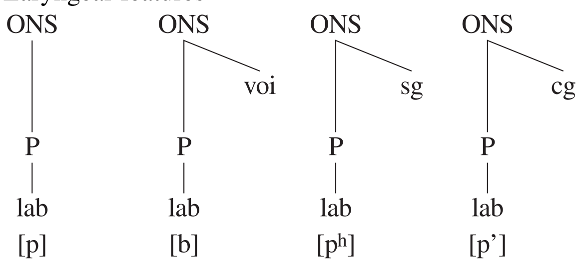
Let us begin with simple onsets. In the unmarked case, an onset has a single uniform place property, a single uniform laryngeal property and a single uniform nasal property. This, we claim, is what makes onsets 'simple,' not some system of segment counting. We begin with place features. Within the onset (and coda) they may be licensed as stops (P), as fricatives (F) or as sonorants (R) as follows:

(12) Place features



Much phonological and phonetic research has shown that laryngeal features like [spread glottis], [constricted glottis] and [voice] are not properly features of segments but of onsets, rhymes, or nuclei (Iverson and Salmons 1995; Golston and Kehrein 1998, 1999). This must be stipulated in a segmental model, which has no other option than to locate laryngeal distinctions on skeletal points, since at UR there is nowhere else to locate features, prosodic structure being absent. The non-segmental nature of laryngeal features is straightforwardly captured in a model such as ours, however, because it recognizes prosodic structure underlyingly. Following Golston and Kehrein (1998, 1999) we assume that laryngeal features are licensed not by stricture nodes but by higher syllabic nodes, as follows:

(13) Laryngeal features

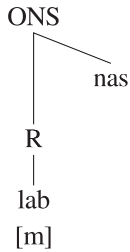


While it is possible to represent complex events like [kt], [ps] and [bd] in our model (cf. 19), it is now not possible to represent *[kd], *[bs] and *[bt] onsets, which are laryngeally disharmonic. The universal absence of such onsets is thus accounted for geometrically and without phonotactic stipulation because all features that associate to the ONS node necessarily have scope over everything that is in it.

We take nasality to be another feature that associates directly with margins and nuclei (Golston 1998). Evidence for this comes from the non-occurrence of contrastive nasal contours within the margin: we know of no language in which

[mb] and [bm] contrast within an onset or coda. This is how we intend to represent simple nasal onsets:

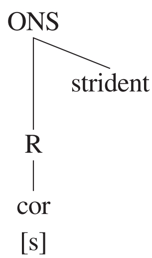
(14) Nasality



Similar representations apply for codas, but we will not go through the details here since the outcome is fairly obvious.

The question arises whether a similar attachment can be defended for features like [lateral], [rhotic] and [strident].

(15) Strident fricative

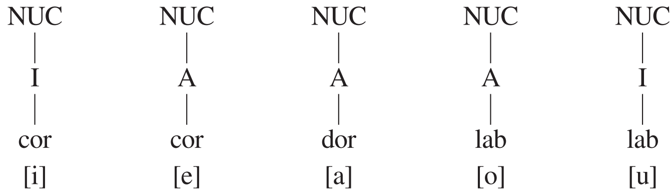


Laterals and rhotics do not often have contrastive places of articulation, but cases are attested. Mid-Waghi contrasts laminal, apical and dorsal laterals (Ladefoged and Maddieson 1996: 190) and Toda contrasts fronted alveolar, alveolar and retroflex trills (ibid: 223). Thus there is no reason not to treat the features [lateral] and [rhotic] analogously with [strident].

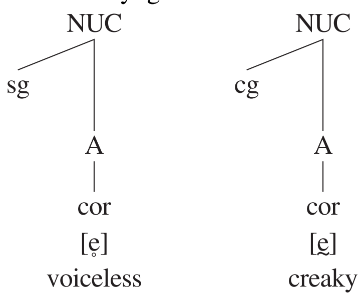
3.2 *Simple Nuclei*

Nuclear representations are similar, with the height encoded in terms of the stricture features A (lo) and I (high). In the unmarked case, a nucleus has a single, uniform place gesture (cf. 16); a single, uniform laryngeal gesture (cf. 17); and a single, uniform nasal gesture (cf. 18). This is what makes nuclei 'simple':

(16) Vocalic place features

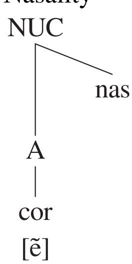


(17) Vocalic laryngeal features



As in syllable margins, nasality is associated directly to nuclei, not to stricture features. Nasalized nuclei are represented thus:

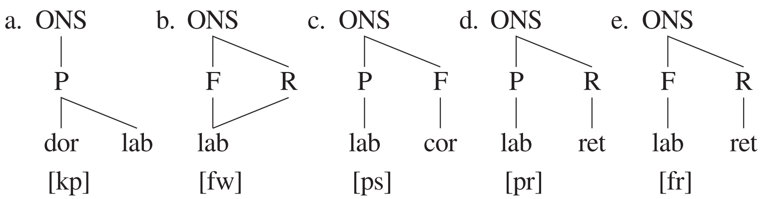
(18) Nasality



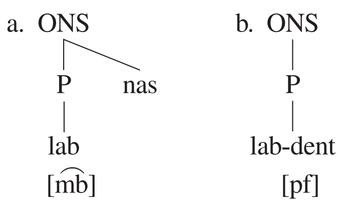
This concludes our discussion of simple onsets and rhymes. We note in passing that each of the representations above would be transcribed in IPA with a single segment (plus diacritics). We turn now to complex syllable structure.

3.3 *Complex Onsets*

We make no principled distinction between traditional branching onsets like [pla], [tra], [kna] and traditional complex segments like [k͡p], [m͡b], [t͡l] since there is little evidence that languages can contrast e.g., [kp] with [k͡p], or [tla] and [t͡la] within an onset or coda (Golston 1998). We represent both types of sound as in (19); the notion ‘complex’ is defined here not in terms of multiple segments but in terms of multiple place features:

(19)


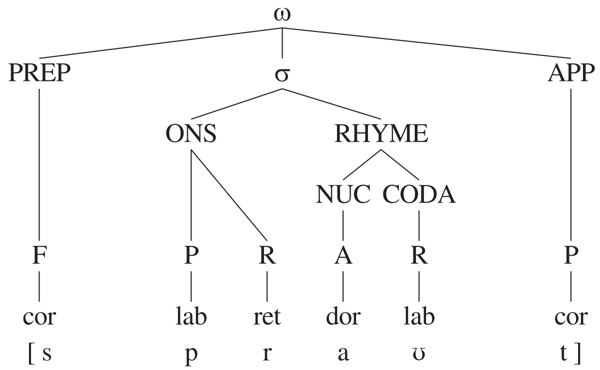
Contour segments invoke an extra cavity feature (prenasalized stops) or a stop at a noisy place of articulation (Kehrein 1998):

(20)


/kn/-type clusters (in Dutch, for example) do not behave like a true onset. Trommelen (1983) showed that they are invariably split up intervocalically. In this model, such clusters cannot be represented as one (complex) onset because the place of the stop and the nasal cannot be disharmonic for place in our model. Hence, the sequence /kn/ must be parsed in terms of two onsets with an intervening empty rhyme, in the style of Government Phonology, or in terms of an extra-syllabic position (prependix) at the word edge.

We have not yet dealt with sC clusters, a topic we feel deserves special consideration. In many languages word-initial obstruent clusters may begin with [s] but with no other consonant: [spa], [sta], [ska] but not *[fpa], *[fta], *[fka] nor *[θpa], *[θta], *[θka] or *[ʃpa], *[ʃta], *[ʃka]. Rather than complicate the prosodic structure of the syllable, we follow a number of researchers in recognizing extra-syllabic metrical positions at word-edge: the word-final appendix and the word-initial prependix (Fudge 1969; Selkirk 1984; Steriade 1982, 1988). We represent words like *sprout* [spraʊt] as follows:

(21) *Prependix, Appendix*



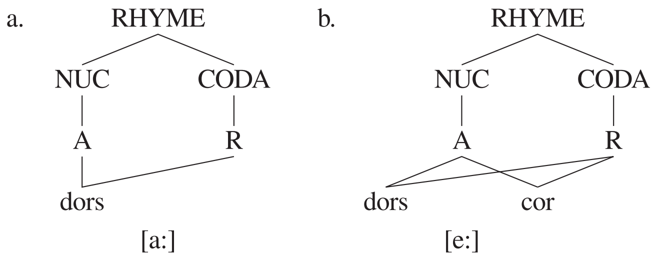
We assume that the appendix can only hold coronals (Fudge 1969; Trommelen 1983; Van der Hulst 1984).[6]

3.4 *Complex Nuclei*

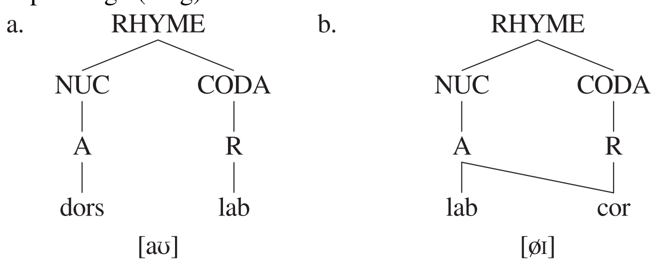
Next, we return to the core of the syllable again. In (22) we represent long vowels that associate their features to both the nuclear and coda position. Associating different features to the nucleus and coda nodes leads to long diphthongs, (23). Short diphthongs involve two place features associated to the nucleus node only, (24).

6. In a word like *ramps*, the [m] might be in the coda, the [p] in a degenerate syllable, and the [s] in the appendix. We will not pursue this here as it leads directly away from the point at hand.

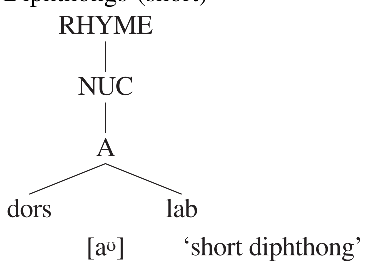
(22) Long vowels



(23) Diphthongs (long)



(24) Diphthongs (short)



This concludes our brief survey of the kinds of stricture that our theory allows. We will now turn to additional evidence supporting our claim that prosodic structure is underlying.

4. Evidence for Underlying Prosody

The phonological representations we have proposed make no use of segmental structure and maximal use of prosodic structure. We take it as given that prosodic structure is well-motivated in surface representations but now need to motivate our claim that it is well-motivated in underyling representations as well.[7] And so we turn now to the issue of underlying prosody.

It is worth pointing out that there is no empirical evidence *against* underlying prosodic stucture. There are a number of quasi-theoretical considerations that have led linguists to eschew underlying syllabification (see Blevins 1995 for recent discussion), but they all address having *both* syllable structure *and* segment strings underlyingly, a possibility obviously fraught with redundancy and one which we will not pursue. We propose replacing segment strings with prosody, not supplementing them with prosody.

So what evidence is there in favor of underlying syllable structure? Here we can distinguish two broad types, linguistic and psycho-linguistic evidence.

4.1 *Linguistic Evidence*

We will adduce no new types of linguistic evidence for underlying prosody in this section. Rather, we will show that our case has already been made for us. The phonological literature of the eighties is replete with arguments that prosody (moras, syllables and feet) is underlying — we merely need to remind the informed reader of these arguments.

4.1.1 *Moras*

A number of phonologists have argued for representing consonant and vowel length in terms of underlying prosody (Hyman 1985; McCarthy and Prince 1986, 1988; Hayes 1989).[8] Long vowels and geminate consonants may be represented as in (25), with a rule that adds a single mora to every vowel, or as in (26) with no such rule:

(25) Minimal moraic specification

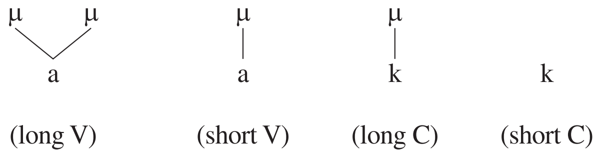
μ		μ	
|		|	
a	a	k	k
(long V)	(short V)	(long C)	(short C)

---

7. See also Dobrin (1993). Inkelas (1994) argues that all non-alternating prosody must be underlying in Optimality Theory, given consistent application of Lexicon Optimization (Prince and Smolensky 1993).

8. See Noske (1992) for critical assessment of the moraic account of compensatory lengthening; we are concerned here with the representation of contrastive length.

(26) Maximal moraic specification



Either way, *some* prosodic information is treated as underlying. Any theory of phonology that embraces such a view of length already makes use of prosody in underlying representation.

More direct evidence for the underlying status of moras comes from minimal requirements on lexical formatives. Extending work on word-minimality (McCarthy and Prince 1986) to the level of roots and affixes, Golston (1990) has argued that English, Latin and Classical Greek require the underlying forms of derivational affixes to be minimally monomoraic. As with contrastive length in moraic theory, this is only statable in a grammar which allows moraic structure in underlying representations.

4.1.2 *Syllables*

There is ample evidence that syllable structure is underlying in many languages. Two general classes of phenomena can be distinguished here: minimality requirements on roots and prosodically defined templatic morphology, both ultimately due to the pioneering work of McCarthy (1979, 1981), Marantz (1982) and McCarthy and Prince (1986, 1990).

Let us take minimality requirements first. Golston (1996b) notes that the well-established shape of Proto-Indo-European roots is exactly the type of evidence one needs for claiming that syllable structure is underlying in a language. PIE had the following types of roots (where R denotes the class of sonorants):

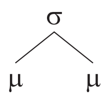
(27) Proto-Indo-European root shapes
(C)VC
(C)VR
(C)VRC

The generalization here is that PIE roots are all single closed syllables: this and nothing else captures the fact that CV, CVV, CVCR and CVCV are impossible PIE roots. A grammar of PIE which does not countenance underlying syllable structure or constraints that regulate it in underlying representation cannot capture this.

Sanskrit provides a similar case. Sanskrit inherited roots from PIE but lost its laryngeal consonants, with compensatory lengthening in coda position. The result, as Steriade (1988) has shown, is that Sanskrit roots are both monosyllabic and bimoraic. Again, this is simply not statable without underlying syllables: if roots are stored simply as segment strings, the fact that the strings all happen to constitute exactly a heavy syllable goes completely unaccounted for.

A similar case can be made for reduplicative morphology. Work by Marantz (1982), McCarthy and Prince (1986), Steriade (1988) and others has shown that reduplication is inherently prosodic. Reduplicative morphemes which consist solely of a syllable template straightforwardly depend on syllable structure being part of lexical representation. McCarthy and Prince's (1986) analysis of Mokilese is typical — the underlying form of the prefix that marks progressive aspect is given as:

(28) Mokilese progressive



The underlying form of this prefix is thus a heavy syllable. Any theory of morphology which countenances prosodic templates such as this supports the claim that prosodic structure can be underlying.[9]

4.1.3 *Feet*

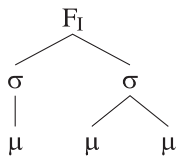
Entirely parallel facts obtain for feet. A number of languages require lexical words to consist of minimally a foot (McCarthy and Prince 1986; Crowhurst 1991). Golston (1990) demonstrates that this can apply to roots as well, suggesting that foot structure must be available underlyingly. Van der Hulst and Klamer (1996) argue for this point using data from Kambera, an Indonesian language, showing that the shapes of roots are often based on the feet used elsewhere in the phonology. Again, such restrictions on underlying forms do not seem to be statable without underlying prosody.

Languages that express morphological category by shape, such as Classical Arabic (McCarthy 1979; McCarthy and Prince 1990), Yawelmani (Archangeli

9. With the advent of OT this might seem to have changed. McCarthy and Prince (1993ab, 1995) propose dealing with the size of reduplicants in terms of constraints: Red = σ, Red = Ft, and so on. But the difference is merely cosmetic: either way, the sole surface regularity of a reduplicative morpheme is its size and this will have to be indicated in any type of grammar.

1991) and Choctaw (Ulrich 1986, 1994; Lombardi and McCarthy 1991; Hammond 1993) also show the necessity of underlying prosody. The point has been made repeatedly in the literature that morphemes in such languages may consist solely of a certain prosodic shape: this *is* underlying foot structure, another clear case of underlying prosody. Consider McCarthy and Prince's analysis of the Arabic Broken Plural as an iambic foot (F_I).

(29) Arabic Broken Plural



Underlying prosody constitutes the underlying form of this morpheme; again, a clearer case of underlying prosody cannot be imagined.

We see then that underlying prosody has become a central part of generative phonology in the past decade. This is true both of pure prosody (reduplicative and templatic morphology) and of prosody attached to melody (the moraic analysis of geminates); and it is true of all well-established levels of prosody — mora, syllable and foot. In this context our proposal that prosodic structure is underlying is utterly banal.

4.2 *Pyscholinguistic Evidence for Underlying Prosody*

Psycholinguistic evidence offers an important check on linguistic theory by casting light on the psychological reality of different grammatical models. The best models of grammar are compatible both with linguistic and with psycholinguistic data. As we will see here, two types of evidence bear on the phonological representations speakers actually store and it is significant that they agree with one another point for point (Cutler 1986: 173; Levelt 1989: 355).

4.2.1 *Tip of the Tongue (TOT) States*

Brown and McNeill (1966) showed that speakers who cannot think of a word tend to know three things about it: the initial segment or onset, the number of syllables and the stress pattern. When a speaker tries to access the phonological form for *sextant*, for instance, words like *secant* and *sextet* come to mind. The results have been confirmed by much subsequent research (Gardiner, Craik and Bleasdale 1973; Yarmey 1973; Koriat and Lieblich 1974, 1977; Rubin 1975; Browman 1978; Reason and Lucas 1984; Kohn, Wingfield, Menn, Goodglass, Berko-Gleason and Hyde 1987; Priller and Mittenecker 1988). This strongly suggests that speakers store words as syllables and do not store them merely as strings of segments.

Consider the alternative. If a speaker stores a word as a string of segments and cannot access (part of) that string, she should not be able to compute the number of syllables or location of stress. For to do so would require access to the string of segments she (ex hypothesi) has no access to. If, on the other hand, prosodic structure is stored, we expect it to be accessible even if the full form of the word is not.

4.2.2 *Malapropisms*

Classifications of speech errors include a category of sound-related substitutions (Fromkin 1973) or malapropisms (Fay and Cutler 1977), involving mis-selection of a word that is phonologically but not semantically similar to the intended word. Typical cases include ('F' from Fromkin; 'FC' from Fay and Cutler):

(30) Sound-related substitutions

Intended		Spoken	
white Anglo-Saxon Protestant	>	white Anglo-Saxon prostitute	F
a routine proposal	>	a routine promotion	F
the conquest of Peru	>	the conquest of Purdue	F
prohibition against incest	>	prohibition against insects	F
week	>	work	FC
open	>	over	FC
constructed	>	corrected	FC

As these cases illustrate, the overall prosody of the target is matched by the overall prosody of the error:

(31)

Intended	Spoken	stress pattern	syllable count
Protestant	prostitute	(x..)	3
proposal	promotion	(.x.)	3
Peru	Purdue	(.x)	2
incest	insects	(xx)	2
week	work	(x)	1
open	over	(x.)	2
constructed	corrected	(.x.)	3

What we do not generally find in sound-related substitutions are cases like *protest* for *Protestant*, *proposition* for *proposal*, *Peruvian* for *Peru*; or *insecticide* for *insects* — forms we have every reason to expect if words are stored merely as segment strings.

The criteria for phonological similarity here are identical to those found in TOT states: same onset, same stress pattern, same number of syllables. Data like this has led researchers like Crompton (1982), Fromkin (1985) and Butterworth (1989) to posit a phonological sub-lexicon within the mental lexicon. *White Anglo-Saxon prostitute* is produced when *prostitute* is mis-selected because of its proximity to *Protestant* in the phonological sub-lexicon. None of this makes any sense if syllable count and stress pattern are not somehow stored.

Again, it is most significant that two quite different sources of evidence converge on the same criteria: word-onsets, stress pattern and number of syllables are used in accessing the phonological forms of words. Any psychologically plausible model of grammar must come to terms with this and admitting prosody into underlying representation seems like the necessary first step.

5. Conclusion

We have tried to provide a psychologically plausible and linguistically reasonable theory of phonological representation that makes maximal use of prosodic structure and minimal use of extrinsic ordering. Our model makes no use of segments, root nodes, C-, V- or X-tiers and the like and minimizes the differences between underlying and surface forms. More specifically we propose that stricture is not encoded featurally in phonological representation but structurally in the form of syllabic sub-constituency.

Acknowledgments

Part of the research carried out here was conducted while the first author was part of the research grant "Theorie des Lexikons" (SFB 282) from the German Research Foundation.

Phonological Restructuring in Yidiɲ and its Theoretical Consequences

Bruce P. Hayes

Detailed study of data in Dixon's (1977) grammar indicates that previous analyses of Yidiɲ have erred in supposing that the synchronic pattern of the language continues the historical pattern, whereby various nominal stems have their underlying final vowels deleted when no suffix follows. Instead, it appears that the system has undergone a radical reanalysis, whereby the suffixed forms are now projectable by general principles from the isolation forms. More precisely, a pattern of *multiple predictability* has developed: the form of suffixed allomorphs is largely predictable from the isolation allomorphs, but the older pattern, whereby isolation allomorphs can be predicted from the suffixed allomorphs, also persists.

From this descriptive result, three principal theoretical consequences are developed: (a) Yidiɲ possesses a fully-productive pattern of alternation that is not driven by markedness-faithfulness interactions; (b) the phonological constraints that are active in Yidiɲ likely include some that are quite unlikely to be members of a universal inventory; (c) there are more relations of predictability among surface forms in Yidiɲ than can be treated by the normal method, namely that of deriving all the surface allomorphs from a single underlying representation.

A tentative suggestion is made for how Optimality Theory might be extended to treat cases of this sort, by means of a class of "Anticorrespondence" constraints.

1. Introduction: Where Does Phonology Come From?

Optimality Theory (Prince and Smolensky 1993) has yielded a compelling picture of phonology, one in which quite intricate data patterns emerge from the relative ranking of markedness and faithfulness constraints. The markedness constraints are characteristically highly general, supported by cross-language typology, and in many instances explicable in terms of functional goals such as those involving articulation or perception. Faithfulness constraints are characteristically atomic, limited to bans on simple insertions, deletions, and feature changes. The theory is receiving increasing support from the study of acquisition (e.g. Gnanadesikan 1995; Pater 1996), from which it appears that many of the markedness constraints are spontaneously manifested in the speech of small children; thus the Stampean idea that phonology is to some degree internally-generated (Stampe 1973) has come again to the fore.

All of this looks like progress, especially when compared with the degree of ad hocness and language-specificity that prevailed in earlier work. The caution introduced here is this: does *all* phonology result from markedness-faithfulness interactions?

To begin, we will consider how exceptional phonology, not treatable under standard Optimality-theoretic assumptions, would be most likely to arise in the course of historical change.

It is a commonplace that phonologies are not transmitted directly from generation to generation; rather, it is language data that are transmitted, and the phonology must be constructed by each new generation on the basis of the input data, along with whatever is provided a priori by the human language faculty.

The process is not necessarily reliable, as language change attests. Consider in particular a case where the older generation comes to apply with ever-greater frequency a process of phonological deletion, plausibly the most dramatic of phonological changes. As the frequency of deletion comes to approach 100%, new language learners will ultimately be faced with a situation in which the recovery of underlying forms is not at all easy: the lost segments of non-alternating forms will not be recoverable at all, and even in alternating forms the character of lost segments might be difficult to obtain. Such situations are ripe for *restructuring*, the creation of a novel phonological system on the basis of the data pattern left behind by an older system.

Thus, there are plausibly two types of phonological change. One is gradual, system-internal change, in which markedness and faithfulness constraints are reranked; this is expected to yield coherent patterns amenable to treatment by standard OT approaches. The other is restructuring, which arises out of the attempt of a new generation to make sense out of a data pattern presented to it by an older generation. The data pattern may be a quite difficult one, due to the accidents of history, and it is an empirical question whether the language learners who carry out restructuring will necessarily limit themselves to systems expressible with only these formal resources.

What all this implies is that if we are to explore the full richness of human phonologies, clearly we should devote some attention to systems that are known to be restructured. It seems quite possible that such systems will put the mechanisms of the theory to the severest test, and suggest directions for revision.

This is what I have tried to do here. I will argue that the phonology of Yidiɲ, an Australian aboriginal language of North Queensland, has undergone a substantial restructuring, one which puts it beyond the reach of the central mechanisms of markedness and faithfulness in OT, and further seems mostly likely to be incompatible with the view of a universal constraint inventory. I will further speculate as to what modifications of the theory might be capable of producing workable OT analyses of Yidiɲ.

The research I report here extends an earlier tradition, whose important works include Vennemann (1972), Schuh (1972), Hale (1973), and Kenstowicz and Kisseberth (1977). Work in this earlier line likewise suggested restructuring as an important origin of phonological processes.

2. Yidiɲ

Yidiɲ phonology was worked out with considerable insight by R. M. W. Dixon (1977, hereafter D) in a intensively detailed description based on work with several of the last native speakers of the language. Dixon's data and generalizations have proven irresistible to phonological theorists, who have principally aimed to elucidate the interesting and unusual metrical structure of the language with particular versions of metrical stress theory.[1]

The focus here lies not in the metrical structures of Yidiɲ per se, but rather in the complex patterns (partly metrically governed) that are found throughout the Yidiɲ nominal and verbal paradigms. Almost every Yidiɲ stem shows alternations of vowel length, and many stems show vowel-zero alternations as well.

---

1. See Nash (1979), Hayes (1980, 1982, 1995a), Halle and Vergnaud (1987), Kirchner (1993), Crowhurst and Hewitt (1995), and further references cited by Crowhurst and Hewitt.

I will first review the Yidiɲ data from a historical viewpoint. This will have two advantages: it avoids an a priori expository prejudice for any particular analytical line, and it helps set the stage for a restructuring account, by showing what kind of data pattern must have faced the innovating generation of Yidiɲ speakers.

2.1 *Historical Yidiɲ*

The dramatic phonological alternations of Yidiɲ are largely the product of two historical changes. One of them is characterized by Dixon (D 43) as follows:

(1) PENULTIMATE LENGTHENING
In every word with an odd number of syllables, the penultimate vowel is lengthened.

Penultimate Lengthening resulted in a huge number of alternations. For example (D 43), the underlying stem for 'mother', /mud^{j}am/, appears as [mud^{j}am] in the absolutive case, where no suffix is added; this is unaltered from its historical form. With the Purposive ending [-gu] added, the resulting form is trisyllabic (odd-syllabled), and thus was eligible for Penultimate Lengthening, which yielded [mud^j**aː**mgu]. The trisyllabic stem for 'dog', historically *[gudaga], appears in modern Yidiɲ as [gud**aː**ga] in the absolutive case, since it is trisyllabic. But when the purposive suffix [-gu] is added to it, the form becomes even-syllabled, and therefore Penultimate Lengthening did not affect it: [gud**a**ga-gu]. A quadrisyllabic stem, [ŋunaŋgara] 'whale' (D 84) was unaltered in the absolutive, but received penultimate length in (for example) the dative, where a suffix renders the form pentasyllabic: [ŋunaŋgar**aː**-nda].

Naturally, one senses that there should be some connection between the interesting environment "penult of an odd-syllabled word" and the alternating stress pattern of Yidiɲ, laid out elsewhere in Dixon's grammar (D 40–41). To establish and formalize this connection is one goal of the many metrical accounts of Yidiɲ phonology.

The pattern expressed by Penultimate Lengthening continues to be highly productive in synchronic Yidiɲ, and is virtually exceptionless on the surface.

The other major phonological change that created modern Yidiɲ was as follows:

(2) FINAL SYLLABLE DELETION
In a word ending in C_1 (C_2) V, delete C_2 and V, if:
a. The form that results ends in a legal word-final consonant (/l,r,ɽ,y,m,n,ɲ,ŋ/)
b. The form that results possesses an even number of syllables.

As one might expect, Final Syllable Deletion led to numerous phonological alternations, many of which persist into the synchronic state which Dixon describes. For example (D 45), if one takes the bare stem [buɲa] 'woman' and adds the basic postvocalic form of the ergative suffix [-ŋgu], one obtains [buɲaːŋ], which may be presumed to have been historically *[buɲaːŋgu] (and somewhat earlier, *[buɲaŋgu]). Similarly, the quadrisyllabic form [ŋunaŋgara] 'whale-absolutive' shows up in the ergative as [ŋunaŋgaraː-ŋ] (D 84), historically *[ŋunaŋgara-ŋgu]. The ergative ending can be seen in its unaltered historical form after a trisyllabic stem, e.g. in [mulari-ŋgu] 'initiated man-erg.' (D 57).

It can be seen that Final Syllable Deletion, taken as a sound change, must have occurred after Penultimate Lengthening, since in [buɲaːŋ] and countless similar words, what was historically the penultimate vowel shows up as long. Indeed, Dixon's synchronic analysis recapitulates the historical ordering, placing Penultimate Lengthening first.

As a result of Final Syllable Deletion, a number of Yidiɲ suffixes have two dramatically different allomorphs, as was just shown for ergative [-ŋgu]/[-ːŋ]; for further cases see (3) below. Moreover, Final Syllable Deletion also yields alternations in stems. For example, the stem meaning 'moon' shows up as trisyllabic when suffixed: [gindanu-ŋgu] 'moon-erg.' (D 57), since quadrisyllables never underwent Final Syllable Deletion. But alone, 'moon' was trisyllabic, and thus underwent Penultimate Lengthening and Final Syllable Deletion to yield modern Yidiɲ [gindaːn].

Unlike Penultimate Lengthening, Final Syllable Deletion has not left a clean, across-the-board data pattern in contemporary Yidiɲ. Rather, there are many exceptions. These arose perhaps as analogical restorations, or perhaps the original process was lexically irregular in the first place.

Among the suffixes, we find that the majority of forms that could in principle alternate actually do (3a). Exceptions, however, are non-negligible (3b).

(3) a. *Suffixes which alternate by Final Syllable Deletion*

[-ŋgu] ~ [-ːŋ]	ergative
[-ɲa] ~ [-ːɲ]	accusative
[-yi] ~ [-ːy]	nominal comitative
[-ɲu] ~ [-ːɲ]	past ([-n-] conjugation)
[-l-ɲu] ~ [-ːl]	past ([-l-] conjugation)
[-ɽ-ɲu] ~ [-ːɽ]	past ([-ɽ-] conjugation)
[-ɲu-nda] ~ [-ɲuː-n]	dative subordinate (verbal)
[-ŋa] ~ [-ːŋ]	verbal comitative
[-ŋa] ~ [-ːŋ]	verbal causative

b. *Suffixes which do not alternate by Final Syllable Deletion*

[-nda]	dative (nominal)
[-na]	purposive
[-n-dʲi], [-l-dʲi], [-ɽ-dʲi]	'lest' inflection (in [-n-], [-l-], and [-ɽ-] conjugations)

Note that this list is incomplete, as there are three other suffixes that alternate by syllable count, but in ways that cannot be dealt with by Final Syllable Deletion alone: locative [-la] ~ [-ː], ablative [-mu] ~ [-m], and genitive [-ni] ~ [-ːn] ~ [-nu] ~ [-nə].

Looking at the inventory of roots, we find a similar bifurcation. Dixon (D 58) counts 80 roots that alternate by Final Syllable Deletion (e.g. [gindanu-ŋgu] ~ [gindaːn]), and 52 roots that do not undergo the process, even though they are phonologically eligible. An example of the latter is the stem for 'initiated man', which undergoes only Penultimate Lengthening where applicable: [mulaːri] (not *[mulaːr]) ~ [mulari-ŋgu].

Due to all of this lexical idiosyncrasy, the synchronic version of Final Syllable Deletion as stated by Dixon ends up incorporating various mechanisms to make reference to individual lexical items possible:

(4) FINAL SYLLABLE DELETION (from D 48)

$XV_1C_1(C_2)V_2\# \rightarrow XV_1C_1$

if (a) $XV_1C_1(C_2)V_2\#$ is an odd-syllabled word;

and (b) C_1 is one of the set of allowable word-final consonants;

and (c) either (i) there is a morpheme boundary between V_1 and C_1

or (ii) V_2 is a "morphophoneme": A, I or U.

In this formulation, a "morphophoneme" is understood to be the final vowel of one of the stems like /gindanu/ that is (more or less idiosyncratically) eligible for deletion; Dixon writes /gindanU/. Moreover, it is understood that case (4c.i), the suffix truncation case, permits of lexical exceptions, namely the suffixes listed in (3b).

Plainly, if Final Syllable Deletion was once a fully-productive, across-the-board process of Yidiɲ, its current status is rather attenuated, with considerable exceptionality and lexical idiosyncrasy.

2.2 *A Traditional Analysis*

Dixon provides a cogent traditional phonological analysis of the Yidiɲ alternations, using a fairly standard post-*SPE* approach. This account has served as the basis of almost all subsequent treatments of Yidiɲ. As with many phonological analyses, the synchronic description recapitulates history. Dixon's basic assumptions are as follows.

– Underlying representations of stems include all the vowels that a stem displays anywhere in its paradigm. Thus the UR of surface [gindaːn] 'moon' is /gindanu/, since the /u/ shows up in suffixed forms like [gindanu-ŋgu].
– Vowels are assumed to be short underlyingly, except in the rare cases where they show up long across the board. Thus the /a/ of [gindaːn] is underlying short /a/, whereas the nonalternating /uː/ of [durguː] 'mopoke owl' (D 84) is underlyingly long.
– The historical sound changes of Penultimate Lengthening and Final Syllable Deletion are continued in synchronic Yidiɲ as phonological rules, applied in the order just given, thus: /gindanu/ → gindaːnu → [gindaːn]. Clearly, rule ordering is required to make this solution work, since the odd-syllabled structural description of Penultimate Lengthening is met only at the deep level of representation, before Final Syllable Deletion has rendered the form even-syllabled.
– Since Final Syllable Deletion is quite irregular in its application, both stems and suffixes are lexically marked for whether they may undergo it.[2]

To my knowledge, all of the many subsequent analyses of Yidiɲ, my own included, have followed Dixon on these basic points. However, the traditional approach has two crucial defects which, to my knowledge, have not been previously noticed. Both indicate that it is not sufficient as a true characterization of the Yidiɲ speaker's knowledge of her language.

2.2.1 *Defects of the Traditional Analysis I: Distribution of Invariant /V:/*
The first problem arises from the distribution of vowel length in Yidiɲ, which is worth examining in detail.

Aside from Penultimate Lengthening, already discussed, there are three sources of long vowels:

1. Certain suffixes idiosyncratically cause the vowel of the preceding syllable to surface as long. For example, when the antipassive suffix [-ːdʲi], a prelengthener, is attached to the stem meaning 'see, look' (and the past tense ending [-ɲu] is

2. Dixon uses "morphophonemes" for stems and exception features for suffixes, but the basic distinction is the same.

added to the result), we get [waw**aː**-dʲi-ɲu], with length on the second syllable (D 218). Dixon plausibly traces the prelengthening property of these suffixes to historical sources in which the lengthening was compensatory, arising to fill the length slot of the deleting segments. However, as he shows, the synchronic situation appears fairly clearly to involve an arbitrary, morphologically-triggered lengthening.

2. Dixon suggests (D 77–83) that some instances of [iː] result from monophthongization of underlying /iy/.[3]

3. Finally, in just a few stems, vowel length is an invariant property of the stem; i.e. it is phonemic under any analysis. Dixon does not speculate on where these long vowels come from historically. They are unlikely to be ancient, as they do not correspond with the length reconstructed from the evidence of neighboring languages (D 70); and indeed this reconstructed length has been lost in Yidiɲ.

Dixon collected sixteen morphemes with underlying length. Thirteen are disyllables, with the long vowel in final position, e.g. [durguː] 'mopoke owl', [giŋaː] 'vine species'; and the remaining three are quadrisyllables, with length in second or fourth position: [galambaɽaː] 'march fly', [waɲinbaraː] 'what's the matter?' and [waɽaːbuga] 'white apple tree'. It can be determined that vowel length in these forms has nothing to do with Penultimate Lengthening, because it is invariant throughout the paradigm: thus [durguː] 'mopoke owl-absolutive', [durguː-n] 'id.-genitive', [durguː-nda] 'id.-dative'; [durguː-nu-la] 'id.-locative of genitive'; etc. (D 84, 137).

By inspecting the forms in Dixon's list of invariant long vowels (D 85–86), one can easily determine the following:

- There are no invariant long vowels found in *odd syllables*. This gap is explainable on metrical grounds,[4] and it is not crucial to present concerns.
- Invariant long vowels may never occur in *closed syllables*. This gap will be quite crucial to the discussion, and should be borne in mind.

---

3. It is not a foregone conclusion that this suggestion should be accepted; in particular, it leads to a puzzling asymmetry in the phonemic long vowel inventory, whereby only two of the three Yidiɲ vowels (/uː/, /aː/) are permitted to occur in underlying forms. If we take [iː] at face value this asymmetry disappears. The cost is a rather peculiar arrangement in suffix allomorphy whereby some stems ending in /i/ (arguably the lowest-sonority vowel) must be lexically marked to take the suffix allomorphs that are otherwise used for consonant-final stems.

4. See the downloadable Appendix to this article, Section 11.3.3; cited in fn. 16.

– Finally, invariant long vowels do not occur in *trisyllabic stems*.[5] A hypothetical case, to show what such a long vowel would look like if it existed, is /nulaːri/, which would show up as [nulaːri] in the unsuffixed absolutive case, and (for example) as [nulaːri-ŋgu] in the ergative.

Note that there are no *phonological* reasons why *[nulaːri-ŋgu] could not exist: its length pattern is completely legal. It arises, for instance, when a disyllabic stem with a final long vowel is followed by two suffixes, as in [durguː-nu-la] 'mopoke owl-genitive-locative', given above. The same length pattern can also be created when one takes a disyllabic stem and adds two suffixes of which the first is a pre-lengthener: a form given earlier, [wawaː-dʲi-ɲu] 'see-antipassive-past', is an example. Finally, the same length pattern also appears in the monomorphemic form [waɽaːbuga] 'white apple tree'.

Could these gaps in the Yidiɲ stem inventory be accidental? I have calculated the expected number of aberrant stems on the following basis. A rough check of Dixon's Yidiɲ glossary[6] yielded 437 disyllabic stems, of which 13 have a long vowel. There are 55 quadrisyllabic stems, of which 3 have a long vowel. Now, there are 206 trisyllabic stems. Following the percentages observed for the other lengths, we should find somewhere between 6 and 11 trisyllabic stems with a long vowel, but in fact there are none (a fact confirmed by Dixon, D 86). Checking with a chi square test, we find that there is only a 1.2% chance that this situation could arise by accident.

The absence of long vowels from closed syllables likewise appears not to be accidental. Consider that of 398 total stems in the glossary that end in an open syllable, 15 have a final long vowel. There are 301 stems that end in a consonant. Of these, we would expect about 11 to have a long vowel, but none do. A chi square test indicates that the probability that this could arise by chance is about 0.07%.

One can actually defend the view that Yidiɲ speakers were *tacitly aware* of these data patterns (though not necessarily using the scheme that the traditional analysis posits). As will be seen below, there is good evidence that Yidiɲ speakers often had to concoct inflected forms of stems that they were only familiar with in the bare absolutive form. When they did this, they generally

---

5. Or more generally, in odd-syllabled stems. There are so few pentasyllabic stems that the absence of invariant long vowels in them could easily be an accident. No stem in Yidiɲ is longer than pentasyllabic.

6. All stems counted except verbs and adverbs, which never have invariant length.

treated surface [V:] as (in Dixon's terms) a derived [V:], not an underlying one; thus [CVC**V:**C] normally appears suffixed as [CVC**V**CV$_i$-CV], where V_i is some vowel, and [CVC**V:**CV] as [CVC**V**CV-CV].

The asymmetry in the distribution of long vowels is an unexplained oddity for the traditional approach. In derivational theories, such gaps must be stated as constraints on underlying representations, as follows:

(5) *Yidiɲ Deep Phonotactics*
In underlying representations:
a. There are no trisyllabic stems with a long vowel.
b. There are no long vowels in closed syllables.

These underlying phonotactic constraints will guarantee that there are no alternations of the type *[CVCV:C] ~ [CVCV:C-CV] or *[CVCV:CV] ~ [CVCV:CV-CV]. I take them to be descriptively accurate, but explanatorily deficient: why should there be a ban on long vowels in specifically trisyllabic stems? Further, why should such a ban occur in (of all languages), Yidiɲ, where because of Penultimate Lengthening, on the surface there is a long vowel in *every* trisyllabic stem? Why, in addition, should long vowels be banned underlyingly in closed syllables, where on the surface it is in fact quite normal for a long vowel to occur in a closed syllable?

Intuitively, there is an answer to these questions. Yidiɲ, in the traditional account, is a language with many derived long vowels, but few underlying ones. In contexts where derived long vowels abound, apparently Yidiɲ speakers have made the assumption that every long vowel is a derived one. This seems plausible, but it has no straightforward translation in the traditional framework.

For accounts using Optimality Theory, the gaps in the invariant long vowel inventory are more embarrassing still, as under the doctrine of the Richness of the Base (Prince and Smolensky 1993; Smolensky 1996), OT has aspired (rightly, in my opinion) to avoid constraints on underlying forms entirely.

2.2.2 *Defects of the Traditional Analysis II: Patterns of Vowel Restoration*

Putting this problem to the side for the moment, let us consider another difficulty for the traditional analysis. One claim that the traditional approach makes quite explicitly is this: for stems that alternate by Final Syllable Deletion, the allomorph of a stem that packs the greatest amount of phonological information is the suffixed allomorph. This is the allomorph that preserves the crucial final vowel (any of the three Yidiɲ vowels /i/, /a/, or /u/) that cannot be determined by inspecting the truncated isolation stem. The basic pattern of predictability is claimed to be "from the context stem, you can predict the isolation stem", but not vice versa.

This is the claim, but the actual facts of Yidiɲ are different. In the great bulk of the cases, the form of the full stem is predictable (up to free variation) from the form of the isolation stem. This view is adumbrated in a fascinating section (D 59–65) of Dixon's grammar.

During the Yidiɲ elicitations, once Dixon had figured out the basic pattern of Final Vowel Deletion, he took care to obtain a suffixed form for every stem he had previously encountered only in the plain absolutive form. In a number of cases, the very same isolation form yielded more than one suffixed form, either from different speakers or from the same speaker on different occasions. The pattern of free variation is complex; and for the moment I will discuss only the statistically predominant case, which I will call the "standard" outcomes.

There are about 79 stems to consider; namely, those which are truncated in their isolation form but show up with an extra vowel in their suffixed form. As Dixon (D 60) notes, there is a strong regularity governing what vowel must be added to the absolutive base to obtain the inflected stem: in 57 out of 79 cases, it is a *copy of the second stem vowel*; thus $[\mathrm{CVC}\mathbf{V}_i\mathrm{ː C}] \sim [\mathrm{CVC}\mathbf{V}_i\mathrm{C}\mathbf{V}_i\text{-CV}]$.[7] Here are six examples (D 65–68) culled from the 57 total cases; [-CV] represents any suffix. The vowel copy relation is emphasized here with boldfacing.

(6)	[bab**aː**l]	~	[bab**a**l**a**-CV]	'bone'
	[bandʲ**aː**ɽ]	~	[bandʲ**a**ɽ**a**-CV]	'madness'
	[dʲig**uː**r]	~	[dʲig**uru**-CV]	'thunderstorm'
	[gab**uː**l]	~	[gab**ulu**-CV]	'stick for carrying fish'
	[gaw**iː**ɽ]	~	[gaw**i**ɽ**i**-CV]	'crescent shaped'
	[mind**iː**r]	~	[mind**iri**-CV]	'salt-water centipede'

This is an interesting observation, and becomes more so if we add a slight emendation to it: when the consonant that appears at the end of the isolation allomorph is a *nasal*, the vowel that is added is virtually always /u/. This provision brings an additional 14 cases into the realm of predictability. Here are five representative examples, taken from the 14:

7. My counts differ slightly from Dixon's due to a different procedure: I omit cases with free variation and cases where the vowel is unknown, but include forms from Dyalŋuy ("mother-in-law language").

(7)	[baɽi:**ɲ**]	~	[baɽi**ɲu**-CV]	'tree used for handles'
	[dʲala:**m**]	~	[dʲala**mu**-CV]	'fresh, young'
	[dʲuri:**n**]	~	[dʲuri**nu**-CV]	'leech'
	[ginda:**n**]	~	[ginda**nu**-CV]	'moon'

As these cases show, the requirement for /u/ after nasals appears to outrank the requirement for the last two stem vowels to be the same; there are no cases at all like "[wug**a:m**] ~ [wug**ama**-CV]" where the vowel identity requirement overrides the post-nasal /u/ requirement. Naturally, there are a number of examples in which both vowel identity and postnasal position lead to the same result.

Putting Dixon's vowel-copy pattern together with the post-nasal principle, we get a striking outcome: the inflected stem is to a fair degree predictable from the truncated absolutive stem. I state this principle in preliminary form below:

(8) STEM VOWEL RESTORATION

a. If the truncated absolutive stem is [CVC_0V:**N**], where N is a nasal consonant, then the full inflected stem must be augmented by /-u/.

b. Otherwise, the full inflected stem of [$CVC_0$$\mathbf{V_i}$:C] must be augmented by /-$\mathbf{V_i}$/.

Of the 79 relevant stems, 71 have their missing vowels filled in correctly by Stem Vowel Restoration (57 by vowel copying, and 14 by post-nasal /u/). Seven additional cases show free variation, all between the vowel predicted by Stem Vowel Restoration and some other vowel.

(9)	[gadʲ**u:**l]	~	[gadʲ**u**l**a**-CV], [gadʲ**u**l**u**-CV]	'dirty (e.g. water)'
	[gaw**u:**l]	~	[gaw**u**l**a**-CV], [gaw**u**l**u**-CV]	'blue gum tree'
	[gunb**u:**l]	~	[gunb**u**l**a**-CV], [gunb**u**l**u**-CV]	'Dyalŋuy ?: billy-can'
	[mag**u:**l]	~	[mag**u**l**a**-CV], [mag**u**l**u**-CV]	'a root food'
	[gangu:**n**]	~	[gangu**na**-CV], [gangu**nu**-CV]	'bushes arranged as trap'
	[ɲag**i:**l]	~	[ɲag**i**l**a**-CV], [ɲag**i**l**i**-CV]	'warm'
	[wag**a:**ɽ]	~	[wag**a**ɽ**i**-CV], [wag**a**ɽ**a**-CV]	'wide'

In three cases, Dixon could only obtain forms in which the consultants followed the alternate vowel restoration strategy of suffix vowel copying (discussed below), so that the inherent stem vowel could not be determined. These cases are listed and discussed in (16) below. Finally, there were only eight cases that are outright exceptions; that is, show an invariant vowel contrary to that predicted by (8). These exceptions are the following:

(10)	[d^{j}amb**u:**l]	~	[d^{j}amb**ula**-CV]	'two'
	[gang**u:**l]	~	[gang**ula**-CV]	'grey wallaby'
	[gamb**i:**ɽ]	~	[gamb**i**ɽ**a**-CV]	'tablelands'
	[gaɲd^j**i:**l]	~	[gaɲd^j**ila**-CV]	'crab'
	[gubu:**m**]	~	[gubu**ma**-CV]	'black pine'
	[gul**a:**ɽ]	~	[gul**a**ɽ**i**-CV]	'big-leaved fig tree'
	[guŋg**a:**ɽ]	~	[guŋg**a**ɽ**i**-CV]	'north'
	[waŋ**a:**ɽ]	~	[waŋ**a**ɽ**i**-CV]	'pre-pubescent boy'

Despite these exceptions, it can be argued that Stem Vowel Restoration is productive in Yidiɲ; that is, "psychologically real". In particular, for some forms Dixon located evidence (D 65–68) from cognate stems of neighboring languages which indicate what the originally truncated final vowel must have been. In most of these, it appears that Yidiɲ has altered the original historical pattern of alternation so as to conform to the Stem Vowel Restoration principle.

(11)	Form	Gloss	Cognate
	[gaw**u:**l] ~ [gaw**u**l(**a**/**u**)-CV]	'blue gum tree'	[gawul**a**] (Dya:bugay)
	[mag**u:**l] ~ [mag**u**l(**a**/**u**)-CV]	'a root food'	[magul**a**] (Dya:bugay)
	[ɲag**i:**l] ~ [ɲag**i**l(**a**/**i**)-CV]	'warm'	[ɲagil**a**] (Gunggay), [ɲigal**a**] (Mamu Dyirbal)
	[yag**u:ɲ**] ~ [yag**uɲu**-CV]	'echidna'	[yuguɲ**a**ŋ] (Gunggay)
	[muɽi:**n**] ~ [muɽ**inu**-CV]	'ashes'	[murin**i**] (Dya:bugay)
	[band^j**a:**ɽ] ~ [band^j**a**ɽ**a**-CV]	'madness in head'	[band^j**a**ɽ] (Dyirbal)
	[d^{j}uŋg**u:m**] ~ [d^{j}uŋg**umu**-CV]	'worm'	[d^{j}uŋgum] (Dya:bugay)
	[gaba:**n**] ~ [gaba**nu**-CV]	'rain'	[gaba:n] (Dya:bugay)
	[ginda:**n**] ~ [ginda**nu**-CV]	'moon'	[ginda:n] (Dya:bugay)
	[gugi:**ɲ**] ~ [gugi**ɲu**-CV]	'flying fox'	[gugi] (Mamu Dyirbal), [gugi:n] (Dya:bugay)
	[mala:**n**] ~ [mala**nu**-CV]	'right hand'	[mala:n] (Dya:bugay)

There are, however, three troublesome forms that work the wrong way:

(12)	Form	Gloss	Cognate
	[gub**u:**l] ~ [gubul(**a**/**u**)-CV]	'billy-can'	[gunbul**u**] (ŋgadjan Dyirbal)
	[wag**a:**ɽ] ~ [wagaɽ(**a**/**i**)-CV]	'wide'	[wagar**a**] (Dya:bugay)
	[gamb**i:**ɽ] ~ [gambiɽ**a-CV**]	'tablelands'	[gamb**i**l] (Dyirbal, Mamu Dyirbal)

Assuming that some explanation is possible for these cases,[8] the main force of the comparative data is that Stem Vowel Restoration has been active in reshaping the inventory of alternating stems in Yidiɲ. Note that several forms in (11) affirm the greater strength of the post-nasal subcase of Stem Vowel Restoration over the vowel copying subcase.

Consider further the cases of free variation seen in (9). Since these forms always involve a vowel of the quality predicted by Stem Vowel Restoration, it seems very likely that these cases illustrate change in progress, with new vowels derived by Stem Vowel Restoration displacing the older, etymologically correct vowels.

This, then, is a novel phenomenon to be accounted for. Previously, we have imagined that the truncated allomorph is largely predictable from the full allomorph; this is just what one would expect from a historical process that deleted all three stem-final vowels. But now we can see that *the full allomorph is also largely predictable from the truncated allomorph*. It will be recalled from above that 52 of the 132 eligible stems simply don't apocopate; an example is [mulaːri] ~ [mulari-ŋgu] 'initiated man.' Because of these cases, it is not particularly more effective to try to predict the absolutive forms from the inflected forms than vice versa.

2.2.3 *An Attempt to Save the Traditional Analysis: Deep Constraints*

The traditional analysis, as it has always been stated, has nothing to say about the data patterns involved in Stem Vowel Restoration. Could it be beefed up to include these crucial and productive generalizations? Earlier, we rather awkwardly covered a similar problem (the distribution of non-alternating long vowels) by adding constraints on underlying forms. Let us consider the same strategy here:

(13) *Deep Constraints as an Attempted Substitute For Stem Vowel Restoration*

a. In an underlying stem of shape /CVC_0VNV_i/, where N is a nasal, V_i must be /u/.

b. Otherwise, in an underlying stem of shape /$CVC_0V_jCV_i$/, V_i and V_j must be identical.

Interestingly, these proposed constraints *fail empirically*. To see this, recall that about 30% of the eligible trisyllabic stems idiosyncratically fail to alternate by

8. For discussion of possible accounts, see the downloadable Appendix to this article, cited in fn. 16.

Final Vowel Deletion, instead having trisyllabic allomorphs across the board. In the traditional approach, these stems have just the same kinds of underlying representation as the alternating stems, except that they possess an exception feature that blocks Final Vowel Deletion.[9] Thus, any constraint on underlying forms that holds true of the alternating stems should hold true just as well of the non-alternating ones.

In fact, this is not so. Among non-alternating stems listed in Dixon's glossary, only 8 out of 21, or 38%, of the stems obey the constraints proposed in (13); this may be compared with the 83–91% obedience rate (depending on how the free variants are counted) in the alternating stems.

(14)

Obey: 8		Disobey: 13	
[baŋga**mu**]	'English potato'	[bibi**ya**]	'coconut tree'
[bud^{j}**ala**]	'fine, finely ground'	[dam**ari**]	'silly'
[dal**iyi**]	'hunger'	[d^{j}iy**uya**]	'catbird'
[dig**ara**]	'coast'	[galg**ali**]	'curlew'
[d^{j}ud**ulu**]	'brown pigeon'	[gaɽa**na**]	'black cockatoo'
[gud^{j}**ara**]	'broom-like implement'	[gum**aɽi**]	'red'
[gug**ulu**]	'recitative style'	[gurg**iya**]	'khaki bream'
[ŋaw**uyu**]	'salt-water turtle'	[mug**aɽu**]	'fish net'
		[mul**ari**]	'initiated man'
		[mulɲ**ari**]	'blanket'
		[ɲiri**ɲi**]	'long peppery fruit'
		[waɲ**ira**]	'what kind of'
		[yuɽ**iya**]	'saltwater snake species'

There is another comparison possible which makes the same point. In Dixon's glossary, there are 33 stems that in principle would be eligible for Final Vowel Deletion, but cannot undergo it because the final vowel is preceded by an obstruent or by a consonant cluster (both of these conditions reliably block the process). Now, of these, only 14, or 42%, obey the deep constraints (for example, [gud**a**g**a**] 'dog-absolutive'), whereas 19 are like [bind**u**b**a**] 'crayfish-absolutive' and disobey them.

What seems to be the correct generalization is this: the final vowel of a

9. Alternatively, as in Dixon's account, the undergoers bear a special diacritic that triggers the rule.

trisyllabic stem is strongly constrained to obey the generalizations of Stem Vowel Restoration only when this vowel alternates with zero. It would seem that Stem Vowel Restoration is a principle governing *the relationships between forms*, not *the shape of underlying forms*. Thus, even with the additional device of constraints on underlying forms, the traditional analysis has no grip on the crucial data. The facts suggest rather forcefully, I think, that this analysis should be reconsidered. Before doing so, we must consider further data that also point in this direction.

2.3 *More on Stem Vowel Restoration: The Alternative Strategies*

The data pattern of Stem Vowel Restoration is actually more intricate than the previous section suggests. Dixon's careful description (D 59–65) actually records three *additional* patterns, which apparently occur largely in free variation with the statistically dominant pattern just noted.

2.3.1 *The Suffix Vowel Copy Pattern*

On some occasions, Dixon's consultants volunteered forms in which the vowel added to the isolation stem was a copy of the vowel of the following suffix. Thus, one consultant gave as the inflected forms of absolutive [gambiːn] 'top-knot pigeon' the following (D 61):

(15)	[gambin**u**-ŋg**u**]	ergative
	[gambin**a**-l**a**]	locative
	[gambin**i**-y**i**]	comitative

In a later elicitation, this consultant provided only forms of the type [gambin**u**-CV], which follow the primary pattern, specifying /u/ post-nasally.[10]

There were three stems Dixon collected which always alternated by this pattern:

(16)	[mugiːɽ]	~	[mugiɽ**V**-C**V**]	'small mussels'
	[wubuːl]	~	[wubul**V**-C**V**]	'lucky (at hunting, etc.)'
	[wurguːl]	~	[wurgul**V**-C**V**]	'pelican'

There is no evidence in Dixon's grammar to support the view that the suffix vowel copy pattern is *obligatory* for these stems, and indeed for the case of [mugiɽ**V**] Dixon explicitly states that other outcomes also were found (D 62).

10. It is unknown whether this vowel is the etymologically correct one.

2.3.2 *The Schwa Pattern*
There also occurred cases (D 62) in which the "restored" vowel was a schwa. Thus a consultant offered as the genitive of [gubuːm] 'black pine' the form [gubum**ə**ni], instead of the primary (and probably, etymologically correct) form [gubum**a**ni]. The appearance of schwa is extraordinary, since schwa does not otherwise occur as a vowel of Yidiɲ.

2.3.3 *Nonalternation*
The last of the three supplementary patterns is the rarest of all: it is simply the attachment of the suffix to the truncated isolation stem itself, as in [gindaːn] ~ [gindaːn-dʲi] 'moon-absolutive, comitative' (D 64). If the suffix involved has a special allomorph for consonant-final stems, such as comitative [-dʲi], that is the allomorph which is used (the postvocalic allomorph of the comitative can be seen in the normal variant [gindanu-yi]).

The nonalternation pattern is only doubtfully well-formed: when Dixon played tapes to his speakers of themselves saying such forms, they sometimes felt that these were errors. Thus whatever analysis is adopted for them should not treat them as fully integrated into the language.

Now, consider again the traditional analysis of the vowel ~ zero alternations. The outlook for this analysis clearly becomes worse when these cases of free variation are considered. Systematic free variation is typically felt to diagnose optional or competing rule systems, but the traditional account has no alternative but to place the variation in the underlying forms themselves, with massive loss of generalization.

2.4 *The Genitive Ending*

The Yidiɲ genitive marker [-ːn] ~ [-ni] ~ [-nu] ~ [-nə] provides one further argument against the traditional analysis. Genitive in Yidiɲ is treated by Dixon as a derivational category. The genitive stem created by affixation of [-ːn] ~ [-ni] ~ [-nu] ~ [-nə] is inflected with the normal nominal cases (absolutive, ergative, dative, etc.). Let us consider the distribution of the allomorphs of the genitive.

(a) [-ːn] shows up in absolutives (i.e., "bare genitives") when two conditions are met: that its use should result in a word that is even-syllabled, in accord with the general Yidiɲ preference, and that the stem be vowel final, so that the Yidiɲ requirements for segment sequencing may be respected.

(17)	[buɲa]	'woman-absolutive'	[buɲa-**ːn**]	'woman-genitive'	D 53
	[dʲadʲa]	'child-absolutive'	[dʲadʲa-**ːn**]	'child-genitive'	D 136
	[ŋunaŋgara]	'whale-absolutive'	[ŋunaŋgara-**ːn**]	'whale-genitive'	D 83

(b) [-ni] appears in absolutives wherever [-ːn] would be disallowed; thus when the base has an odd number of syllables (18a), or when the base ends in a consonant (18b):

(18)	a.	[gudaːga]	'dog-absolutive'	
		[gudaga-**ni**]	'dog-genitive'	D 53
		[guŋgaːdʲi]	'tribal name-abs.'	
		[guŋgadʲi-**ni**]	'tribal name-genitive'	D 134
	b.	[dumbul]	'blue-tongued lizard-absolutive'	
		[dumbuːl-**ni**]	'b.t. lizard-genitive'	D 134
		[gunduy]	'brown snake-abs.'	
		[gunduuy-**ni**]	'brown snake-genitive'	D 135

Moreover, any forms created by further suffixation to such bases likewise take [-ni], suitably altered by Penultimate Lengthening (1) where appropriate: [gudaga-**niː**-nda] 'man-genitive-ergative' (D 53), [gunduy-**ni**-la] 'brown snake-genitive-locative' (D 135).

(c) The third allomorph, [-nu] shows up when further suffixes are added to plain genitives that end in [-ːn]. Thus paradigms like those of (17), which show [-ːn] in the plain genitive, may be amplified as follows:

(19)	[buɲa]	'woman-abs.'	
	[buɲa-**ːn**]	'genitive'	
	[buɲa-**nu**-nda]	'gen.+dative'	D 53
	[dʲadʲa]	'child-abs.'	
	[dʲadʲa-**ːn**]	'genitive'	
	[dʲadʲa-**nu**-ŋgu]	'gen.+ergative'	D 136

(d) The fourth and final allomorph of the genitive is [-nə]. It occurs in free variation with /-nu/.

What are we to make of this pattern? First, the distribution of [-ni] vs. [-ːn] plainly follows the normal pattern in Yidiɲ: allomorphy is guided by an even-syllable target. The remaining, more puzzling, cases of (c) and (d) above follow under the assumption that the [u] and [ə] of [-nu] and [-nə] are *restored vowels*, just like the restored vowels of monomorphemic stems. Thus, [buɲaː-n] restores

/u/ postnasally in [buɲa-n**u**-ŋgu], but may also restore schwa in [buɲa-n**ə**-ŋgu] (form implied in D 54).[11]

The genitive data are of interest because they illustrate the relative strength of two generalizations that compete. Had Yidiɲ speakers opted for an *invariant underlying form* for the genitive, then the abundant instances of [-ni] (that is, with any odd-syllabled or consonant-final stem) would imply that this allomorph should also appear when Final Syllable Deletion is blocked by the presence of an additional suffix. In actual fact, the inserted [u] and [ə] vowels show that the extra vowel is being projected on the basis of the phonological shape of the *bare genitive stem form*. Thus, Stem Vowel Restoration takes place even when there is substantial evidence available for what the underlying stem vowel is supposed to be.

I take this to be one further reason to hold the traditional account in grave doubt: it must treat the genitive suffix with allomorphy, replicating the pattern that occurs independently with simple noun stems.

3. Theoretical Consequences

All the above discussion has as its general conclusion that Yidiɲ is not fully or properly analyzed under the traditional account. The traditional account, because it merely recapitulates history, does not capture the new patterns that arose as further generations of Yidiɲ speakers reanalyzed the system.

Is this conclusion of importance just to Yidiɲists, or does it have more general theoretical consequences? I would suggest the latter, on several grounds.

3.1 *The "Inside-Out" Preference in Phonology*

In Hayes (1995b) I suggested, following much earlier work, that phonological systems tend to organize themselves in ways that permit derived forms (such as the suffixed case forms of Yidiɲ) to be predicted from the base forms (usually, as in Yidiɲ, isolation forms). In that article, I presented a couple of examples suggesting that languages often rearrange their phonologies so that this will be true. Further support for this position is presented by Kenstowicz (1996).

11. Dixon does not report cases in which the genitive suffix vowel is restored by Suffix Vowel Copy (section 2.3.1), or is not restored (non-alternation; section 2.3.3). I conjecture that this is an accidental data gap: suffix-copy forms, and especially non-vowel-restored forms, are rare, and it could easily have happened that within the relatively rare morphological category of suffixed genitives, Dixon did not find any cases.

Yidiɲ provides the most intricate case of this sort I have yet seen. It would appear that a serious reorganization of the phonology has taken place, permitting the suffixed allomorphs of stems to be predicted from their isolation allomorphs. In other words, the "direction of predictability" in the system has to a fair degree reversed. In the older system, the isolation form of a stem like [gindaːn] could be predicted from its suffixed form [gindanu-CV], by means of Final Syllable Deletion. But following restructuring, it became largely possible to predict the suffixed form on the basis of the isolation form, using general and productive principles. Thus Yidiɲ phonology went from having an "outside-in" character to the (I claim, less marked) "inside-out" type.

It is easy to imagine when this change must have happened: "outside-in" phonology would have been quite stable while Final Syllable Deletion remained as an *optional* pattern of alternation. At one point there would have occurred archaizing free variants like [gindaːnu], alongside innovating [gindaːn]. Archaic [gindaːnu] would have provided ample evidence to learners for the underlying form of the stem. It is when Final Syllable Deletion became entirely obligatory that the restructuring process would likely have begun.

3.2 *Rule Inversion and the Basis of Phonological Constraints*

It has been suggested, e.g. by Tesar and Smolensky (1996), that phonology is learned simply by using the input data to rank a set of constraints given *a priori*; that is, as part of Universal Grammar. Under this approach, one would expect all phonologies to be highly principled in character, given that all of their ingredients are universal and only their ranking is language-specific.

The facts considered here do not necessarily refute this view. But they do encourage one to think of the acquisition problem in broader terms. Consider the pattern of nominal inflection in Yidiɲ, at the stage in the language's history just after Final Syllable Deletion became obligatory. I will use the term "Pre-Yidiɲ" to refer to this crucial stage. As Dixon suggests (D 59, 64–65), the Pre-Yidiɲ data pattern must have been quite hard to learn, because the unsuffixed absolutive stems occur in text more often than all other inflected forms put together. The restructuring that took place, converting Pre-Yidiɲ to Modern Yidiɲ, presumably was a direct response to this difficulty.

Let us hypothesize that a major goal of the language-acquiring child is to produce *correct and accurate reproductions* of the inflected forms spoken around her. Given that phonology is often cleanly patterned and based on markedness constraints of wide applicability, the Tesar/Smolensky strategy of ranking only principled constraints will often suffice, or come close to it, in attaining accurate reproduction. But as we have seen, sometimes historical change deals the child a difficult hand. My contention is that in such cases the child does what is necessary, which includes the creation of relatively ad hoc, language-specific constraints. Yidiɲ seems to be an example of this, especially in its requirement that [u] be the restored vowel after nasals.

The specific scheme I suggest is as follows: in cases of great difficulty, such as was found in Pre-Yidiɲ, the language learner will seize upon generalizations that are statistically useful, albeit imperfect, and make use of them to improve her ability to guess unknown inflected forms. There are several cases where this appears to have happened here.

Consider first the observation made earlier that (under the traditional analysis) underlying long vowels are excluded from trisyllabic stems. Here, the crucial statistical pattern was that the great majority of long vowels in Yidiɲ nouns were due historically to Penultimate Lengthening. Thus, if a language learner heard [CVC**V**ːCV] in an isolation stem, it was a truly excellent bet that the second stem vowel would appear as short under suffixation: [CVC**V**CV-CV]. I believe that this is the origin of (what has been called) the ban on underlying long vowels in trisyllables. Whatever invariant long vowels may have existed in trisyllabic stems in Pre-Yidiɲ (these would have had paradigms like [CVC**V**ːCV] ~ [CVC**V**ːCV-CV]) were drowned in the statistically predominant pattern of alternating long vowels.[12] The need for language learners to be able to project the vowel lengths of suffixed forms from isolation forms thus rendered it impossible to sustain a vowel length contrast in trisyllabic stems.

The second case to consider is the principles of Stem Vowel Restoration, covered in sections 2.2.2. and 2.3. above. I believe it plausible that these principles also originated as exaggerations of statistical patterns already present in the Yidiɲ lexicon. In attempting to provide the missing vowel, Yidiɲ learners relied on a slight pre-existing tendency toward vowel harmony in the final syllables of stems, as well as a slight pre-existing preference for /u/ after nasals (see (14) and immediately following discussion). These slight tendencies, which may well have been true *by sheerest accident* in Pre-Yidiɲ, were exaggerated and employed as the best available means of predicting the quality of the inserted vowel in the reanalyzed modern language.

At the time Dixon collected his data, the principles of Stem Vowel Restoration had not yet stabilized; they still competed with rival strategies for stem

12. There is every reason to believe that such long vowels once existed, since the historical processes that create long vowels (loss of C from VCV, compensatory lengthening) are not ordinarily sensitive to syllable count.

vowel restoration. These competing strategies, laid out in Section 2.3., were presumably created out of whole cloth by Yidiɲ learners, and indeed they seem to have a more principled, markedness-driven character. The tendency towards vocalic harmony in epenthetic vowels is widespread in languages, and vocalic harmony also exists as a typical constraint in child phonology. Schwa is likewise a typical quality for epenthetic vowels.[13]

Consider finally the vowel length alternation in earlier stems of the type [CVCVːC] ~ [CVCVCV-CV]. Here the projection of suffixed form from isolation form worked with a combination of the previous principles. The exceptionless pattern of inserting a vowel after /…CVːC/ arose from the fact that all, or virtually all isolation stems ending in /…CVːC/ were derived historically by Penultimate Lengthening and Final Syllable Deletion. Any historical stems that had once alternated as [CVCVːC] ~ [CVCVːC-CV] would have been drowned in the vastly larger statistical pattern, and would have undergone readjustment to [CVCVːC] ~ [CVCVC**V**-CV], using the same principles for choice of epenthetic vowel just discussed.

The upshot of all this is, perhaps, an extension of one's conceptions of "where phonology comes from". Rather than constituting just a language-particular prioritization of general, a priori principles, *some* phonology seems to represent the relatively ad hoc response of learners to conundrums presented to them by historical change. Faced with such a conundrum, learners are capable of fabricating entirely new phonology, which has no direct connection with a language's earlier, more "motivated" form. This conclusion follows earlier research on "rule inversion", cited above in the introduction.

It is true that *most* phonology does seem to have the principled character that arises from well-motivated constraints. The reason for this statistical predominance is that the rather more ad hoc principles of the type discussed here arise only in the context of restructuring, where peculiar historical evolution forces the language learner to develop odd constraints to deal with the resulting data conundrums.

---

13. Before concluding that these strategies necessarily constitute "UG in action," however, one would want to rule out an alternative suggested by Dixon (D 62): that the restored vowels are meant to be inconspicuous in their context, thus helping to avoid embarrassment at being unable to remember the "correct" vowel. Schwa is a good candidate here, since it is roughly equidistant perceptually from the three full-vowel possibilities that could otherwise occur. Copying of a neighbor vowel is also a good choice: given the pervasive existence of vowel-to-vowel coarticulation (Öhman 1966), any vowel in the crucial position will be shifted somewhat in the direction of its neighbors. Thus a vowel copying a neighbor is a good bet for resembling the unknown "correct" vowel.

Moreover, we can expect that even the constraints that arise from restructuring will not be utterly unprincipled; they too must fall at least loosely within the set of constraints possible within phonological theory in general.

3.3 *Alternation Not Driven by Markedness-Faithfulness Interactions*

In the introduction to this article I mentioned that it is a virtue of Optimality Theory that it ties the pattern of alternation in a language to basic principles of markedness. At the same time, it does seem that there are cases of productive alternation that cannot be reduced to the simple interaction of markedness and faithfulness. Yidiɲ provides a fairly clear example.

Consider the vowel ~ zero alternations we have been examining. They may usefully be compared with another type of alternation, which is extremely common: the appearance of derived vowel length whenever a [-CV] suffix is attached to a [CVCVC] stem (cf. (1), Penultimate Lengthening). An example given earlier is [mud^{j}am] 'mother-absolutive' ~ [mud^{j}a:m-gu] 'mother-purposive'. Curiously, in this pattern of alternation, Yidiɲ *creates* a word shape ([CVCV:CCV]) which it elsewhere tries to avoid. Specifically, where [CVCV:CCV] would arise simply from the concatenation of [CVCV:C] and [-CV], Yidiɲ productively inserts vowels to avoid this outcome, in the cases of Stem Vowel Restoration, discussed at length above.

In brief, we have two patterns of alternation:

(20) a. [CVC**V:**C] ~ [CVC**V**C**V**-CV]
(Stem Vowel Restoration cases)
b. [CVC**V**C] ~ [CVC**V:**C-CV]
(ordinary suffixation to /CVCVC/)

The pattern in (20a) cannot be driven by orthodox markedness-faithfulness interactions, as the following reasoning shows. Assume for purposes of argument that there exist highly-ranked markedness constraints that force [CVCV:C] + [CV] in (20a) to surface as [CVCVCV-CV]. The penalty in faithfulness for doing this is (as stated in the system of McCarthy and Prince 1995) a violation of IDENT([+LONG]), due to the loss of vowel length, and a violation of DEP(V), due to the insertion of a vowel. Yet, in (20b), where our starting point is [CVCVC] plus [-CV], the very same markedness constraints are satisfied by the candidate *[CVCVCV-CV], which is not the winner; and the penalty in faithfulness is *less*, namely just a violation of DEP(V). Thus it is not at all clear how the real winner [CVCV:C-CV] could ever be made to defeat its ill-formed rival, at least without serious amplifications of the theory.

The reader might object at this point that the [CVC**V**C**V**-CV] output in forms of the type (20a) is not the result of GEN altering the underlying representation, but simply of the underlying representation surfacing unaltered. But the whole point of the above discussion was to show that this point of view is wrong: the traditional analysis fails to capture the facts in precisely this area. It is more plausible to suppose that alternations like [CVCVːC] ~ [CVCVC**V**-C**V**] really are due to processes of epenthesis and vowel shortening, somehow expressed appropriately within Optimality Theory. It is these processes that cannot be treated as straightforward markedness-faithfulness interaction.

The Yidiɲ case differs from other examples of this type in a crucial way. In many languages, there is problematic phonology that involves lexical exceptions or special vocabulary strata. It is easy to imagine a theory that sets up a treatment for exceptions (e.g. Pater in preparation) or strata (Itô and Mester 1995), so that such alternations do indeed follow from markedness-faithfulness interaction. The Yidiɲ case is different because the problem is entirely general (involving essentially all forms of the relevant shape), and not limited to any kind of stratum or exception class.

Thus, classical OT needs amplification; a means of driving alternations that cannot be the result of markedness-faithfulness interaction. This is not to deny that this interaction is the central and normal cause of alternation in phonology: cases like Yidiɲ are the exception, with an origin in historical restructuring.

3.4 *Underlying Representations*

Consider the task of taking the "Wug" test in Yidiɲ. In this test, named after the seminal work of Berko (1958), an experimental subject is given one inflected form of a novel stem, and is asked to say what the other inflected forms would be; thus, in the classic case for English, one is asked "What is the plural of [wʌg]?" and replies "[wʌgz]."

For a hypothetical Yidiɲ form like "[bagaːn]", we already know that Yidiɲ speakers would pass the Wug test: as shown above, in the past various forms along the lines of [bagan**u**-nda] have already been invented, "correctly", by Yidiɲ speakers. The speakers succeeded in passing the Wug test by inventing the principles of Stem Vowel Restoration, which redefined what counts as the correct answer.

What, then, of the *other* direction for Wug testing, namely suffixed form to plain stem? The sort of question asked here is: "If there were a particular thing that in the dative was called a [baganu-nda], what might its absolutive form be?"

Here, alas, we lack data on Yidiɲ speakers' Wug-testing performance, but it seems fairly likely (based on my own informal experience in Wug-testing speakers of other languages) that they would be able to name the two most likely candidates, namely [baga:n] and [baga:nu]. Let us suppose tentatively that this is the case, and see what would follow.

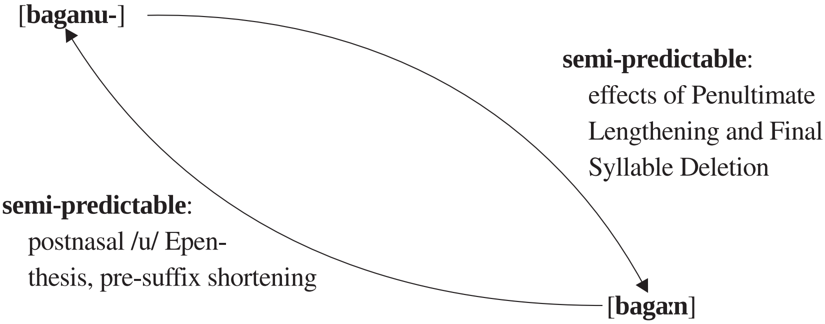
Here is the crucial observation. With roughly 90% certainty, it is predictable in Yidiɲ that the dative form of a hypothetical absolutive [baga:n] will be [baganu-nda], at least as a free variant. This is because the principles of Stem Vowel Restoration work correctly about 90% of the time. But simultaneously, by Dixon's figures (D 58) it is about 70% predictable that the absolutive form of a hypothetical dative [baganu-nda] will be [baga:n] (and not [baga:nu]).

With this in mind, we can consider the question: How do speakers pass the Wug test? Under traditional accounts of phonology (both Optimality-theoretic and earlier generative ones), the crucial mediating element is the underlying representation. The taker of the Wug test uses her phonological and morphological knowledge in the "backwards" direction to deduce what could be the underlying form or forms of what she is hearing, then applies the morphology and phonology in the forward direction to figure out the predicted surface form (or forms) that answer the original question.

In the Wug test for Yidiɲ considered here, the two reasonable choices for an underlying representation of a surface absolutive [baga:n] are /baga:n/, if the alternation is due to epenthesis and a vowel shortening process, or /baganu/, if the alternation is actually due to the lengthening and Final Syllable Deletion posited in the traditional account. But neither choice does real justice to the pattern of predictability. The underlying representation /baga:n/ claims that the isolation stem allomorph [baga:n] is arbitrary from the viewpoint of the inflected form [baganunda] (it is not; it is predictable on about a 70% basis); while /baganu/ claims that the inflected form [baganunda] is arbitrary from the viewpoint of the plain stem [baga:n] (it is not; it is predictable on about a 90% basis).

This situation, in principle, has major theoretical consequences. As textbooks like Kenstowicz and Kisseberth (1979) ably lay out, the central role of phonological underlying representations is to serve as a repository of unpredictable information concerning the pronunciation of a given morpheme. The patterns of predictability among the members of the paradigm of that morpheme are supposed to follow from the phonological rules (or other principles, such as GEN and constraints in OT), which act to derive all surface allomorphs of a stem from the unique underlying form. What is special about Yidiɲ is that there is *more predictability* present in the system than can be accounted for under the usual assumption of derivation from a single underlying form. The following diagram illustrates this point.

(21)



This situation is not unique to Yidiɲ. In Hayes (1995b), I laid out the evidence for a comparable situation in English. In English, it is largely predictable that forms ending in […ənt] will alternate with […ɛnt] upon suffixation, as in numerous pairs like ˈ*elem*[ə]*nt* ~ *ele*ˈ*m*[ɛ]*ntal*, ˈ*accid*[ə]*nt* ~ *acci*ˈ*d*[ɛ]*ntal*, etc. (Here is the Wug test: if X = [ˈpɛɹədənt], what is [Xəl]? Answer: [ˌpɛɹəˈdɛntəl]). But it is *also* predictable that stems which end in […ɛnt] when affixed will show […ənt] in their isolation forms (Wug test: if X+əl = [ˌpɛɹəˈdɛntəl], what is X? Answer: [ˈpɛɹədənt]). Derivation from a single underlying form does not suffice to account for both relations.[14]

---

14. Some further details, to nail down the argument:

a. Suppose we use a rule (or equivalent) of Vowel Reduction to guarantee the […ɛnt] → […ənt] mapping. Then, to get the […ənt] → […ˈɛnt] mapping, we need to restrict underlying forms so that final /nt/ is always preceded by /ɛ/. But this can't be right, because English has many words like *st*[ʌ]*nt,* [æ]*nt*, *st*[ɪ]*nt*, ˈ*galli*ˌ*v*[æ]*nt*, etc. It is only where the principles of English stress happen to make the final syllable stressless (i.e., in polysyllabic nouns) that the underlying vowel must be limited to /ɛ/. So the underlying structure condition would have to duplicate the stress rules in its structural description, a highly undesirable move.

b. Alternatively, we could suppose that the underlying form of the relevant words ends in /…ənt/, and posit a /ə/ → ɛ / ___ nt] rule, the "Full Vowel Restoration" of Hayes (1995b). But here again, we must keep vowels other than schwa from occurring in this underlying position, else they would show up on the surface when assigned stress. This forces the need for an underlying constraint, just as before, only this time requiring final /nt/ to be preceded by /ə/. Obviously, this constraint, too, is falsified by *stunt*, *stint*, etc., unless it duplicates the stress rules.

c. A final option is to make Vowel Reduction cyclic, forcing […ənt] on the first cycle, with Full Vowel Restoration applying on the second cycle. This fails, because Vowel Reduction cannot in general be cyclic (compare ˈ*at*[ə]*m* ~ *a*ˈ*t*[ɑ]*mic*, ˈ*syst*[ə]*m* ~ *sys*ˈ*t*[ɛ]*mic*.

The conclusion is that, just as in Yidiɲ, the crucial generalization relates surface forms to each other, not surface forms to underlying forms.

3.5 *Summary*

Optimality Theory, at least as it is usually practiced today, is in one sense a conventional theory of derivational phonology, in which surface forms are related to one another by deriving them from a common underlying form. As such, it suffers from the problems that beset all such theories. These are: that no means is provided to account for the characteristic "inside-out" bias of phonology, and that no means is provided to account for hypercharacterized phonological systems like Yidiɲ in which there is more interparadigm predictability than can be handled by deriving all allomorphs of a morpheme from a single underlying form.

I have also suggested two lessons from Yidiɲ that pertain specifically to OT. First, it is unrealistic to expect all constraints to be utterly principled manifestations of UG: some constraints result from the efforts of language learners to pass the Wug test (that is, to project novel members of paradigms) in cases where historical change has served them up a difficult data pattern. It seems that in such cases learners will take advantage of regularities present in the data, even relatively arbitrary ones. Second, it is unrealistic to view all alternation in phonology as driven by the relative ranking of faithfulness and markedness constraints (though obviously, much alternation is). Here again, the need for language learners to deal with the Wug test for difficult data patterns dealt them by the accidents of history leads the learners to create fully-productive and widespread patterns of alternation that go beyond what can be treated with the resources of contemporary OT.

4. Where to Go From Here?

A clear picture of how phonological theory should respond to cases like Yidiɲ is not yet in view. But a few speculative remarks might be helpful.

First, in line with the title of this book, I would assert that while the Yidiɲ data involve a kind of "residue" (set of unaccountable phenomena) for standard Optimality Theory, it is hardly a *derivational* residue. Nothing in traditional derivational phonology seems to promise any better account of the data here, and indeed there are phenomena in Yidiɲ for which non-derivational OT seems very well suited: Yidiɲ phonology is quite "conspiratorial," in the sense of Kisseberth (1970), and conspiracies are just what OT was designed to handle. Some notable conspiracies in Yidiɲ are the "look-ahead" property of Final Syllable Deletion (2), which only applies when a legitimate final consonant will remain after truncation, and the pervasive phonotactic target of even-syllabledness.

Second, as a way of shedding light on the problems raised here, it might be worthwhile briefly to situate Optimality Theory in its historical context. The generative phonology of the 1960's and early 1970's was very much focused on "deep" phenomena: patterns of alternation that (it was claimed) could be treated only with quite complex and intricate mappings from underlying to surface representations. This led to a strong emphasis on language-particular derivations as the central element of the theory; and this in turn led to two long-term crises that were ultimately resolved only with the introduction of OT. These crises were (i) the conspiracy problem and (ii) an inability to relate the content of language-particular analyses to general principles of markedness.

The approach that OT used to solve these two problems was to turn the research strategy on its head: instead of positing rules to account for alternation, with the hope of somehow finessing the problems of conspiracies and markedness, classical OT let the crucial principles that govern conspiracies and markedness serve as the heart of the theory, and hoped instead to finesse the problem of complex patterns of alternation.

Specifically, the device that OT provides to permit alternation is the ranking of particular faithfulness constraints sufficiently low in the grammar. Faithfulness constraints are characteristically atomistic, banning simple feature switches, insertions, and deletions. Given that an earlier research program felt it necessary to devote massively complex resources to the description of alternation, we should not be surprised if the much more impoverished capacities of contemporary OT are not up to the task of handling the more complex alternation patterns.[15]

Is this cause for despair? My own judgment is that it is certainly not. Arguably, we are very much better off under OT than we were before. We now have a theory that promises an adequate account of the very general and pervasive phenomena of conspiracies and markedness; earlier, we had a theory that had very little to say about these matters. The payoff for returning to earlier approaches would be quite small, and would primarily concern only the residual class of phonological phenomena that arise from restructuring.

---

15. A large further class of cases that should be considered in this context are those previously treated with opaque rule ordering. These are now the subject of a growing literature, which attempts to reconcile opacity with OT.

4.1 *Anticorrespondence*

As for what is needed in OT to treat facts like those given in this article, I tentatively suggest that we should address the question of patterns of alternation rather more directly than they are addressed with simple faithfulness constraints. I conjecture that what is needed is constraints of "Anticorrespondence," which would *actively require* morphemes to alternate in particular ways. An Anticorrespondence constraint would say: "if morpheme μ appears with shape X in a particular context C, it must appear with shape X′ in a distinct context C′". One might call the allomorph of context C the *base allomorph* and the allomorph of context C′ the *projected allomorph*.

Here is a brief Yidiɲ example. As noted above, all isolation [CVCVːC] stems in Yidiɲ are amplified by a vowel when a suffix is added to them, as in [bab**aː**l] ~ [bab**ala**-CV] 'bone' from (6). As was observed in Section 3.3., it does not appear to be possible to treat this as a simple assertion of markedness over faithfulness, since parallel cases where the isolation form is [CVCVC] do not undergo a similar shift. An Anticorrespondence constraint to treat this is stated in (22), where "X" designates some non-null string:

(22) VOWEL RESTORATION

$$\emptyset \ / \ \mathrm{V{:}C} \ \_\_\_ \ ] \ // \ \_\_\_ \ ]_{\mathrm{Word}} \rightarrow \mathrm{V} \ // \ \_\_\_ \ \mathrm{X}$$

The notation here involves both single and double environment slashes.

– The single slash refers to the applicable context *within the base allomorph*. Thus, in (22) the single slash indicates that the constraint is focusing on the right edge (expressed as the null string) of a base allomorph that ends in /…VːC/.
– The first double slash refers to the *context in which the base form appears*, which in this case is isolation (i.e., no suffix is present). As noted earlier, the use of isolation allomorphs as base allomorphs is quite characteristic.
– The second double slash specifies the context of the projected allomorph. In this case, the variable X indicates that the projected allomorph occupies nonfinal position.
– Lastly, the material at the right side of the arrow indicates what must occur in the projected allomorph as a replacement for what occurs as the initial symbol of the constraint. Here, zero must correspond to some vowel.

Thus, we can read the VOWEL RESTORATION constraint (22) as follows: "zero at the right edge of an isolation base allomorph ending in […VːC] must correspond with a vowel for allomorphs occurring with a suffix." The actual quality of the inserted vowel would be determined by additional constraints, ranked freely in order to account for free variation (see Section 2.).

A constraint like VOWEL RESTORATION has a rather crudely descriptive character. It should be remembered, however, that Anticorrespondence constraints arise, according to the present view, only in acquisitional conundrums, where historical change has created a tough data pattern. In such cases, the ordinary analytical approach of ranking faithfulness constraints against markedness constraints has already been tried by the learner and has failed. In such an instance, the simple task of collecting stem allomorphs, and the contexts where they occur, at least has the advantage of straightforwardness and directness.

Anticorrespondence constraints, quite naturally, resemble rules. A crucial difference is that they refer to surface members of a paradigm, rather than to underlying forms. In the present case, this would seem to be a clear virtue. As argued in Section 2, Yidiɲ shows more relations of predictability among surface forms than can be accounted for by the method of deriving all surface forms from a single underlying representation. In contrast, Anticorrespondence constraints may be stated for *all* the relations of mutual predictability among surface allomorphs, and thus are capable in principle of characterizing such relations completely and coherently.

At this point, I must refer the interested reader to a longish downloadable Appendix[16] for a tentative analysis of Yidiɲ that makes use of Anticorrespondence constraints.

5. Conclusions

In terms of "good news" and "bad news," the inspection of Yidiɲ data carried out here has yielded rather mixed results. While Yidiɲ emerges as an even more interesting language than has previously been thought (or so I believe), *phonology* comes out looking rather less principled and harder to do research in. If I am right in supposing that something like Anticorrespondence constraints are needed for Yidiɲ phonology, then it would seem that phonology simply cannot be done on the fully-principled basis laid out in classical Optimality Theory, and must indeed on occasion make use of a certain amount of brute-force descriptive

16. Available at the author's Web site: http://www.humnet.ucla.edu/humnet/linguistics/people/hayes/

power. Is phonology still worth doing in this case?

My own opinion is that it most certainly is, but we will have to be shrewder in the research strategies we adopt. One possibility is to locate particular empirical domains in which "pure" phonology of the type identified by classical OT can be isolated and studied; for discussion of such cases, see Hayes (1999). Another possibility is to let considerations of *learnability* serve as a guide to research. Suppose we developed the ability to write computational algorithms that could learn entire phonological systems. Any proposed model of this type faces a serious danger, namely that the number of hypotheses it must consider and test against data might be impossibly large. If it were possible to develop an algorithm that could learn, say, the Yidiɲ data pattern using only limited time and computational resources, then whatever the model assumed about the nature of phonological constraints would receive a kind of semi-empirical support. It seems likely that only a very thoughtful, highly principled theory of constraints could constrain the hypothesis set to the point that learning could occur in a feasible amount of time.

The crucial point here is that the "brute force" objection that can be made to constraints like Anticorrespondence is fundamentally an objection about learnability. It is only direct research on learning that will determine to what extent such an objection is legitimate.

The philosophical conclusion here, which is hardly a novel one, is that languages are often more interesting and challenging than our current theories allow for. The proper response when one discovers this is not (usually) to abandon the theory in despair, but to exercise some imagination and initiative: what modifications to the theory, and what novel research strategies, would bring the new phenomena into line? Ultimately, taking on such challenges seems likely to be a very fruitful approach.

Acknowledgments

I would like to thank René Kager, Donca Steriade, the volume editors, and the participants in a UCLA seminar on Correspondence Constraints for helpful input on this research project.

Surface Opacity of Metrical Structure in Optimality Theory

René Kager

1. Introduction

A central hypothesis of rule-based generative phonology is that rules have exclusive access to representations that occur in the stage of the derivation at which they apply. Whether or not a rule's conditioning environment is satisfied in the surface form is totally irrelevant to its actual application. Or to put it differently, derivational phonology is 'output-blind'. Of course it may well occur that some rule is ordered so late that no other rule will destroy its context in the output. But such surface transparency is accidental only (a mere side-effect of rule-ordering), and not principled. Analogously, derivational phonology is 'input-blind' as well, since for a rule's application it is irrelevant whether or not its structural description had been satisfied as early as in the underlying representation. In sum, rules have access to lexical representations only *indirectly*, and (in principle at least) cannot distinguish derived from non-derived properties[1].

A derivational theory with extrinsic rule ordering easily expresses generalisations that are *not* surface-true, and it specifically predicts *opaque* surface forms. I will use the notion 'opacity' as referring to situations in which a disparity exists between a surface form and the context of application of a phonological process. Two types of opacity are known from the literature (see Kenstowicz and Kisseberth 1979, and references cited there). First, a form may display the effects of some phonological rule, but fail to meet its context of application at the surface. This situation arises when, at some point in its

1. Constraining rules such that they apply only to 'derived forms' (e.g. strict cyclicity) can be seen as breaking away from derivationalism in a strict sense.

derivation, a form matches a rule's structural description (hence undergoes the structural change), after which some subsequent rule destroys the former rule's context of application. The effect is an 'overapplication' of the rule with respect to its output phonological shape. This kind of interaction is illustrated by an example from Turkish (Kenstowicz and Kisseberth 1979: 192). First a minor rule deletes the initial consonant of the possessive suffix /-sɪ/ after consonant-final stems. The triggering consonant is subsequently deleted by a second minor rule of intervocalic *k-Deletion*[2]:

(1)	/ayak-sɪ/	'his foot'
	ayak-ɪ	Consonant deletion in /-sɪ/
	aya-ɪ	k-Deletion
	[aya-ɪ]	

The second class of cases has precisely the opposite characteristics: a surface form meets the structural description of a rule, but it fails to show its effects. This situation arises when the form fails to meet a rule's structural description at its proper point of application in the derivation, and comes to match it only later, due to subsequent rules (a situation of 'counter-feeding'). The effect is an '*underapplication*' of the rule to a form, in terms of its output phonology. Consider the following example from French (Schane 1968). Schane's rule of *Nasalisation* (nasalising vowels before nasal consonants in syllable 'Coda' position) underapplies to /ɔ/ since at the point in the derivation where Nasalisation applies this vowel fails to match the rule's context. It only comes to match this context after a subsequent rule of *Schwa deletion*:

(2)	/bɔn-ə sœr/	'good sister'
	---	Nasalisation (V $\rightarrow$ [+nas] / ___ [+nas] {#, C})
	bɔn sœr	Schwa deletion
	[bɔn sœr]	(*[bɔ̃sœr])

It has been pointed out by Hooper (1976) and many others that this 'abstract' analysis only serves to maintain the otherwise unsupported assumption that nasalisation is not distinctive in vowels in French. These criticists argue that Schane's rule-ordering analysis may capture the historical developments, rather than the synchronic situation, in which nasalisation has become distinctive in vowels — due to opacity.

A special case of opacity is that involving *metrical structure*. For example,

2. Zimmer (1975) states the rule as deleting the final *k* of a polysyllabic stem when it is intervocalic.

a stress rule may place stress on a vowel at some point in the derivation, after which a subsequent rule wipes out its context of application, for example by a metathesis of consonants which affects syllable weight[3]. The example below illustrates this kind of interaction of stress and metathesis in Palestinian Arabic (Kenstowicz and Kisseberth 1979). Stress is assigned to the penult if it is heavy (e.g. *darásti* 'you fem. studied'), and otherwise to the antepenult (e.g. *dárasat* 'she studied'). Metathesis, a rule restricted to stems ending in the sequence CCVC, renders the stress pattern opaque by closing the penult:

(3)	/b-tu-drus-u/	'you masc. study it'
	b-tú-drus-u	Accent
	b-tú-durs-u	Metathesis (in CCVC stems)
	[btú-durs-u]	

Opacity of *metrical* structure appears to make a very strong case for derivational theory. In Palestinian Arabic, as in many other languages, stress is a predictable property, reflecting a fully productive set of generalisations. It's context is created by syllabification, another fully predictable property. If the stress pattern reflects a syllable structure that coincides with the segmental structure of the lexical representation, but not with that of the surface form, then it seems almost forced to assume some intermediate level in the derivation at which stress is assigned. This level cannot coincide with the 'input' (since stress is fed by syllabification), nor with the 'output' (since syllabification may have changed at the surface).

Proponents of derivational theory have not failed to point out that opacity presents difficulties to theories that make alternative assumptions, such as Hooper's (1976) 'True Generalization Condition', according to which all rules must express surface generalisations. In the seventies this issue led to an intense debate in the phonological literature, until other issues became the focus of attention. But recently, 'opacity' has become a potential issue in phonology again, due to the rise of Optimality Theory (Prince and Smolensky 1993). In a sense, OT itself is a logical development from the discussion on rule interactions in the seventies, specifically that on 'conspiracies'. It was first observed by Kisseberth (1970) that grammars (i.e. 'sets of rules') strongly favor rule interactions that produce specific output targets, or that avoid specific output characteristics. Such 'rule-conspiracies' are found in one language after another, and point

3. Some prosody-dependent processes actually *destroy their own context of application*. For example, foot-governed vowel deletion rules affect syllable structure, destroying their conditioning foot structure (see Kager 1997 for an approach in OT).

to a serious flaw of derivational theory, whose basic tenet is that rules are output-blind.

OT views the output level as the 'priviliged' level at which significant linguistic generalisations are expressed. On the view that grammars map underlying representations (inputs) into surface forms (outputs), this reduces derivations to *one-step* mappings. Any differences between the input and output (effects of rule application in derivational theory) are due to universal *well-formedness* constraints that favor (or reject) output properties. Well-formedness constraints are always in competition with a second class of constraints, the so-called *faithfulness* constraints, which require that the input and output be identical (McCarthy and Prince 1995). In a typical way, phonology (or language) is a conflict between 'contrast' and 'well-formedness'. The formal equivalent of the notion of 'conflict' in OT is constraint ranking: the best possible output form is the one that maximally satisfies higher ranked constraints, inherently at the expense of lower ranked ones. Much like derivational theory, OT finds its explanatory value in maximising the scope of linguistic generalisations. The difference between both theories resides in how interactions between generalisations take place: by linear precedence (derivational theory), or by hierarchical ranking (OT).

The cases of opacity that I have discussed under the headings of 'overapplication' and 'underapplication' pose challenges for Optimality Theory. These seem to require a level of generalisation that does not coincide with the input, nor the output. Intermediary levels of derivations are excluded by OT, under its most straightforward interpretation. Below we will see that it is actually possible to refer to a 'level' that does not occur with input, nor with output, nor with any intermediary level, but that is still independently motivated. This is the *paradigm*, the set of morphologically related forms whose output phonology can be taken as a basis of comparison. For example, we may say that a form *F* has a property *P* because there exists a morphologically related form *F'* that has this property *P*. In the example from Palestinian Arabic, finding such a related form is actually not difficult. The form *btúdursu* 'you masc. study it' is morphologically related to the form /b-tu-drus/ *btúdrus* 'you masc. study', which has stress on the 'corresponding' vowel [u]. Observe that the latter form can be viewed as the 'base' of the former in a 'compositional' way — in terms of its morpheme make up and the resulting feature composition.

The OT notion of faithfulness is generalised to cover not only identity relationships between input and output, but also between morphologically related output forms (Burzio 1994; McCarthy 1995; Benua 1995; Buckley, this volume; Orgun, this volume). Opacity, or 'underapplication' and 'overapplication' of phonology, can be said to be due to paradigm force: the domination of well-formedness constraints by *paradigm identity constraints*. The OT theory of 'opacity' therefore predicts that overapplication and underapplication always occur in the context of a morphologically related form that displays the relevant phonological property. Derivational theory makes no such prediction. A comparison of both theories should be based on various criteria, but the empirical investigation of this prediction should be among these criteria. This paper may in fact form a starting point.

This paper is organised as follows. In Section 2, I will compare derivational theory and OT with respect to paradigmatic ('transderivational') relationships, on the basis of data from Palestinian Arabic that involve opacity of vowel deletion with respect to the surface stress pattern. I will propose a definition of 'paradigmatic relatedness' that predicts the morphological relationships under which two forms display phonological identity effects. A comparison is made with derivational theory, which uses the cycle rather than paradigmatic identity. Section 3 discusses base identity effects in Tripoli Arabic, another Levantine dialect. I focus on the possibility, allowed in OT, that a form's surface phonology reflects both faithfulness constraints (input identity) and paradigmatic constraints (base identity). I will argue for parallel evaluation of output forms, in the sense that both base and lexical input are accessible simultaneously. Interactions between I/O and B/O faithfulness constraints are due to regular constraint ranking.

2. The Cycle vs. Correspondence

2.1 *Syncope in Palestinian Arabic*

In rule-based theory, transderivational relationships between morphologically related forms were in fact recognised. This notion is shaped as the *transformational cycle* (Chomsky and Halle 1968), a mode of rule application in morphologically complex words. Ordered rules (R_1-R_n) first apply to the minimal domain, and then to successively larger domains. Each language has a subset of rules that apply cyclically, and another set of rules that apply noncyclically.

A famous example of cyclic rule application (again) comes from Palestinian Arabic, and it was pointed out by Brame (1974). The stress rule interacts with a rule of *i-Syncope*, that deletes /i/ in an open unstressed nonfinal syllable. It is stated below:

(4) i-SYNCOPE
i [-stress] → Ø / ___ CV

Since i-Syncope preserves stressed vowels, it must be ordered after stress. The stress rule of Palestinian Arabic (introduced informally above for example 3) places stress on a heavy penult, and otherwise on the antepenult. Verbal forms that are inflected for subject (person, number, and gender) illustrate the application of stress and i-Syncope:

(5) *Verb plus subject suffix*

a.	/fihim/		'to understand' (verb stem)
b.i	/fihim/	fíhim	'he understood'
b.ii	/fihim-na/	fhím-na	'we understood'
b.iii	/fihim-u/	fíhm-u	'they understood'

Derivations of these forms are presented in (6):

(6)

	/fihim/	/fihim-na/	/fihim-u/
Stress	fíhim	fihím-na	fíhim-u
i-Syncope	---	fhím-na	fíhm-u
	[fíhim]	[fhím-na]	[fíhm-u]

Verbal forms containing 'accusative' suffixes display transderivational preservation of stress. These forms are inflected for person, number, and gender of the *object* by a suffix that is added to a verb form inflected for subject. Observe that bold-face [i] in the forms in (7c) fails to delete, even though it stands in an open unstressed nonfinal syllable.

(7) *Verb plus accusative suffix*

a.	/fihim/	fíhim		'he understood'
b.i	/fihim-ak/	fíhm-ak		'he understood you m.'
b.ii	/fihim-ik/	fíhm-ik		'he understood you f.'
b.iii	/fihim-u/	fíhm-u		'he understood him'
c.i	/fihim-ni/	f**i**hím-ni	*fhím-ni	'he understood me'
c.ii	/fihim-ha/	f**i**hím-ha	*fhím-ka	'he understood her'
c.iii	/fihim-na/	f**i**hím-na	*fhím-na	'he understood us'

Brame observes that the accusative forms in which [i] fails to delete are all based on a free form (7a) in which [i] is stressed. The generalisation is that i-Syncope 'under-applies' to forms that have a morphological base form in which [i] is stressed.

The correctness of the generalisation is shown by possessives, which have morphologically related nouns in which the relevant [i] is stressed (Kenstowicz and Abdul-Karim 1980):

(8) *Possessives*

a.	/birak/	bírak		'pools'
b.	/birak-u/	bírak-u		'his pools'
c.	/birak-na/	b**i**rák-na	*brák-na	'our pools'

Again, underapplication of i-Syncope correlates with the addition of a (possessive) suffix to a free form (or 'base') that is inflected itself.

Underapplication of i-Syncope cannot be due to an actual (secondary) stress on the initial syllable of accusatives and possessives, cf. [f**ì**hím-na] and [b**ì**rák-na]. Kenstowicz and Abdul-Karim (1980) found that for speakers of Palestinian Arabic the analogous form that is based on a CaCaC verb, e.g. [ḍaráb-na], is *ambiguous* between 'we hit' and 'he hit us'. This makes the property that is responsible for blocking i-Syncope in accusatives and possessives 'abstract' to a certain degree.

In derivational theory the notion of 'relatedness' between forms is characteristically modelled in derivational terms, that is, *cyclically*. Brame (1974) assumes that accusatives and possessives have an additional internal layer of morphological structure, which triggers a cyclic application of the stress rule. The cyclically assigned stress on the first syllable of [fíhim] is carried over to the second cycle [[fìhím]-na], even though it is subordinated to the new main stress. The secondary stress protects the initial syllable's vowel against postcyclic i-Syncope. Finally, initial secondary stress is erased by a postcyclic rule of *Destressing*:

(9)

Input	$[\text{fihim-na}]_{\text{Subj}}$	$[\text{fihim-u}]_{\text{Subj}}$	$[[\text{fihim}]\text{na}]_{\text{Acc}}$	$[[\text{fihim}]\text{u}]_{\text{Acc}}$
Cycle1				
Stress	fihím-na	fíhim-u	fíhim	fíhim
Cycle2				
Stress	---	---	fìhím-na	(*vacuous*)
Postcyclic				
i-Syncope	fhím-na	fíhm-u	*blocked*	fíhm-u
Destressing	*n.a.*	*n.a.*	fihím-na	*n.a.*
Output	fhímna	fíhmu	fihímna	fíhmu

This analysis was regarded as strong evidence for extrinsic rule ordering, in the sense that a phonological property that is acquired in the course of the derivation (stress) blocks a rule that is sensitive to its presence (i-Syncope), although it is *absent* from the surface form due to a subsequent rule that deletes it (Destressing). This is achieved by linearly ordered rules that are 'blind' to underlying representations ('no globality'), and have access only to the representation that arises at the point in the derivation at which they apply.

How could this underapplication of i-Syncope be analysed in OT, a theory that has no derivations, hence lacks the intermediary level of representation which Brame argued for? The central idea is that underapplication is due to paradigm regularity.[4] More specifically, a *well-formedness* constraint that militates against light syllables is dominated by an '*identity*' constraint requiring that vowels which are 'prosodic heads' in basic forms should have 'correspondents' in morphologically related forms.

Let us first carefully consider the assumptions and general theoretical framework in which this analysis is embedded. Setting up identity constraints that compare the identity of morphologically related forms requires the notion of *correspondence*. McCarthy and Prince (1995), in a work that introduces the notion, define it as a relationship between elements that are part of two strings. For example, between elements in an input string and elements in an output string:

(10) I/O-CORRESPONDENCE
Given two strings S_1 and S_2, **correspondence** is a relation $\Re$ from the elements of S_1 to those of S_2. Segments α (an element of an input string S_1) and β (an element of an output string S_2) are referred to as **correspondents** of one another when $\alpha\Re\beta$.

Correspondence relationships hold between segments (an extension to prosodic elements is proposed by McCarthy 1995a). The actual constraints which produce faithfulness effects between Input and Output are of the types in (11):

(11) a. DEPENDENCE
Every element of S_2 has a correspondent in S_1.
b. MAXIMALITY
Every element of S_1 has a correspondent in S_2.
c. IDENTITY(γF)
Let α be a segment in S_1 and β be a correspondent of α in S_2. If α is [γF], then β is [γF].

In sum, a central assumption of Correspondence Theory is that constraint-based evaluation of an output form may have direct access to its input lexical representation. But this notion of correspondence has been generalised to relationships between an Output form and other Output forms. In particular, constraint

4. Since this paper was first presented (at the Tilburg conference "The derivational residue", fall 1995) two researchers have, independently, proposed a similar analysis of the Palestinian data: Kenstowicz (1996) and Steriade (1996). Moreover, Orgun (1996) has developed a declarative theory of cyclic phenomena that, in principle at least, offers an alternative approach to the data discussed here.

evaluation of an output candidate may have access to the output of a morphologically related output form — its 'base'. This is so-called B/O-Correspondence (Benua 1995; McCarthy 1995a; Burzio 1994)[5].

Benua (1995: 51) argues for English that "Class 2 affixation is derived through an O/O-correspondence with the unaffixed word." An example is a phonological process that is typical of New York-Philadelphia English: æ-Tensing in closed syllables, e.g. *pass* [pEs], but *passive* [pæ.sɪv]. In forms that contain Class 2 affixes, such as *-ing*, this 'overapplies' in the sense that the vowel in the open syllable of *passing* [pE.sɪŋ] surfaces as tense. Benua argues that this overapplication in *passing* is due to its relatedness to its base, *pass* [pEs]. In a diagram, this can be portrayed as follows ('B' abbreviates the base, while 'A' abbreviates 'affixed form'):

(12) *B/A-Identity* (B=Base, A=Affixed form)

[pEs] ----------- [pE.sɪŋ]

I/O-Faith |

/pæs/ (cf. *passive*)

Two constraint interactions are relevant. First, the well-formedness constraint $*æC]_\sigma$ that rules out [æ] in closed syllables outranks a well-formedness constraint *TENSE-low, requiring low vowels to be lax. This ranking is supported by the observation that [pEs] is selected as the optimal candidate, rather than [pæs]. Second, the B/A-identity constraint with respect to the feature [tense] outranks *TENSE-low. This is supported by the selection of [pE.sɪŋ] rather than [pæ.sɪŋ].

Here and in the rest of this paper, 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.

This definition of 'base' is precise enough to capture the distinction among verbal forms in Palestinian Arabic in the way that is required. Subject forms such as [fhím-na] 'we understood' (5bii) have no base, since no free form occurs in the language that matches the criterion of semantic compositionality. In particular, the verb stem /fihim/, which would be appropriate in the compositional sense, fails to occur without inflection, failing the second criterion for

5. The terminology of 'base' has been chosen to reflect the strong similarities with other types of O/O-Correspondence, for example reduplication and truncation.

'base-hood'. Nor can the free form [fíhim] 'he understood' serve as a base, since this is not compositionally related to [fhím-na] 'we understood'. In contrast, all object forms have a base by these criteria. For example, [fihím-na] 'he understood us' has as its base the free form [fíhim] 'he understood', of which it contains all grammatical features.

The generalisation is that i-Syncope 'underapplies' to vowels that are *stressed* in the *base*:

(13) a. [f**í**him] 'he understood' ↔ [f**i**hím-na] 'he understood us'
b. [b**í**rak] 'pools' ↔ [b**i**rák-na] 'our pools'

From here on I will use graphic means to indicate the correspondence relationships between an affixed form and (on the one hand) its input, and (on the other hand) its base. I will mark these relationships by vertical lines between correspondents at three levels (Input, Output, and Base). In (14a), bold-face [**i**] indicates the underapplication of i-Syncope in the output:

(14)

```
a.  'he understood us'    b.  'we understood'
    /f i h i m -n a/          /f i h i m -n a/    Input
     | | | | |   | |           |   | | |   | |    I/O Correspondence
    [f i h í m -n a]          [f   h í m -n a]    Output
     | | | | |                                    B/O Correspondence
    [f í h i m]                                   Base
    'he understood'
```

The correspondence-based perspective of this pattern is that syncope 'underapplies' in the accusative and possessive because the relevant vowels have *stressed correspondents* in the base. This requires an extension of correspondence to stress properties of segments. Such an extension was proposed by McCarthy (1996b), Alderete (1995b, this volume), and Kager (forthcoming):

(15) HEAD-MAX(B/O)
Every segment in the base prosodic head has a correspondent in the output.

The effects of what I have called 'i-Syncope' so far are due to a well-formedness constraint that disallows [i] in open syllables[6].

(16) NO [i]
/i/ is not allowed in light syllables.

Note that no reference is made to stress, a result that will be confirmed in Section 2.2, where we discuss the interaction between syncope, stress and epenthesis.

NO [i] obviously outranks the identity constraint requiring that input segments have correspondents in the output (cf. McCarthy and Prince 1995):

(17) MAX(I/O)
Every segment in the input has a correspondent in the output.

This ranking is motivated by the fact that deletion is possible in forms such as [fhím-na] 'we understood', which satisfy NO [i] at the expense of MAX(I/O). The constraint ranking that now arises is (18):

(18) HEAD-MAX(B/O) » NO [i] » MAX(I/O)

The analysis is illustrated by the tableaux (19–20). Only candidates are considered that satisfy the canonical stress patterns of Palestinian Arabic; the issue of how stress is predictable by constraints, and its interaction with the processes of syncope and epenthesis will be taken up in Section 2.2 below.

Tableau (19) illustrates the underapplication of i-Syncope due to base identity, while tableau (20) shows how syncope occurs in a form that lacks a base:

(19)

Input: /fihim, -na/ Base: [fí.him]	HEAD-MAX(B/O)	NO [i]	MAX(I/O)
a. ☞[fi.hím.na]		*	
b. [fhím.na]	*!		*

6. This constraint is modelled after an analogous constraint NO [a] in Orgun (1995a). An interesting question is whether NO [i] can be decomposed into general constraints which, taken together, produce its effects. I believe that this is possible, if we assume a constraint against monomoraic syllables (argued for by Broselow 1992: 32), and rank this in between faithfulness constraints for specific vowels, e.g. IDENT-[a] » $*\sigma_{\mu}$ » IDENT-[i]. I will not pursue this issue here, and maintain the formulation of NO [i] as it is.

(20)

Input: /fihim, -na/ Base: *none*	HEAD-MAX(B/O)	NO [i]	MAX(I/O)
a. [fi.hím.na]		*!	
b. ☞ [fhím.na]			*

Turning to the forms that have vowel-initial suffixes, we find that application of i-Syncope is blocked in neither form.

(21)

Input: /fihim, -u/ Base: [fí.him]	HEAD-MAX(B/O)	NO [i]	MAX(I/O)
a. [fí.hi.mu]		*!*	
b. ☞ [fíh.mu]			*
c. [fhí.mu]	*!	*	*

(22)

Input: /fihim, -u/ Base: *none*	HEAD-MAX(B/O)	NO [i]	MAX(I/O)
a. [fí.hi.mu]		*!*	
b. ☞ [fíh.mu]			*
c. [fhí.mu]		*!	*

If we compare this output-based analysis of (22) to the derivational analysis of such forms, we find that the latter analysis imposes a condition on i-Syncope that it applies in unstressed syllables. This condition protects the vowel in the initial syllable of (22) against i-Syncope, in order to rule out the incorrect form *[fhímu]. But in the OT analysis, no need arises for such a condition, and actually it would be impossible to state it. It is impossible to refer to lack of stress in the target vowel of syncope, since this vowel does not appear in the output. Nor is there an input correspondent that is stressed, and to which the blocking of syncope could be attributed through a 'faithfulness' constraint. But on the other hand, the OT analysis simply does without a condition on the stress value of the syncopated vowel. It blocks *[fhí.mu] by the very constraint that triggers syncope in the first place: No [i]. If the 'goal' of syncope is avoidance of light syllables containing [i], then it is preferable to maximally achieve this goal in the output (cf. [fíh.mu], 22b), rather than succeeding only halfway (cf. 22c).

From the above discussion and analysis, a number of preliminary conclusions can be drawn. First, the notion of 'base' in B/O-correspondence is firmly linked to compositionality. Second, the requirement that the base must be an output form correctly predicts that trans-derivational relationships must involve output forms. This prediction does not follow from derivational theory, since what counts as a *cycle* is inherently unconstrained by the criterion of 'occurrence as a free form'. In fact, any layer of morphological structure regardless of its relationship to word morphology is predicted to display cyclic properties. In order to obtain the relevant distinction between subject morphology and object morphology in Palestinian Arabic verbs, Brame had to make the arbitrary assumption that only the latter type involves cyclically layered structure. If matters had been reversed (with subject morphology invoking cyclic layering), this would have been equally natural on the derivational analysis. Third, the underapplication of a phonological 'process' is modelled as a domination of 'base identity' constraints over well-formedness constraints in the hierarchy.

2.2 *Opacity of Metrical Structure Due to i-Epenthesis*

We now turn to opacity of metrical structure, which arises by various modifications of the syllable structure on which stress is based. I will first discuss the process of epenthesis that is the main source of metrical opacity, and the way in which it is affected by base identity and faithfulness. Then I will discuss the metrical constraints proper, and rank them with respect to those responsible for vowel-zero alternations and base identity effects.

Like the Levantine dialects to which it is related, Palestinian Arabic has a process of i-*Epenthesis*, which inserts [i] between the first and second consonant in a sequence of three consonants, or between two consonants at the end of the word:

(23) i-EPENTHESIS
$\emptyset \rightarrow$ i / C ___ C {C, #}

This process is the source of opaque stress in examples such as those in (24a.iii, b.iii) below:

(24)	a.i	/fihm/	fí.h**i**m	'understanding'
	a.ii	/fihm-u/	fíh.mu	'his understanding'
	a.iii	/fihm-na/	fí.h**i**m.na *fi.h**í**m.na	'our understanding'
	b.i	/ʔakl/	ʔá.k**i**l	'food'
	b.ii	/ʔakl-u/	ʔák.lu	'his food'

b.iii /ʔakl-ha/ ʔá.k**i**l.ha *ʔa.k**í**l.ha 'her food'

In a derivational analysis, this interaction involves ordering i-Epenthesis *after* stress. Cyclic application of stress is essentially irrelevant to its interaction with epenthesis, but for reasons of similarity to earlier derivations I have indicated it in the derivations below:

(25)			/fihm/	/fihm-u/	/fihm-na/
	Stress	Cycle 1	fíhm	fíhm	fíhm
		Cycle 2	fíhm	fíhm-u	fíhm-na
	i-Epenthesis		fíh**i**m	---	fíh**i**m-na
			[fíhim]	[fíhmu]	[fíhimna]

Again the question arises as to how OT captures the opacity. Let us first see what constraint interaction produces epenthesis of [i].

A 'repair' in the form of vowel epenthesis is triggered by a combination of two high-ranked constraints on syllable well-formedness (26a-b), which I will combine into a cover constraint *CLUSTER:

(26) *CLUSTER
a. *COMPLEX-CODA (cf. Prince and Smolensky 1993)
No complex syllable codas.
b. SONORITY SEQUENCING PRINCIPLE (cf. Clements 1990)
Segments of higher ranking sonority stand closer to the centre of the syllable than segments of lower ranking sonority.

Avoidance of syllabic ill-formedness takes priority over avoidance of epenthetic vowels. Or, to put it into terms of constraint ranking: *CLUSTER » DEP(I/O).[7]

Violations of *CLUSTER can be avoided in two different ways, that is, by epenthesis between the first and the second consonant, or after the second consonant.

(27) a. C_CC fí.h**i**m.na
b. CC_C *fíh.m**i**.na

In a derivational analysis, the locus of epenthesis is stated in the rule, or derived indirectly by a statement that epenthesis applies *directionally* (cf. Itô 1989). But neither of these options is available in an OT analysis. Once more, base identity seems to be involved. If we consider the output form [fí.him.na], we find that

7. Notice that at the word beginning, complex onsets are allowed, and actually created by the deletion of input vowels (cf. [fhím-na]).

this maximally resembles the base in the sense that the vowels that appear in the base [fí.him] are both preserved in the output. The fact that in (28b) the underlined vowel is 'epenthetic' with respect to the *input* is totally irrelevant for *base* identity:

```
(28)   a.   /f i h   m -n a/       b.   /f i h   m   -n a/      Input
             | | |    |  | |              |   |    |     | |
            [f í h i m -n a]            [f í h   m i -n a]      Output
             | | | | |                    | | |    |
            [f í h i m]                 [f í h i m]             Base
```

Correspondence between base and output is broken in (28b), where the epenthetic vowel has no correspondent in the base. This points to the following base-identity constraint requiring that base segments have correspondents in the output:

(29) MAX(B/O)
Every segment in the **base** has a correspondent in the **output**.

This constraint is more general than HEAD-MAX(B/O), since it does not mention the stressed status of the vowel in the base. MAX(B/O) is also distinguished from HEAD-MAX(B/O) in its position in the constraint hierarchy, specifically with respect to NO [i]. While the latter must be ranked above NO [i], the former ranks below it, as the following syncope data show:

(30)	a.	HEAD-MAX(B/O) » NO [i]	/fihim-na/ 'he understood us' [fihím-na] > (*[fhím-na])
	b.	NO [i] » MAX(B/O)	/fihim-u/ 'he understood him' [fíhm-u] > (*[fíhim-u])

The next question is: if the locus of epenthesis in [fi.h**i**m.na] 'our understanding' depends on that in its base [fí.h**i**m] 'understanding', then what predicts the epenthesis site in the base? Here we must consider two candidates, one with stem-internal epenthesis, [fí.h**i**m], the other with stem-external epenthesis, [fíh.m**i**]. The former is selected by a well-known generalised alignment constraint (McCarthy and Prince 1993):

(31) ALIGN-R
$]_{Stem} =]_{\sigma}$

This constraint is undominated in Palestinian Arabic, as far as I can determine.

In Correspondence Theory, what it means for [i] to be the epenthetic vowel is a matter of strictness of 'identity' between an output vowel and its input correspondent. That is, it is ‘less costly’ to have [i] as an output vowel lacking an input correspondent, than it is to have any other vowel in such a situation. This result follows from the ranking of the O/I-Correspondence constraint IDENT-[a] at the top of the constraint hierarchy, guaranteeing that every surface [a] has an input correspondent. Accordingly, the choice of epenthetic vowel is restricted to [i] (and possibly [u][8]). See (32):

(32) a. IDENT-[a](O/I)
Output [a] must have an input correspondent.
b. IDENT-[a] » IDENT-[i], IDENT-[u]

In the tableaux below, I will not include this ranking, and tacitly assume it by including only [i] as an epenthetic vowel.

We are now in a position to localise the source of metrical opacity. In Palestinian Arabic, heavy penultimate syllables are stressed in the ‘regular’ case. Then why is epenthetic [**i**] in [fí.h**i**m.na] and [ʔá.k**i**l.ha] *unstressed* even though it stands in a closed penult? Actually the rejection of stress by epenthetic vowels is a cross-linguistically common phenomenon (cf. Piggott 1995; Alderete 1995b). Accordingly, Alderete (1995b, this volume) proposes a constraint to the effect that stressed vowels must have input correspondents:

(33) HEAD-DEP(I/O)
Every vowel in the output prosodic head has a correspondent in the input.

Epenthetic [**i**] in [fí.h**i**m.na] lacks a correspondent in the input, as illustrated by the diagram in (34a). Therefore stressing it would violate HEAD-DEP(I/O). In contrast, the stressed vowel in [ba.kár.na] ‘our cattle’ (34b), with a base [bá.kar] ‘cattle’, has an input correspondent and hence stressing it does not violate HEAD-DEP(I/O).

(34)

	a.	b.	
	/f i h m -n a/	/b a k a r -n a/	Input
	[f í h i̲ m -n a]	[b a k á r -n a]	Output
	[f í h i m]	[b á k a r]	Base

Indirectly, form (34b) rules out an alternative hypothesis about opaque stress in [fí.h**i**m.na], according to which a vowel that is stressed in the output must have

8. See Abu-Salim (1980) for data showing that [u] patterns much like [i] in syncope and epenthesis.

a *stressed* correspondent in the base. This hypothesis is ruled out since the stressed output vowel of form (34b) has no stressed correspondent in the base.

'Metrical opacity' is modelled as a domination of a faithfulness constraint over a well-formedness constraint. Here HEAD-DEP(I/O) dominates the well-formedness constraint that is responsible for stress on heavy (penultimate) syllables (the Weight-to-Stress Principle, or WSP, Prince and Smolensky 1993; more will be said about this constraint directly below):

(35) HEAD-DEP(I/O) » WSP

The reverse ranking would produce penultimate stress, e.g. *[fih**í**m-na], rather than [fíh**im**-na]. I rank HEAD-DEP(I/O) directly above WSP, postponing arguments for this ranking to Section 3.

Before I discuss the actual metrical constraints in Section 2.3, I will first demonstrate the analysis of vowel-zero alternations that we have available now by three tableaux of the minimal contrasts in [fihím-na] 'he understood us' (36), [fhím-na] 'we understood' (37), and [fíhim-na] 'our understanding' (38).

(36)

I: /fihim-na/ B: [fí.him]	*CLUSTER	HEAD-MAX(B/O)	NO [i]	HEAD-DEP(I/O)	MAX (B/O)	WSP
a. ☞[fi.hím.na]			*			
b. [fí.him.na]			*			*!
c. [fíh.mi.na]			*		*!	
d. [fhím.na]		*!			*	
e. [fíhm.na]	*!				*	

(37)

I: /fihim-na/ B: *none*	*CLUSTER	HEAD-MAX-(B/O)	NO [i]	HEAD-DEP(I/O)	MAX (B/O)	WSP
a. [fi.hím.na]			*!			
b. [fí.him.na]			*!			*
c. [fíh.mi.na]			*!			
d. ☞[fhím.na]						
e. [fíhm.na]	*!					

(38)

I: /fihm-na/ B: [fí.him]	*COMPLEX	HEAD-MAX(B/O)	NO [i]	HEAD-DEP(I/O)	MAX (B/O)	WSP
a. [fi.hím.na]			*	*!		
b. ☞[fí.him.na]			*			*
c. [fíh.mi.na]			*		*!	
d. [fhím.na]		*!			*	
e. [fíhm.na]	*!				*	

Finally we are in a position to substantiate the claim that stress opacity of Palestinian Arabic involves the parallel evaluation of faithfulness ('input-identity') and paradigm regularity ('base-identity'). Parallelism is demonstrated by the activity, within the same constraint hierarchy, of constraints evaluating I/O-correspondence and B/O-correspondence. Note that within one output form [fí.h**i**m.na] (38), the boldface vowel is treated as epenthetic with respect to the input (which is why it cannot surface as stressed — due to HEAD-DEP(I/O)), while it is paradoxically treated as non-epenthetic with respect to the base (which is why it must be retained — due to MAX(B/O)).

2.3 *The Stress Pattern of Palestinian Arabic*

In order to fathom to which extent metrical constraints are subordinated to correspondence constraints, we must now take a closer look at the Palestinian Arabic stress system. Section 2.4 will integrate these constraints with the correspondence constraints of Section 2.2.

Main stress falls on the penultimate syllable if it is heavy, otherwise on the antepenult (as in Latin, the difference being that final 'superheavy' syllables may be stressed, see (39e)).

(39)	a.	ḍá.rab	'he hit'	(K 207)
		báa.rak	'he blessed'	(K 207)
		ʕál.lam	'he taught'	(K 207)
	b.	ḍa.ra.bu	'they hit'	(K 207)
		dá.ra.sat	'she studied'	(K&K 230)
		ʕál.la.mu	'they taught'	(K 207)
		báa.ra.ku	'they blessed'	(K 207)

c.	da.rá.sa.tu	'she studied it'	(K&K 230)
	baa.rá.ka.tu	'she blessed him'	(K 207)
	ʕal.lá.ma.tu	'she taught him'	(K 207)
d.	da.rás.ti	'you fem. studied'	(K&K 229)
	ʔa.zúu.ru	'I visit him'	(K&K 229)
	ḍar.bát.na	'she hit us'	(K 207)
	baa.ra.kát.na	'she blessed us'	(K 207)
e.	da.rást	'I studied'	(K&K 229)
	ka.máan	'also'	(K&K 229)

According to (Kenstowicz 1983: 208), "reliable judgments of secondary stress are difficult to obtain." Apparently there is no secondary stress in immediately pretonic position[9]. This I attribute to an undominated constraint *CLASH.

(40) *CLASH
No adjacent syllables are stressed.

This accords with the rule of prestress destressing in Kenstowicz (1983)[10].

The foot of Palestinian is the quantitative or *moraic trochee* (McCarthy 1979; Hayes 1995a)–(**L**L) or (**H**). With Dresher and Lahiri (1991) and Kiparsky (1995b), I assume that a third foot is universally analysable as a moraic trochee, i.e. the 'resolved' foot (**L**H) that is composed of a light plus heavy sequence. Evidence for (**L**H) as a unit that is quantitatively equivalent to (**H**) and (**L**L) comes from Old English (cf. Dresher and Lahiri 1991), where high vowels were deleted that immediately followed precisely these three quantitative sequences, but not any other sequence, such as HL or HH:

(41)	a.	'Pure'	(**H**)	(**L**L)	but *not*	*(**H**L)
			[(wór).du]	[(wé.ru).du]		[(níi).(tè.nu)]
	b.	'Resolved'	(**L**H)		but *not*	*(**H**H)
			[(fǽ.rel).du]			[(fúl).(wìh).tu]

---

9. This is diagnosed by the fact that long vowels in pretonic position shorten (Abu-Salim 1983):

(i)	a.	báab	'door'	c.	ma.káa.tib	'offices'
	b.	bab-éen	'two doors'	d.	ma.ka.tíb.na	'our offices'

In the Tripoli dialect, to be discussed in Section 3, an analogous process occurs by which /a/ reduces to [i] in unstressed syllables. Again, reduction applies to pretonic heavy syllables (Kenstowicz & Abdul-Karim 1980).

10. Kenstowicz and Abdul-Karim (1980) assume that all non-primary stresses are suppressed by a rule of stress deletion.

This resolved foot occurs under duress, and is involved in ‘opaque stress’, as will be shown below. These three feet are precisely the feet that occur extensively in Arabic morphology (McCarthy and Prince 1990).

Accordingly, I assume a high-ranking constraint FOOT-FORM:

(42) FOOT-FORM
Feet are moraic trochees (H), (LL) or (LH).

The fact that the resolved foot (**L**H) is observeable only under ‘special’ conditions is due to the Weight-to-Stress Principle (Prince and Smolensky 1993):

(43) WSP
Heavy syllables are prominent within the foot and on the grid.

The WSP is the constraint responsible for stress on heavy penults, e.g. [da.(rás).ti], rather than *[(dá.ras).ti].

In surface forms of Palestinian Arabic, stress always appears on one of the final three syllables of the word (Kenstowicz 1983: 207). Independent evidence bearing on this (cited by Kenstowicz) is the fact that English words ending in four light syllables are mispronounced so as to match the window requirement, e.g. *necéssary, obligátory*[11]. Antepenultimate stress in words ending in two or more light syllables is due to NONFINALITY:

(44) NONFINALITY
The head of the PrWd must not be final.

According to Prince and Smolensky (1993), who argue for NONFINALITY on the basis of the Latin stress pattern, it serves two purposes. First, the final syllable must not be the head of the main stress foot, and second, the final syllable must not be part of the main stress foot (‘no final foot’). Violations are counted separately for these two requirements.

---

11. The Palestinian speaker consulted by Halle and Kenstowicz (1989) produced initial stress in Classical Arabic forms such as [ʃájaratun] ‘a tree’, which apparently argues against a trisyllabic window. However, the argument is flawed by the fact that this form in all respects (including stress) matches the classical language. Presumably any educated speaker would know the classical pronunciation.

(45)

I: /darasat/	*CLASH	FOOT-FORM	NON-FINALITY	WSP	PARSE-σ
a. ☞[(dá.ra).sat]				*	*
b. [da.(rá.sat)]			*!	*	*
c. [da.ra.(sát)]			*!*		**
d. [(dá.ra).(sát)]			*!*		

(46)

I: /baarak/	*CLASH	FOOT-FORM	NON-FINALITY	WSP	PARSE-σ
a. ☞[(báa).rak]				*	*
b. [baa.(rák)]			*!*	*	*
c. [(báa.rak)]		*!	*	*	
d. [(bàa).(rák)]	*!		**		

For CvCvC words NONFINALITY *forces* a (LH) trochee:

(47)

I: / ḍarab/	*CLASH	FOOT-FORM	NON-FINALITY	WSP	PARSE-σ
a. ☞[(ḍá.rab)]			*	*	
b. [ḍa.(ráb)]			**!		*
c. [(ḍá).rab]		*!		*	*

But NONFINALITY is insufficient to predict stress in words that end in three or more light syllables, such as [da.(rá.sa).tu]. This is due to PARSE-2 (Kager 1994):

(48) PARSE-2
One of two adjacent syllables must be parsed by a foot.

For reasons that will become clear immediately below, PARSE-2 must be ranked below NONFINALITY, but above WSP. The tableau in (49) shows how antepenultimate stress is predicted in [da.(rá.sa).tu]:

(49)

I: /darasatu/	*CLASH	FOOT-FORM	NON-FINALITY	PARSE-2	WSP	PARSE-σ
a. ☞ [da.(rá.sa).tu]						**
b. [(dá.ra).sa.tu]				*!		**
c. [da.ra.(sá.tu)]			*!	*		**
d. [(dà.ra).(sá.tu)]			*!			

The next tableau shows why PARSE-2 must dominate WSP:

(50)

I: /baarakatu/	*CLASH	FOOT-FORM	NON-FINALITY	PARSE-2	WSP	PARSE-σ
a. ☞[baa.(rá.ka).tu]					*	**
b. [(báa).ra.ka.tu]				*!*		***
c. [(bàa).ra.(ká.tu)]			*!			*
d. [(bàa).(rá.ka).tu]	*!					*

With undominated FOOT-FORM, ruling out (HL) trochees, PARSE-2 is necessarily dominated by NONFINALITY.

(51)

I: /baaraku/	*CLASH	FOOT-FORM	NON-FINALITY	PARSE-2	WSP	PARSE-σ
a. ☞[(báa).ra.ku]				*		**
b. [baa.(rá.ku)]			*!		*	*
c. [(báa.ra).ku]		*!				*
d. [(bàa).(rá.ku)]	*!		*			

Forms such as [da.(rás).ti] show that WSP must dominate PARSE-σ:

(52)

I: /darasti/	*CLASH	FOOT-FORM	NON-FINALITY	PARSE-2	WSP	PARSE-σ
a. ☞[da.(rás).ti]						**
b. [(dá.ras).ti]					*!	*
c. [da.(rás.ti)]		*!	*			

In [ḍar.(bát).na] two undominated constraints (*CLASH and FOOT-FORM) force a parsing in which one of the heavy syllables is outside the foot. The choice which one is unstressed (and which one is stressed) is made by PARSE-2:

(53)

I: /ḍarbatna/	*CLASH	FOOT-FORM	NON-FINALITY	PARSE-2	WSP	PARSE-σ
a. ☞[ḍar.(bát).na]					*	**
b. [(ḍár).bat.na]				*!	*	**
c. [(ḍár.bat).na]		*!			*	*
d. [(ḍàr).(bát).na]	*!					*

Words ending in CvvC# and CvCC# are stressed on their final syllables. Kenstowicz (1986) argues that the final consonant in these sequences is *extra-prosodic*, that is, outside the syllable. The preceding syllable may thus satisfy NONFINALITY if stressed (Hayes 1995).

(54)

I: /darast/	*CLASH	FOOT-FORM	NON-FINALITY	PARSE-2	WSP	PARSE-σ
a. ☞[da.(rás).t]						*
b. [(dá.ras).t]					*!	

A summary of rankings and an illustration by some forms on which they are based is given below:

(55)	a.i	*Clash » WSP	[ḍar.(bát).na] > [(ḍàr).(bát).na]
	a.ii	*Clash » Parse-σ	[ar.(bát).na] > [(ḍàr).(bát).na]
	b.i	Foot-Form » NonFinality	[(ḍá.rab)] > [(ḍá).rab]
	b.ii	Foot-Form » Parse-2	[(mák).ta.bi] > [(mák.ta).bi]
	b.iii	Foot-Form » WSP	[(ḍá.rab)] > [(ḍa.ráb)]
	b.iv	Foot-Form » Parse-σ	[ḍar.(bát).na] > [(ḍár.bat).na]
	c.i	NonFinality » Parse-2	[(mák).ta.bi] > [mak.(tá.bi)]
	c.ii	NonFinality » WSP	[(ḍá.rab)] > [ḍa.(ráb)]
	c.iii	NonFinality » Parse-σ	[(mák).ta.bi] > [mak.(tá.bi)]
	d.	Parse-2 » WSP	[baa.(rá.ka).tu] > [(báa).ra.ka.tu]
	e.	WSP » Parse-σ	[da.(rás).ti] > [(dá.ras).ti]

These rankings of individual constraints are integrated into the total ranking below:

(56) *Clash, Foot-Form » NonFinality » Parse-2 » WSP » Parse-σ

2.4 *Integrating the Metrical Constraints and the Correspondence Constraints*

What remains to be demonstrated is that the metrical constraints of the previous section can be integrated into a total ranking together with the correspondence constraints that I argued for in Section 2.2 on the basis of vowel-zero alternations and identity effects. The hierarchy that was reached at the end of Section 2.2 is repeated below:

(57) *Cluster, Head-Max(B/O) » No [i], Head-Dep(I/O) » Max(B/O), WSP

Factoring out the undominated constraints from both hierarchies (the metrical hierarchy and (57)), we find that we face the task of integrating the following two partial rankings:

(58) a. NonFinality » Parse-2 » WSP » Parse-σ
b. No [i], Head-Dep(I/O) » Max(B/O), WSP

An interesting aspect of this task is the fact that vowel-zero alternations may be partially conditioned by metrical constraints. That is, the (non-)application of syncope and epenthesis strives towards specific metrical targets in the output.

For instance, consider the fact that syncope is blocked in the first vowel in /fihim/ (*[(fhím)]), even though this would satisfy No [i]. This candidate is rejected by a metrical constraint, viz. NonFinality. Formally this means that

NonFinality dominates No [i], as in (59). We have seen no evidence (yet) to rank No [i] with respect to Parse-2 and Head-Dep(I/O), nor to rank Max(B/O) with respect to WSP:

(59)

I: /fihim/ B: *none*	Non-Finality	No [i]	Parse-2	Head-Dep(I/O)	Max (B/O)	WSP	Parse-σ
a. ☞[(fí.him)]	*	*				*	
b. [(fhím)]	**!						

Another illustration of metrical conditioning of vowel-zero alternations is that of imperfect verbs with the input shape /yi-CCvC-v/, such as /yi-ktib-u/ [yí.k**i**t.bu] 'they write'. Here i-Syncope interacts with i-Epenthesis in a very interesting way. The deletion of /i/ in an open syllable is compensated for by epenthesis of [i], thus avoiding syllable ill-formedness. Observe that it would apparently have been more 'economic' to retain the input vowel in its proper position (*[yík.ti.bu]), rather than to syncopate it, only to substitute another [i] by epenthesis. (This interaction has been called 'promiscuous syncope' by Broselow 1992). The interaction of syncope and epenthesis produces a striking confirmation of the sub-ranking of Parse-2, Head-Dep(I/O) » WSP, which was originally motivated for stress 'proper':

(60) I: /yi-ktib-u/ B: *none*	Non-Finality	No [i]	Parse-2	Head-Dep (I/O)	Max (B/O)	WSP	Parse-σ
a. ☞[(yí.k**i**t).bu]		*				*	*
b. [yi.(k**í**t).bu]		*		*!			**
c. [(yík).ti.bu]		*	*!				**
d. [yik.(tí.bu)]	*!	*				*	*

Notice that a derivational theory is faced with an uncomfortable uncertainty: is this a combined case of i-Syncope and i-Epenthesis, or is it a single metathesis rule? Kenstowicz and Kisseberth (1979: 230) select the second descriptive option: "The Palestinian dialect also has a rule whereby a stem ending in the sequence CCVC is metathesised to CVCC when a vowel-initial suffix is added." Additional examples of 'metathesis' are given in (61):

(61)	a.	/simsim-e/	[sí.mis.me]	'a sesame seed'
	b.	/b-tu-drus-i/	[btú.dur.si]	'you fem. study'
	c.	/zuʔruḍ-a/	[zú.ʔur.ḍa]	'a bee'

We now see the pressure behind 'metathesis': avoiding violation of two metrical constraints, NonFinality and Parse-2. Actually the forces that lead to 'metathesis' are just the same as those leading to other vowel-zero alternations that are named 'syncope' and 'epenthesis' in a derivational theory. The OT analysis unifies the formerly disjoint conditions under which syncope and epenthesis apply. Finally, note that the choice of the epenthetic vowel in (61), [i] or [u], depends on the rounding of the vowel that precedes (Abu-Salim 1980). This is due to a harmony constraint that I will not state here.

The final ranking appears in (62):

(62) *Clash, Foot-Form, *Cluster, Head-Max(B/O) » NonFinality » No [i], Parse-2, Head-Dep(I/O) » Max(B/O), WSP » Parse-σ

This concludes the analysis of metrical opacity in Palestinian Arabic. I have argued that metrical opacity, a phenomenon that appears to form strong evidence for a derivational theory of phonology, can actually be re-interpreted as an 'identity' phenomenon, analysed in a constraint-based theory. I proposed a number of correspondence constraints that spell out requirements of identity holding between an output form on the one hand, and its base and its input on the other hand. We saw that evaluation of output forms must be parallel, in the sense that simultaneous reference is made to the base and the input. These correspondence constraints turned out to interact with constraints of syllable well-formedness, and metrical constraints that govern the shape and position of feet.

Section 3 will extend the analysis to a process of *a-syncope* in another Arabic dialect, that of Tripoli. This analysis will confirm two general points. First, we will find additional motivation for the notion of 'base' as I have defined it in Section 1, as a free form that is compositionally related to the output. Second, it will provide evidence for the metrical analysis and its interaction with correspondence constraints. Cases are predicted of epenthetic vowels that are stressed under duress — because of foot form constraints which dominate 'epenthetic unstressability' constraints.

3. Metrical Opacity, Syncope and Epenthesis in Tripoli Arabic

The stress-related phonology of the Tripoli dialect of Arabic has been documented in much detail by Kenstowicz and Abdul-Karim (1980), and the following discussion will incorporate many of their insights. I will change, however, from a rule-based perspective into one based on paradigm relationships. The resulting OT analysis will improve in a number of ways over Kenstowicz and Abdul-Karim's original analysis. Most notably, it will not use morphological structure diacritically to mark off cyclic domains. Instead the notion of 'base', as proposed earlier in Section 1, makes correct predictions about which morphologically related forms are in a correspondence relationship.

Three 'rules' of Tripoli Arabic will play an important role in the discussion. Below these are formulated according to Kenstowicz and Abdul-Karim (1980):

(63) a. a-SYNCOPE
  a [−stress] → Ø / ___ C + V 'non-cyclic suffix'
 b. i-SYNCOPE
  i [−stress] → Ø / ___ CV
 c. i-EPENTHESIS
  Ø → i / C ___ C {C, #}

Two of these rules we have already seen active in Palestinian Arabic. The difference between Palestinian and the Tripoli dialect is that the latter has a rule of a-Syncope which applies in a wide variety of contexts, whereas Palestinian restricts it to a small number of morphological contexts. This dialectal difference can be attributed to a small difference in the ranking of the constraint triggering a-Syncope. In the following sections I will discuss the highly interesting interactions between the three 'processes', and their interactions with (input) faithfulness and (base) identity.

3.1 *Underapplication of a-Syncope*

A rule of a-Syncope deletes /a/ in open unstressed nonfinal syllables (cf. 64a.ii, iii, 65a.ii). Kenstowicz and Abdul-Karim (1980: 59) observe that application of this rule is blocked before a 'cyclic' suffix, that is, in accusative (64b) and possessive (65b) forms:

(64)	a.i	/ḍarab/	ḍárab		'he hit'
	a.ii	/ḍara̲b-it/	ḍárb-et[12]		'she hit'
	a.iii	/ḍa̲rab-ti/	ḍráb-ti		'you f. sg. hit'
	b.i	/ḍarab-ik/	ḍár**a**b-ik	*ḍárb-ik	'he hit you f.'
	b.ii	/ḍarab-ni/	ḍ**a**ráb-ni	*ḍráb-ni	'he hit me'
(65)	a.i	/baʔar/	báʔar		'cattle'
	a.ii	/baʔa̲r-a/	báʔr-a		'a cow'
	b.i	/baʔar-i/	báʔ**a**r-i	*báʔr-i	'my cattle'
	b.ii	/baʔar-na/	b**a**ʔár-na	*bʔár-na	'our cattle'

The blocking of a-Syncope in accusatives and possessives is problematic from a derivational viewpoint. Note that if a-Syncope is a cyclic rule, then it should apply within the outer cycle, e.g. [[baʔar]-i], producing *[báʔri]. No appeal can be made to strict cyclicity, since the rule's context arises by morpheme concatenation, stress, and resyllabification.

In setting up an OT analysis, the first question is: what triggers a-Syncope, a process that does not occur in Palestinian Arabic? We readily identify the 'trigger' of deletion of /a/ as a constraint that is analogous to No [i]:

(66) No [a]
[a] is not allowed in open syllables.

For obvious reasons, this constraint must dominate 'faithfulness to the input' in the Tripoli dialect, while in Palestinian (which does not have a-Syncope) the ranking must be reverse:

(67)	a.	Tripoli:	No [i], No [a] » Max(I/O)
	b.	Palestinian:	No [i] » Max (I/O) » No [a]

Now we may ask why a-Syncope underapplies to accusatives and possessives in the Tripoli dialect. The generalisation is that these forms are the ones that have *bases*. Criteria for base-hood point to (64a.i) [ḍárab] as the base of the accusatives (64b.i, ii) [ḍár**a**b-ik] and [ḍ**a**ráb-ni], since this is a free form to which they are compositionally related. In contrast, the forms (64a.ii, iii) [ḍárb-et] and [ḍráb-ti] have no base, because their verbal stem, /ḍarab/ 'to hit', does not occur uninflected.

12. The ending [–et] is an allomorph of the 3sg. fem. suffix that appears word-finally. Interestingly, a-Syncope is blocked in the 3.pl. form, which is [ḍára̲b-u] rather than *[ḍárb-u]. This might point to the 3.sg. form [ḍárab] as the base of the 3.pl., to which it is compositionally related by morpheme composition, as well as by person and gender features, although not by number features.

Similar reasoning applies to the possessives [báʔ**ar**-i] and [b**a**ʔár-na] in (65b), both of which are compositionally related to a free form (65a.i) [báʔar] 'cattle'. In contrast, no base can be identified for the singular form [báʔr-a] 'a cow', because it cannot be compositionally related (due to a conflict in number inflection with that of the plural [báʔar] 'cattle').

Given the pairing of output forms with bases, we directly find an explanation for the 'underapplication' of a-Syncope to the first vowels in forms (64b.ii) [ḍ**a**ráb-ni] and (65b.ii) [b**a**ʔár-na]. These vowels have stressed correspondents in the base. Hence, deleting them would violate HEAD-MAX(B/O), which must therefore be undominated, as in Palestinian.

Interestingly, we also have an explanation for the 'underapplication' of a-Syncope to the second vowel in the accusative form (64b.i) [ḍár**a**b-ik] and the possessive form (65b.i) [báʔ**a**r-i]. In contrast to the initial vowels, these vowels lack stressed correspondents in their bases. We observe that /a/ is protected from deletion whenever it has a correspondent in the base, as in (68a), but that it deletes where no base occurs, as in (68b):

```
(68)  a.  /ḍ a r a b -i k/      b.  /ḍ a r a b -i t/      Input
           | | | | | |  | |            | | |   | | |
          [ḍ a r a b -i k]          [ḍ a r   b -e t]      Output
           | | | | |
          [ḍ á r a b]                                     Base
```

This points to a base-identity constraint requiring that base segments have correspondents in the output. This, of course, is MAX(B/O), repeated below from (29):

(69) MAX(B/O)
Every segment in the base has a correspondent in the output.

MAX(B/O) dominates NO [a], as motivated by the 'underapplication' effects that we saw in the examples above[13]. And as I have argued above, NO [a] itself dominates MAX(I/O), since we find normal application of a-Syncope at the expense of (input) faithfulness in forms such as [ḍár.bet] 'she hit'. In Section 2.1 I argued for Palestinian Arabic that MAX(B/O) is ranked below NO [i], while HEAD-

13. Interestingly, the relative ranking of MAX(B/O) and NO [a] is a source of dialectal variation. As Kenstowicz and Abdul-Karim (1980: 59) report, the Kfar-Ṣghāb dialect (another Levantine dialect, Fleisch 1963) has extended a-Syncope to apply 'before cyclic suffixes', that is, in accusative and possessive forms. This dialect has forms such as [ḍárb-ak] 'he hit you m.' and [sámak] 'fish, pl.', [sámk-u] 'his fish, pl.'. The application of a-Syncope in these forms follows from a demotion of MAX(B/O) below NO [a].

MAX(B/O) is ranked above it. (This ranking actually reoccurs in Tripoli, and I will give evidence for it below.) This subset of constraints is ranked as in (70):

(70) HEAD-MAX(B/O) » NO [i] » MAX(B/O) » NO [a] » MAX(I/O)

This partial ranking is consistent with the complete ranking of Palestinian except for the sub-ranking of NO [a] with respect to MAX(I/O). If we tentatively assume the remaining rankings of Palestinian for the Tripoli dialect, we arrive at:

(71) *CLASH, FOOT-FORM, *CLUSTER, **HEAD-MAX(B/O)** » NONFINALITY » **NO [i]**, PARSE-2, HEAD-DEP(O/I) » **MAX(B/O)**, WSP » **NO [a]**, PARSE-σ » **MAX(I/O)**

The tableaux in (72) and (73) illustrate the analysis. Constraints appearing in boldface in (71) re-appear as such below:

(72)

I: /ḍarab-ik/ B: [ḍá.rab]	HEAD-MAX (B/O)	NON-FINALITY	NO [i]	PARSE-2	HEAD-DEP (I/O)	MAX (B/O)	WSP	NO [a]	MAX (I/O)
a. [(ḍár).bik]						*!	*		*
b. ☞[(ḍá.ra).bik]							*	**	
c. [(ḍrá.bik)]	*!	*				*	*	*	*

(73)

I: /ḍarab-et/ B: *none*	HEAD-MAX (B/O)	NON-FINAL-ITY	NO [i]	PARSE-2	HEAD-DEP (I/O)	MAX (B/O)	WSP	NO [a]	MAX (I/O)
a. ☞[(ḍár).bet]							*		*
b. [(ḍá.ra).bet]							*	*!*	
c. [(ḍrá.bet)]		*!					*	*	*

Like Palestinian Arabic, the Tripoli dialect has epenthesis, and opaque stress as a result of it. Consider the following tableau of /madrs-e/ [má.d**i**r.se] 'school':

(74)

I: /madrs-e/ B: *none*	HEAD-MAX (B/O)	NON-FINAL-ITY	NO [i]	PARSE-2	HEAD-DEP (I/O)	MAX (B/O)	WSP	NO [a]	MAX (I/O)
a. ☞[(má.d**i**r).se]							*	*	
b. [ma.(d**í**r).se]					*!			*	
c. [(mád).r**i**.se]			*!	*					
d. [mad.(r**í**.se)]		*!	*		*		*		

Tripoli Arabic also shares with Palestinian a process of i-Syncope, patterning as discussed in Section 2. Again, we find the 'promiscuous' interaction between i-Syncope and i-Epenthesis that occurs in Palestinian examples such as /yi-ktib-u/, [yí.k**i**t.bu]. See (75):

(75)	a.	/ ṭifl-e/	ṭíf.le		'child, f.'
	b.	/ ṭifl-i̱t-i/	ṭí.f**i**l.ti	* ṭif.li.ti	'my child, f.'
	c.	/ ṭifl-it-na/	ṭif.lít.na		'our child, f.'

This supports the assumption, made immediately above, that the full constraint ranking of Palestinian, reached at the end of Section 2, holds for Tripoli as well. (As argued earlier, the single difference is the ranking of NO [a], see again (67)).

(76)

I: /ṭifl-it-i/ B: [ṭíf.le]	HEAD-MAX (B/O)	NON-FINAL-ITY	NO [i]	PARSE-2	HEAD-DEP (I/O)	MAX (B/O)	WSP	NO [a]	MAX (I/O)
a. ☞[(ṭí.f**i**l).ti]			**			*	*		*
b. [ṭi.(f**í**l).ti]			**		*!	*			*
c. [(ṭíf).li.ti]			**	*!					
d. [ṭif.(lí.ti)]		*!	**				*		
e. [(ṭf**í**l).ti]	*!		*		*	**			*

3.2 *a-Syncope and its Interactions with i-Syncope and i-Epenthesis*

The most interesting aspect of Tripoli is its *triple* interaction of a-Syncope, i-Syncope, and i-Epenthesis. This is exemplified in (77) by 3sg. fem. accusative forms and possessives. The relevant forms in which all three processes apply are

(77a.ii, b.ii). Syncopated vowels have been underlined in input forms while epenthetic vowels appear in boldface in the outputs[14]:

(77)		*Input*	*Output*	*Retain /a/*	*Retain /i/*	*Gloss*
	a.i	/ḍarab-it/	ḍár.bet			‘she hit’
	a.ii	/ḍarab-it-u/	ḍá.r**i**b.tu	*ḍa.ráb.tu	*ḍár.bi.tu	‘she hit him’
	a.iii	/ḍarab-it-na/	ḍar.bít.na			‘she hit us’
	b.i	/baʔar-a/	báʔ.ra			‘cow’
	b.ii	/baʔar-it-i/	bá.ʔ**i**r.ti	*ba.ʔár.ti	*báʔ.ri.ti	‘my cow’
	b.iii	/baʔar-it-na/	baʔ.rít.na			‘our cow’

Apparently it would have been more ‘economic’ to syncopate only one vowel, rather than to syncopate both /a/ and /i/, and substitute one by epenthesis. What causes this exchange of /a/ for [i]? In order to find out we must look at competitors that retain input /a/, as well as those that retain input /i/.

As should be clear by now, starred candidates that retain input /i/ fatally violate NO [i] or PARSE-2. These constraints outrank those that are violated in actual outputs, WSP in particular. But the nonoccurrence of the candidates that retain input /a/ is more surprising. Unlike the forms that retain /i/, these do not violate NO [i] nor PARSE-2. They have skeletal forms that are identical to those of the actual output forms, the single difference being that they retain the input vowel /a/, rather than deleting it and replacing it by an epenthetic [i]. The question is: why are these ‘faithful’ forms rejected at the expense of ‘unfaithful’ forms? Preserving the input vowel must be dispreferred for some reason. What can it be?

Observe that epenthetic [**i**] in [ḍá.r**i**b.tu] and [bá.ʔ**i**r.ti] has no correspondent in the bases of these forms, [ḍár.bet] and [báʔ.ra]. If an output vowel lacks a B*ase* correspondent, then it is ‘epenthetic’ from the viewpoint of any constraint evaluating B/O-correspondence, regardless of whether it possesses an *Input* correspondent. Therefore the alternative forms that retain input /a/, [ḍá.r**a**b.tu] and [bá.ʔ**a**r.ti], would also be treated as ‘epenthetic’. We already know from the previous discussion that the epenthetic vowel is [i], a result of the undominated ranking of IDENT-[a]. Conseqently an output [i] lacking a correspondent in the input is preferred over an output [a] lacking a correspondent in the input. If we only generalise correspondence relationships that are evaluated by these constraints from Output-Input to Output-Base, then we are able to explain the choice of (78a) over (78b):

14. The question of what distinguishes *opaque* stress in [ḍár**i**b-t-u] and *non-opaque* stress in [baa.r**í**k.tu] and [ʕal.l**í**m.tu] will be addressed in Section 4 below.

(78)	a.		b.		
		/ḍ a r a b -i t-u/		/ḍ a r a b -i t-u/	Input
		[ḍ a r **i** b t-u]		[ḍ a r **a** b t-u]	Output
		[ḍ a r b-e t]		[ḍ a r b-e t]	Base

Form (78b) fatally violates IDENT-[a](B/O), the constraint requiring that output [a] must have a base correspondent. In contrast, form (78a) only violates the lower-ranked IDENT-[i](I/O). This completes the answer to the original question why 'less faithful' outputs are preferred to 'faithful' outputs.

The single remaining question is why [ḍá.r**i**b.tu] and [bá.ʔ**i**r.ti] have opaque stress on their initial syllables, even though their second syllables would be qualified bearers of stress from the viewpoint of HEAD-DEP(I/O). Because [i] in (78a) has a correspondent in the input, no violation of HEAD-DEP(I/O) would arise by stressing it, as in [ḍa.r**í**b.tu]. However, as we observed immediately above, this output [i] has no correspondent in the base, and therefore it is treated as epenthetic by any constraint evaluating B/O-correspondence.

The answer therefore must be that opaque stress is due to non-correspondence of the output prosodic head with a vowel in the *base*, as a comparison of (79a, b) shows:

(79)	a.		b.		
		/ḍ a r a b -i t-u/		/ḍ a r a b -i t-u/	Input
		[ḍ <u>á</u> r **i** b t-u]		[ḍ a r <u>**í**</u> b t-u]	Output
		[ḍ <u>á</u> r b-e t]		[ḍ á r b-e t]	Base

That is, a stressed output vowel must have a *base* correspondent. This is stated in:

(80) HEAD-DEP(B/O)
Every vowel in the output prosodic head has a correspondent in the base.

If we rank this constraint in the same position as HEAD-DEP(I/O), we arrive at an analysis of opaque stress in [ḍár**i**btu] that is illustrated in tableau (81). In the column HEAD-DEP(IB/O), I will mark violations by indicating whether a stressed output vowel lacks a correspondent in the base by 'B'.

(81)

I: /ḍarab-it-u/ B: [ḍár.bet]	IDENT [a] (B/O)	HEAD-MAX (B/O)	NON-FINAL-ITY	NO [i]	PARSE-2	HEAD-DEP (IB/O)	MAX (B/O)	WSP	NO [a]
a. ☞[(ḍá.r**i**b).tu]							*	*	*
b. [ḍa.(r**í**b).tu]						B!	*		*
c. [(ḍár).bi.tu]				*!	*				
d. [ḍar.(bí.tu)]			*!	*				*	
e. [ḍa.(ráb).tu]	*!					B	*		*
f. [ḍa.(rá.bi).tu]	*!			*		B			**

For comparison and to wind up this section, I give the tableau for [ḍar.bít.na] in (82):

(82)

I: /ḍarab-it na/ B: [ḍár.bet]	IDENT [a] (B/O)	HEAD-MAX (B/O)	NON-FINAL-ITY	NO [i]	PARSE-2	HEAD-DEP (IB/O)	MAX (B/O)	WSP	NO [a]
a. ☞[ḍar.(bít).na]								*	*
b. [ḍár).bit.na]					*!			*	*
c. [(ḍà.ra).(bít).na]	*!								***

The final section of this article will demonstrate that this analysis correctly predicts a property that caused great troubles to earlier derivational analyses: epenthetic vowels may be stressed 'under duress', due to foot well-formedness.

4. Stressed Epenthetic Vowels in Tripoli Arabic

Consider the following partial verbal paradigm of the CvCvC stem /ḍarab/ 'to hit', and compare it to that of a CvvCvC stem /baarak/ 'to bless' and a CvCCvC stem /ʕallam/ 'to teach':

(83)	a.i	/ḍar<u>a</u>b-it/	ḍár.bet	'she hit'
	a.ii	/ḍar<u>a</u>b-<u>i</u>t-u/	ḍá.r**i**b.tu	'she hit him'
	a.iii	/ḍar<u>a</u>b-it-na/	ḍar.bít.na	'she hit us'

b.i	/baarak-it/	báar.ket	'she blessed'
b.ii	/baarak-it-u/	baa.r**í**k.tu	'she blessed him'
b.iii	/baarak-it-na/	baar.kít.na	'she blessed us'
c.i	/ʕallam-it/	ʕáll.met	'she taught'
c.ii	/ʕallam-it-u/	ʕal.l**í**m.tu	'she taught him'
c.iii	/ʕallam-it-na/	ʕall.mít.na	'she taught us'

When we compare the accusative forms (83a.ii) to (83b.ii, c.ii) we observe that the position of the epenthetic vowels is identical, but that there is a difference in the position of stress. In the former this vowel is unstressed, whereas in the latter the epenthetic vowels are stressed. Kenstowicz and Abdul-Karim (1980) observe that where an epenthetic vowel is stressed, the syllable that precedes it is *heavy*.

A similar relationship holds between the possessives based on CvCvC nouns (/baʔar/ 'cow', (84a)) and those based on shapes CvCCvC (/maktab/ 'library', (84b)).

(84)	a.i	/baʔar-a/	báʔ.ra	'cow'
	a.ii	/baʔar-it-i/	bá.ʔ**ir**.ti	'my cow'
	a.iii	/baʔar-it-na/	baʔ.rít.na	'our cow'
	b.i	/maktab-e/	má.k**i**t.be	'library'
	b.ii	/maktab-it-i/	mak.t**í**b.ti	'my library'
	b.iii	/maktab-it-na/	ma.k**i**t.bít.na	'our library'

Finally this relationship holds between possessives of CvCC nouns (/ṭifl/ 'child', (85a) and those of CvCCC nouns (/madrs/ 'school', (85b), and CvvCC nouns (/taawl/ 'table', (85c)):

(85)	a.i	/ṭifl-e/	ṭíf.le	'child, f.'
	a.ii	/ṭifl-it-i/	ṭí.f**il**.ti	'my child, f.'
	a.iii	/ṭifl-it-na/	ṭif.lít.na	'our child, f.
	b.i	/madrs-e/	má.d**ir**.se	'school'
	b.ii	/madrs-it-i/	mad.r**í**s.ti	'my school'
	b.iii	/madrs-it-na/	ma.d**ir**.sít.na	'our school'
	c.i	/taawl-e/	táaw.le	'table'
	c.ii	/taawl-it-i/	taa.w**í**l.ti	'my table'
	c.iii	/taawl-it-na/	taaw.lít.na	'our table'

The question is, why do we find *opaque* stress in [ḍá.r**i**b.tu], but *transparent* stress in [baa.r**í**k.tu]? If epenthetic vowels consistently reject stress, then we would expect stress to fall on the initial syllable regardless of its weight:

(86)	a.	*[(báa).r**i**k.tu]	>	[baa.(r**í**k).tu]
	b.	*[(mák).t**i**b.ti]	>	[mak.(t**í**b).ti]

The generalisation that a heavy syllable must precede in order for an epenthetic vowel to be stressed points to an interaction between metrical constraints and the constraints that require that epenthetic vowels must be unstressed, that is, HEAD-DEP(IB/O).

Only a small modification of the analysis is required to produce this interaction. Both of the metrical constraints FOOT-FORM and PARSE-2 must come to dominate HEAD-DEP(I/O). The constraint hierarchy can now be stated in its final form in (87):

(87) *CLASH, IDENT[a](B/O), FOOT-FORM, *CLUSTER, HEAD-MAX(B/O) » NONFINALITY » NO [i], PARSE-2 » HEAD-DEP(IB/O) » MAX(B/O), WSP » NO [a], PARSE-σ » MAX(I/O)

To illustrate this analysis I present the following tableaux:

(88)

I: /baarak-it-u/ B: [báar.ket]	IDENT [a] (B/O)	FOOT-FORM	NON-FINAL-ITY	NO [i]	PARSE-2	HEAD-DEP (IB/O)	MAX (B/O)	WSP	NO [a]
a. ☞[baa.(r**í**k).tu]						*	*	*	
b. [(báa).r**i**k.tu]					*!		*	*	
c. [baa.(rá.ki).tu]	*!			*		*		*	*
d. [(báar).ki.tu]				*!	*				
e. [baar.(kí.tu)]			*!	*				*	
f. [(báa.r**i**k).tu]		*!					*	*	
g. [baa.(rák).tu]	*!					*	*	*	

(89)

I: /madrs-it-i/ B: [má.d**i**r.se]	IDENT [a] (B/O)	FOOT-FORM	NON-FINAL-ITY	NO [i]	PARSE-2	HEAD-DEP (IB/O)	MAX (B/O)	WSP	NO [a]
a. ☞[mad.(r**í**s).ti]				*		*	**	*	
b. [(mád).r**i**s.ti]				*	*!		**	*	
c. [mad.(r**í**.si).ti]				**!*		*	*	*	
d. [ma.(d**í**r).si.ti]				**!	*	*			*
e. [(mád.r**i**s).ti]		*!		*			**	*	

(90)

I: /maktab-it-i/ B: [má.k**i**t.be]	IDENT [a] (B/O)	FOOT-FORM	NON-FINAL-ITY	NO [i]	PARSE-2	HEAD-DEP (IB/O)	MAX (B/O)	WSP	NO [a]
a. ☞[mak.(t**í**b).ti]				*		*	**	*	
b. [(mák).t**i**b.ti]				*	*!		**	*	
c. [ma.(k**í**t).bi.ti]				**!	*	*			*
d. [(mák.t**i**b).ti]		*!		*			**	*	
e. [mak.(táb).ti]	*!			*		*	**	*	
f. [mak.(tá.bi).ti]	*!			**		*	*	*	*

4.1 *Alternative Accounts*

Kenstowicz and Abdul-Karim (1980) base their analysis on the observation that epenthetic [i] in Tripoli is stressed *iff* it lands in a position in which a vowel is stressed in the (Ramallah) Palestinian dialect:

(91) a. Tripoli: [ḍá.r**i**b.tu] [baa.r**í**k.tu] [ʔal.l**í**m.tu]
b. Palestinian: [ḍár.ba.to] [baa.rá.ka.to] [ʔal.lá.ma.to]

The idea is that the basic metrical structures of the dialects are identical, but that dialects have different syncope rules. Metrification is due to a left-to-right assignment of a quantitative foot, which may be seen as an ancestor of the moraic trochee. Constituency is preserved through cyclic stress and syncope, 'unifying' with epenthetic [i] post-cyclically. An epenthetic vowel surfaces as stressed if it 'lands' into a metrical position that was strong in the original parsing. This requires an intermediate step of 'floating feet' for Tripoli:

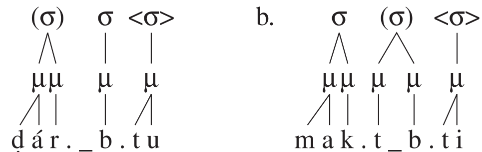
(92) a. [(**LL**)(LL)] [(ḍá.r_).(b .tu)]
b. [(H)(**LL**) L] [(baa).(r_.k).tu] [(ʔal).(l_.m).tu]

Similar analyses of related Levantine dialects are Irshied and Kenstowicz (1984), Al-Mozainy, Bley-Vroman and McCarthy (1985), and Hayes (1995a).

The central conceptual problem to this type of analysis is that the foot parsings that 'explain' the stress values of epenthetic vowels never surface in Tripoli, because of excessive syncope in long strings of light syllables. Therefore the position of epenthetic vowels within feet may be inferred *only* from the stress value of these vowels, producing circularity. (As Kenstowicz and Abdul-Karim notice: "[…] suffixed 3 sg.f. perfects and feminine nouns in construct are the only places where the morphology of Levantine Arabic permits words of four light syllables to be constructed.")

A second type of analysis is based on 'degenerate' syllables rather than floating feet (Broselow 1992). Epenthesis is the assignment of a syllable nucleus (i.e. minimally a mora) in a string of consonants that cannot be syllabified (Itô 1989). Consonants syllabify as moraic in degenerate syllables, but there is pressure on degenerate syllables to be bimoraic, due to a *Bimoraicity Constraint* (a slightly different analysis is presented by Piggott 1995):

(93) a. b.



[ḍár.ib.tu] [mak.tíb.ti]

The stress contrast results from the assumption that the /r/ in (93) syllabifies 'backward', so as to contribute to the bimoraicity of the preceding syllable. In contrast, the first /t/ in (93b) is syllabified as an onset of the degenerate syllable [t_b], since the preceding syllable [mak] is already maximal as it is. The result is a monomoraic epenthetic syllable in (93a), and a bimoraic one in (93b).

I identify three problems for this analysis. First, the bimoraic syllabification of /tb/ in (93b) requires a *moraic onset*. Even as an intermediate step in a derivation, this is a highly questionable assumption. Second, this analysis fails to predict non-opaque stress in Tripoli forms in which a long vowel is involved, rather than a triconsonantal cluster. The underlying representation /baarak-it-u/, (83bii), might syllabify as [báar._k.tu] since CvvC syllables are allowed in word-internal positions, cf. [báar.ket]. Third, why should epenthetic vowels be inserted in *pre-consonantal* position, e.g. [ḍár**i**btu] > *[ḍárb**i**tu], in violation of the constraint ONSET? This cannot follow from the Bimoraicity Constraint. (In my analysis, which does not make the assumption that /r/ syllabifies 'backward', this follows from NO [i].)

5. Conclusions

The OT analysis of Levantine Arabic stress and vowel-zero alternations presented in this paper has lead to the following conclusions. Although metrical opacity apparently gives severe problems to Optimality Theory, there is in fact an OT counterpart to the derivational mechanism of the cycle: Base/Output-correspondence. This requires no abstract intermediate levels of representation in accounting for opaque stress as in rule-based analyses. I proposed a definition of '*base*' as a compositionally related, free form. Finally, I have argued that the evaluation of output forms is *parallel* in B/O and I/O correspondence.

Acknowledgments

This paper has benefitted from discussions with John Alderete, Laura Benua, Jan Don, Anthony Dubach Green, Bruce Hayes, Ben Hermans, John McCarthy, and Wim Zonneveld. Full responsibility for any mistakes is mine, of course. The research for this paper was sponsored by the Royal Netherlands Academy of Sciences (KNAW).

Sign-Based Morphology

A declarative theory of phonology-morphology interleaving

Cemil Orhan Orgun

1. Introduction

This paper investigates the implications of cyclic phonological effects for nonderivational theories of phonology, and proposes a declarative theory of the phonology-morphology interface called Sign-Based Morphology that accounts for such effects. The term 'cyclicity' refers to the state of affairs in which a subpart of a linguistic form may be subject to phonological constraints on its own, in addition to constraints enforced on the whole form. There may be a number of such embedded phonological domains in morphologically complex forms. The number and location of such domains is determined by the morphological structure. There are two common ways in which cyclicity has been implemented in theories of phonology. These are summarized in (1) and (2).

(1) Phonology applies to fully formed morphological structures. The most deeply embedded constituent undergoes phonology first, phonology then applies to successively larger constituents: Chomsky and Halle (1968), Halle and Kenstowicz (1991), Halle and Vergnaud (1987), Odden (1993).

(2) Inputs to some morphological constructions may be subject to phonology on their own (interleaving): Lexical Phonology (Pesetsky 1979; Kiparsky 1982a, 1983, 1985; Mohanan 1986), Prosodic Lexical Phonology (Inkelas 1990).

Theories such as Lexical Phonology have made a distinction between two kinds of phonology-morphology interleaving: cyclicity and level ordering. In this paper, the term ‘interleaving’ covers what is traditionally called cyclicity as well as what is traditionally called level ordering. The difference between cyclicity and level ordering will not be relevant.[1]

2. Cyclicity in the Age of Nonderivationalism

In this section I discuss the implications of the rise of nonderivational phonology for the analysis of apparently cyclic phenomena.

2.1 *Background: Nonderivational Approaches to Phonology*

Recent years have witnessed a growing movement towards nonderivational approaches to linguistics. The first strictly declarative approach to phonology was proposed by Johnson (1972), who launched the small but healthy field of declarative phonology based on finite state transducers, represented by such works as Koskenniemi (1983), Kaplan and Kay (1981). More recently, monostratal constraint-based approaches to phonology have been proposed by such researchers as Bird (1990) and Scobbie (1991). However, it is only during the past few years that constraint-based approaches such as Harmonic Phonology (Goldsmith 1989), the Theory of Constraints and Repair Strategies (Paradis 1988) and Optimality Theory (Prince and Smolensky 1993) have come to dominate the field.

Constraint-based theories of phonology have largely focused on the phenomena once believed to motivate rule ordering, demonstrating that work earlier attributed to derivation can be handled by nonderivational theories as well. However, the status of cyclic phonology remains somewhat less clear. The present paper is devoted to exploring this issue.

2.2 *Cyclicity as a Problem for Nonderivational Phonology*

Many proponents of nonderivational phonology summarily equate cyclic phonology with rule ordering. For example, Goldsmith summarizes the operation of the phonology-morphology interface in Lexical Phonology: “First add an affix, then send that material through a set of rules which modifies the resultant form; then go to the next level, add another affix, and finally string all the words together, only after which do we reach a point where the postlexical rules get a chance to

1. Detailed discussion of the status of level ordering can be found in Orgun (1996)

apply" (1993: 21). The presupposition here is clear: cyclicity is intrinsically derivational. Under this view, development of a true nonderivational approach to phonology requires finding alternatives to cyclicity. The sample slogans below illustrate this and some other common reactions to interleaving.

(3) "Interleaving is not necessary" (Lakoff 1993; Karttunen 1993; Kennedy 1994; Myers 1994; Zec 1994; Cole 1990; Cole and Coleman 1992)

(4) "Interleaving is a device every analysis should try to do without" (Lakoff 1993; Karttunen 1993; McCarthy and Prince 1993b; Kennedy 1994; Myers 1994; Zec 1994; Benua 1995; McCarthy 1995; Kenstowicz 1996)

(5) "Interleaving is cognitively implausible" (Goldsmith 1993a; Lakoff 1993)

(6) "Interleaving is computationally intractable" (Sproat 1992)

(7) "Interleaving is inherently derivational" (Bird 1990; Scobbie 1991; Lakoff 1993; Karttunen 1993; McCarthy and Prince 1993; Kennedy 1994, Myers 1994; Zec 1994; Benua 1995; McCarthy 1995; Kenstowicz 1996; etc.)

2.3 *Doing without Cyclicity*

The common stance among nonderivational phonologists is, as we have seen, that interleaving is inconsistent in spirit (as well, presumably, as in letter) with a strictly declarative approach to phonology. Some researchers holding this view have optimistically assumed that the demonstrated alternatives to rule ordering will, with little or no modification, account for cyclic effects as well. To take one example, Karttunen (1993) writes in a paper defending a nonderivational approach to phonology that "We have not mentioned the issue of *cyclic* ordering … Let us simply state that, in our opinion, the arguments for cyclic ordering are weaker then the ones [for rule ordering within a cycle] we discuss" (1993: 194). This optimism has proved unwarranted, however, as other, more empirically oriented work has recognized that many cyclic effects prove resistant to reanalysis using the tools developed to replace rule ordering. A number of different approaches have been taken to the challenge posed by cyclic effects. I survey two general classes of approaches in sections 2.3.1 and 2.3.2.

2.3.1 *Cyclic Effects are Illusory and Reducible to Other Things*
Some researchers have claimed that cyclic phonological effects are only epiphenomenal and looked for other mechanisms to derive such effects.

One approach is to multiply the number of levels to which declarative constraints apply. Both Goldsmith (1993a) and Lakoff (1993) propose that constraints relate three parallel phonological representations, the (implicit) expectation being that cyclic effects never require more than two cycles of phonology (a claim made independently by Cole 1991). Another, totally different, line of attack is taken by Prince and Smolensky (1993) and McCarthy and Prince (1994). In this Optimality-theoretic approach, cyclic effects are claimed to reduce to the result of constraints requiring alignment between edges of morphologically defined domains and edges of metrical elements. These two approaches essentially claim that true cyclic effects, in which there are potentially as many applications of phonology in a complex word as there are morphological subconstituents in the word, are epiphenomenal. What appears to be the application of phonology to a subword constituent is really something else entirely.

What all these approaches share is the view that phonology applies only once (twice in Lakoff's and Goldsmith's theories) to the underlying representations supplied by the morphemes in a linguistic form.

2.3.2 *Cyclic Effects are Real but Paradigmatic*
Another type of approach, represented by Becker 1993; Bochner 1993; Burzio 1994; Steriade 1994b; Benua 1995; Buckley 1995b; Flemming 1995; Kenstowicz 1996; and McCarthy 1996b, admits the validity of evidence for cyclic phonology but proposes an alternative interpretation in which cyclic effects are claimed to follow from correspondence constraints holding between paradigmatically related lexical items. The lexical items in question are not morphologically derived from one another. The correspondence constraints simply relate existing lexical items.

2.4 *Cyclicity as Nonderivational*

All the approaches described so far have taken it for granted that cyclic phonology, like rule ordering, is derivational and that this is sufficient reason to look for alternatives to cyclicity. The approaches vary only in the nature of the alternative they propose.

In this paper, I reject the presupposition underlying these approaches, contending that there is an important distinction between rule ordering and phonology-morphology interleaving. I argue that, contrary to popular belief, there is nothing inherently derivational about the latter.

Similar arguments can be found in two strains of past work. Cole and Coleman (1992) show that interleaving effects can be captured in a monostratal approach to phonology by enforcing declarative constraints on morphologically defined subparts of a phonological string (the approach is similar in this regard to the one taken by Buckley 1995b). Following a different line of attack, Orgun (1994b, c, 1995b, c) and Riehemann (1994) propose to integrate two-level approaches to phonology such as Generalized Correspondence Theory (McCarthy and Prince 1994, 1995) with declarative approaches to phrase structure such as HPSG (Pollard and Sag 1994). Such declarative approaches derive interleaving effects as a direct consequence of their basic architecture. In this paper I build on the approaches of Orgun and Riehemann, developing a full-fledged nonderivational theory of the phonology-morphology interface called Sign-Based Morphology.

Sign-Based Morphology, the theory proposed in this paper, falls into the category in (2). It is a theory of phonology-morphology interleaving. However, it differs from past interleaving approaches in several important ways. First, it is declarative; second, it derives interleaving effects from constituent structure configurations, rather than stipulating them in derivational terms, as past approaches have done. The aim of this paper is to show that Sign-Based Morphology, in addition to being nonderivational, provides an empirically and theoretically superior account of the phonology-morphology interface.

The paper begins with Section 3 presenting a number of examples showing the need for interleaving. Then, in Section 4, I present Sign-Based Morphology, a declarative approach to morphology that builds on the findings of Riehemann (1994) and Orgun (1994b, c, 1995b, c). In Section 5 I compare Sign-Based Morphology with current paradigmatic approaches to the phonology-morphology interface, showing that Sign-Based Morphology is superior on both empirical and theoretical grounds. I conclude that whether or not there is a derivational residue in phonology is entirely a question for phonological theory proper. Phonology-morphology interleaving is not a source of derivationalism.

Due to lack of space, I do not take up the cognitive and computational implications of Sign-Based Morphology in this paper. See Riehemann (1994), Koenig and Jurafsky (1994), and Orgun, Koenig and Jurafsky (1996) for discussion.

3. Empirical Necessity for Interleaving: Schizophrenia and Sensitivity to Branching

The purpose of this section is to show that constraints on intermediate morphological constituents are relevant to the ultimate surface outcome — i.e. that interleaving is necessary. This conclusion is entirely independent of the particular theory of phonology one assumes. It is also independent of one's particular approach to the phonology-morphology interface.

I discuss two types of phenomena that point to the need for interleaving. Section 3.1 presents a case of schizophrenia, the term applied by Orgun (1994a) to cases in which a segment is subject to the constraints of a syllabic position that differs from its surface one, but matches the syllable position of the corresponding segment in a morphologically related form (see Hall 1994 for examples). Section 3.2 discusses phonological sensitivity to the direction of branching in the morphological constituent structure.

3.1 *Schizophrenia*

Schizophrenic segments include consonants that are subject to the effects of a syllabic position that they do not surface in. Schizophrenia requires interleaving because the conditioning for the alternation in question is not present in the surface form. The constraint (or rule) must have applied in a morphologically related form, which the surface form is derived from (or related to).

Schizophrenia arises in Uighur (Orgun 1994b) through the interaction of vowel raising and elision. The first alternation, raising, applies to vowels in stem-final open syllables when followed by a suffix (i.e., it does not apply word-finally)

(8)	*Plain noun (no raising)*		*Suffixed noun (raising applies)*	
	qazan	‘pot’	qazɨn-i	‘pot-POSSESSIVE’
	bala	‘child’	balɨ-si	‘child-POSSESSIVE’
	ana	‘mother’	anɨ-lar	‘mother-PLURAL’

The second alternation of interest is elision of high unrounded vowels between identical consonants when permitted by syllable structure (i.e., in two-sided open syllable environments). The underlined vowel in the input deletes in the output:

(9)	balɨ̲lar + i	→	ballɨri	‘child-PL-POSSESSIVE’
	balɨ̲lar + i + ni	→	ballɨrɨni	‘child-PL-POSSESSIVE-ACC’

Schizophrenia arises through the interaction of raising with elision. Through elision, a vowel that has undergone open syllable raising may end up in a surface closed syllable. The schizophrenic vowel in (10) is underlined:.

(10) qazan + i + ni → qazɨ̲nni ‘pot-POSSESSIVE-ACC’

Why is the underlined vowel high? Note that the corresponding vowel in the form *qaza̲nni* ‘pot-ACC′ is low.

Interleaving provides a simple answer to this question: N+POSS+CASE is morphologically derived from (or related to) N+POSS. In particular, *qazɨnni* (10) is derived from (or related to) *qazɨni* (8). The reason *qazɨnni* has a high vowel is that the input to this particular morphological form itself has a high vowel, as illustrated below:

(11)			qazan	‘pot’
	qazan + i	→	qazɨni	‘pot-POSS’
	qazɨni + ni	→	qazɨnni	‘pot-POSS-ACC’

Many more examples of schizophrenia can be found in Hargus (1993), Hall (1994) and Orgun (1994a), where interleaving accounts are also offered.

3.2 *Sensitivity to the Direction of Branching: Slave Continuant Voicing Alternations*

The second type of evidence for interleaving is presented by phonological alternations which apply differently to left- and right-branching morphological structures. The example I present here involves continuant voicing alternations in Slave. The data are from Rice 1988, 1989.

Example (12) contrasts possessed and nonpossessed forms of nouns. The alternation of interest is that noun-initial continuants, voiceless in the nonpossessed forms, are voiced in the possessed forms. (Other alternations, with which we are not concerned here, also occur in these data, namely: coda consonants other than [n] neutralize to [h]; coda [n] deletes with concomitant nasalization of the preceding vowel).

(12) *Possessed nouns with voiced initial continuants (Rice 1988: 376)*

Nonpossessed	Possessed	
ʃį	se-ʒin-é	‘my song’
łuh	dezonah luz-é	‘(child’s) spoon’

The sensitivity of these voicing alternations is best illustrated by the behavior of three-morpheme structures.

Example (13) illustrates possessed and nonpossessed forms of compound and noncompound nouns. The contrast of interest occurs in the possessed forms: the initial continuant of a noncompound noun is voiced when the noun is possessed, but the initial continuant of a compound noun is voiceless in the same environment. The failure of voicing to apply in possessed compounds is not due to an idiosyncratic property of the noun stems in question: observe, in each of (13i, ii), that the initial continuant of the same stem in isolation will undergo voicing when possessed. The alternating continuants are shown in boldface:

(13) *Failure of voicing to apply to possessed compound nouns (Rice 1989: 34, 189, 190)*

	Nonpossessed	Possessed	Gloss
i.	**s**a-dzeé	se-**s**a-dzeé	'(my) watch, clock'
cf.	**s**a	se-**z**a-á	'(my) sun, month'
ii.	**s**ah-ðeh	se-**s**ah-ðéh-é	'(my) bear skin'
cf.	**s**ah	se-**z**ah-é	'(my) bear'

Moreover, it is not a general fact that the second morpheme in a three-morpheme word will fail to undergo initial continuant voicing. In (left-branching) three-member compounds, the (boldfaced) initial continuants of both the second and third stems undergo voicing:

(14) *Voicing applies in left-branching compounds (Rice 1989: 186, 187)*

da	'face'	**x**á	'hair'	bee	'knife'	da-**ɣ**á-bee	'razor'
deʃį	'wood'	tée	'mat'	ʍé	'skin'	deʃį -té-**w**é	'rug'

The behavior of the complex words in (13) and (14) makes sense once morphological constituent structure is taken into account. Voicing applies to the second and third stems in a left-branching compound, but only to the second stem in a right-branching compound.

Rice (1988) exploits this generalization by presenting a cyclic analysis of these data, which I will closely follow here. However, I will depart from Rice's theory-internal choices regarding underspecification, and, in so doing, better capture the relationship between the cyclic voicing alternations and the treatment of underlying specification of [voice].

The essence of Rice's analysis is that, on each cycle, domain-initial continuants are voiced, while domain-internal ones are devoiced (one can assume either that there is no root cycle, or, like Rice, that the voicing alternations are not active on the root cycle). Crucially, voicing alternations are structure-filling in Rice's analysis, allowing any input voicing specifications to be kept. Voicing alternations apply only to those continuants that are unspecified for voicing in the input. The derivations in (15) illustrate the analysis. Upper case letters indicate segments that are unspecified for voicing:

(15)		[Se [**S**a dzeé]]	[Se **S**aá]	[[da **X**á] bee]	[[deʃin té] **W**é]
	Cycle 1	**s**adzeé	se**z**aá	da**ɣ**á	deʃį té
	Cycle 2	se**s**adzeé	_	da**ɣ**ábee	deʃį té**w**é

The crucial assumption that voicing alternations are structure-filling makes an important prediction: underlying voicing specifications should be respected as well. That is, underlyingly voiceless initial continuants (if any) should always surface as voiceless, regardless of morphological structure; likewise, any underlyingly voiced initial continuants should always surface as voiced. Both predictions turn out to be correct, as shown below:

(16) *Support: Underlying voicing specifications are also respected*

Consistently voiced		Consistently voiceless		
Nonpossessed	Gloss	Nonpossessed	Possessed	Gloss
jah	'snow'	sǫ́	se-sóné	'(my) excrement'
ledʒai	'window'	łe	se-łé	'(my) lard'
zǫ	'only'	sámbaa	se-sámbaa	'(my) money'

The assumption that voicing is structure-filling allows us to capture a generalization over underlying forms and morphologically complex ones. In both cases, input voicing specifications are respected by further morphology. A cyclic, structure-filling account nicely captures the intrinsic relationship between the sensitivity to the direction of branchingness and the preservation of underlying voicing specifications. This insight is not available in a noncyclic account.

Having demonstrated the need for interleaving, we are now ready to develop a theory of phonology-morphology interaction that incorporates interleaving.

4. Sign-Based Morphology

In this section I develop a theory of the phonology-morphology interface, called Sign-Based Morphology, which draws upon two different lines of work. The first is the structural approach to interleaving proposed (in slightly different forms) by Sproat (1985), Cohn (1989) and Inkelas (1990, 1993a). The second is the unification-based approach to grammar, especially the line of work represented by Kay (1983), Gazdar et al. (1985), Pollard and Sag (1994), Fillmore and Kay (in progrress), Koenig and Jurafsky (1994). Interleaving follows as an inevitable consequence of constituent structure in the theory I develop, thus overcoming past objections to interleaving as an extraneous device which phonological theory should try to eliminate.

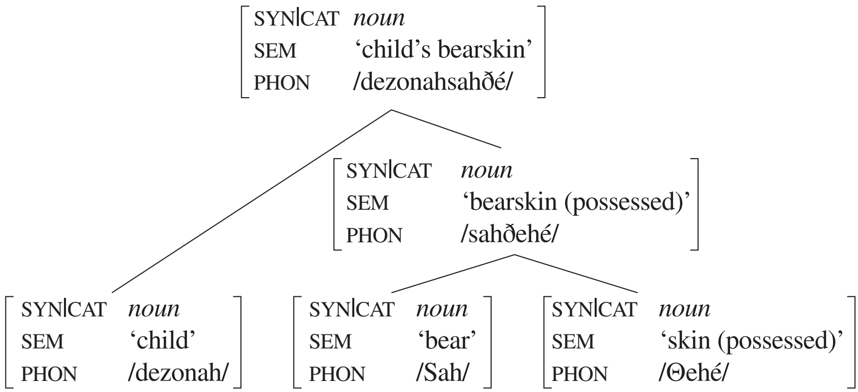
I begin with a brief discussion of the difference between ‘sign-based’ and ‘terminal-based’ approaches to linguistics. This is a critical contrast to draw, as past criticisms of interleaving as ‘extraneous’ crucially, if implicitly, assume a terminal-based approach to grammar. In this section, I will demonstrate first that there are no truly terminal-based approaches to linguistics. I will conclude that interleaving effects can be viewed as a consequence of using constituent structures. Objections to interleaving only make sense if constituent structures are to be dispensed with as well.

In the terminal-based approach, which underlies work in the structuralist tradition, terminal nodes are the only information-bearing elements in a constituent structure. The sole role of nonterminal nodes is to organize the terminal nodes into groups. The meaning of a form is assembled from the semantic information in the terminal nodes, while the phonology is determined by some phonological system operating on the strings supplied by the terminal nodes, which are the underlying representations of the morphemes that occupy those nodes.

Sign-based theories of linguistics differ from terminal-based ones in assuming that every node in a constituent structure, including nonterminal nodes, is an information-bearing element. That is, nonterminal nodes as well as terminal ones carry syntactic, semantic, and phonological information. The following discussion of sign-based linguistics highlights what is important for the purposes of this paper (for a general introduction, see Shieber 1986).

A ‘sign’ is defined as a Saussurean pairing between some phonological shape and some semantic information. In sign-based theories, a constituent structure is a statement of how the grammar justifies (licenses) the sign represented by the top node. Example (17) shows a sign-based representation of the Slave form *dezonahsahðehé* ‘child’s spoon’.

(17) *'Sign-based' approach*



The proper interpretation of the sign-based constituent structure in (17) is the following: The sign [dezonahsahðehé] is a possible pairing of form and meaning for the following reasons: (i) the signs [dezonah] and [sahðehé] exist, and (ii) the grammar allows, given a nominal sign and a possessed nominal sign, for there to be another nominal sign that combines the forms of the possessed nominal and noun in some appropriate way (to be dealt with by phonological theory) and inherits appropriate semantic information from the possessed noun and possessor noun. Constituent structures thus have a dual interpretation: they can be seen as representing the internal part-whole structure of a sign (the syntagmatic interpretation), or as a statement of what in the lexicon and grammar makes it possible to have the sign represented by the top node (the paradigmatic interpretation).

In a sign-based approach, the features of a mother node must be related to the features of its immediate constituents. This relation can be controlled by a set of constraints. Since these constraints apply to the phonological string of each node in a given constituent structure, complex constituent structures automatically give rise to interleaving effects. For example, the failure of continuant voicing to apply to the *sah* portion of the top node in (17) is due to the fact that the intermediate constituent *sahðehé* is subject to phonological constraints which require its initial continuant to be voiceless.

Most work in linguistics implicitly assumes a terminal-based approach; theories which are explicitly sign-based are a distinct minority. However, this contrast is in fact illusory. I am aware of no linguistic theory since structuralism which attributes no information to nonterminal nodes. *All* current constituent-based approaches to linguistics use some kind of feature percolation, thereby locating at least some information on the nonterminal nodes. The fact that nonterminal nodes bear category features is enough to illustrate this point. The need for assigning featural information to nonterminal nodes in a constituent structure was recognized even within the structuralist tradition by Hockett (1954), who observed that a pure item-and-arrangement view (a pure terminal-based approach in the terminology I use here) is therefore untenable.

Since most existing theories of linguistics use constituent structures, and all that do assume at least some feature percolation, it is fair to say that the only device needed to derive cyclicity comes for free in all existing theories of linguistics, even if not explicitly sign-based.

Thus, far from being an extra 'tool' that adds complexity to a theory (Zec 1994; Kennedy 1994), interleaving is a direct consequence of using constituent structures. Anyone who is committed to avoiding interleaving must do without constituent structures.[2, 3] Criticisms of interleaving as 'derivational' or 'formally extraneous' (see (3)–(7)) appear to be based on a lack of understanding of the formal properties of a constituent structure.

5. Comparison of Sign-Based Morphology with Paradigmatic Approaches

Having shown that interleaving is (a) necessary, (b) a natural consequence of constituent structures and thus (c) nonderivational, I will now show that the theory of Sign-Based Morphology provides a superior account of the phonology-morphology interaction than can be found in the recent paradigmatic approach, represented by Steriade (1994b), McCarthy (1994), Benua (1995), Kenstowicz (1996), and others, which holds that interleaving effects are only apparent. According to the advocates of the paradigmatic approach to the phonology-morphology interaction, interleaving effects result from paradigm uniformity requirements: morphologically related words must be phonologically similar. These effects hold only between words (i.e., 'surface' or 'output' forms).

I illustrate this approach by summarizing Kenstowicz's (1996) Optimality Theoretic analysis of Northern Italian s-voicing, based on data from Nespor and Vogel (1986). In the relevant dialects, [s] and [z] are in complementary distribution, with [z] appearing intervocalically:

---

2. This of course does not mean that constituent structures are required in order to derive interleaving effects. Interleaving is in fact also an automatic consequence of realizational approaches to morphology such as that in Anderson (1992).

3. Becker (1993) makes this same point, and develops a paradigmatic theory of morphology that does not use constituent structures.

(18) azola 'button hole'
kaz-a 'house'
kaz-ina 'house-DIMINUTIVE'

As noted by Nespor and Vogel, s-voicing does not apply consistently across morpheme boundaries. The rule applies in (19a, c), but not in (19b):

(19) a. diz-onesto 'dishonest'
b. a-sotʃale 'asocial'
pre-sentire 'to hear in advance'
c. pre-zentire 'to have a presentiment'

Kenstowicz claims, following Nespor and Vogel, that the failure of s-voicing to apply in (19b) is connected to the fact that the stem is an independent word in these forms. The contrast between *rezistenza* and *asotʃale* is to be explained by the fact that the stem is an independent word in the latter but not in the former.[4] In the paradigmatic approach, this idea is implemented by invoking correspondence constraints between related words. Thus, identity constraints are enforced between *sotʃale* and *asotʃale*. By ranking the identity constraints higher than the phonotactic constraint responsible for s-voicing, the failure of voicing to apply to *asotʃale* can be accounted for. In the case of *dizonesto*, *dis* is not an independent word. Therefore, no paradigmatic correspondence constraints apply. There is nothing to block s-voicing.[5]

In this section, I will present challenges to this kind of approach and demonstrate that Sign-Based Morphology deals successfully with these challenges.

Arguments against the paradigmatic approach come from four general sources: (1) the 'inside-out' nature of interleaving effects, (2) the need to contrast cyclic and noncyclic phonology, (3) the fact that morphological constituents which are not possible words can nonetheless function as cyclic domains, and (4) underspecification effects (in which material in a daughter constituent is

4. This difference between bound and free morphs was noted by Kiparsky (1982a), who proposed to account for it by assuming that free morphs undergo a root cycle while bound morphs do not (see also Inkelas 1990). See Orgun (1994a, 1995c) for discussion of this issue from a Sign-Based Morphology perspective.

5. This account cannot, however, handle the contrast between *presentire* and *prezentire*, which both involve the same stem *sentire*, which occurs independently as a word. In Lexical Phonology terms, this contrast is accounted by invoking a pre-affixal stem cycle in one case but not the other (Inkelas 1990). In Sign-Based Morphology, this can be handled by enforcing phonological constraints on the daughter node of constructions. See Stump (1996) and Orgun (1995c) for discussion of this possibility.

underspecified; thus the daughter is not a possible word). I will discuss the first two of these in detail in the following sections, illustrating how Sign-Based Morphology copes successfully with the appropriate data.[6]

5.1 *Inside-Out Nature of Interleaving Effects*

Inside-out effects are those in which a morphologically simpler constituent affects the form of a morphologically more complex constituent of which it is a part, but not vice versa. This section illustrates the inside-out nature of interleaving on the basis of Turkish data. The phenomenon in question is a disyllabic size condition, which certain speakers of Turkish impose on suffixed forms (Itô and Hankamer (1989), Orgun and Inkelas (1992), Inkelas and Orgun (1995a).

(20) *Suffixed words must contain at least two syllables*

sol^j	'musical note "sol"'	do:	'musical note "do"'
sol^j-ym	'my "sol"'	*do:-m	Intended: 'my "do"'
sol^j-ym-y	'my "sol"-ACC'	*do:-m-u	Intended: 'my "do"-ACC'
sol^j-y	'"sol"-ACC'	do:-ju	'"do"-ACC'

The data in (20) show that monosyllabic roots may surface as words on their own. However, suffixed forms must contain two syllables. The form **do:-m-u* 'my "do"-ACC' is ungrammatical even though it contains two syllables by virtue of the extra accusative suffix it carries.

The paradigmatic approach would have to deal with this problem as follows: The form **do:-m-u* 'my "do"-ACC' is ungrammatical because the morphologically related form **do:-m* 'my "do"' is ungrammatical. Paradigm uniformity results in uniform ungrammaticality.

This account sounds fine until we confront the following question, so simple that one runs the risk of overlooking it: The ungrammaticality of the morphologically complex form **do:-m* 'my "do"' does not make the root *do* ungrammatical. Why?

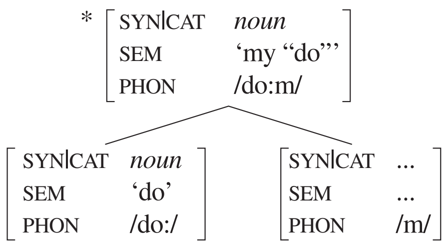
What we are seeing here is that the ungrammaticality of one form results only in the ungrammaticality of *more complex* related forms, not *less complex* related forms within the same paradigm. The immunity of the morphologically simpler form from paradigm uniformity effects follows from nothing in the theory. It must be stipulated (as 'primacy of the base' in Benua 1995 and

6. A good example of a bound stem acting as a cyclic domain is found in Dolbey (1996): allomorph selection in Sami needs to make reference to a bound stem that is not a possible word on its own.

McCarthy 1995). That is, the paradigmatic approach must *stipulate* the most basic property of interleaving effects, namely their inside-out nature.

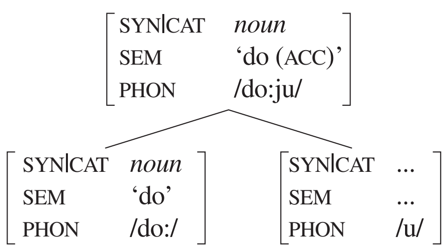
Not stipulation needs to be made in Sign-Based Morphology. Inside-out effects are a result of the basic architecture of the theory. Example (21) illustrates the structure of the ungrammatical form *do:m* 'my "do"'. This form is ungrammatical because the phonological string of the top node violates the disyllabic minimal size condition, which applies to all branching nodes.

(21)



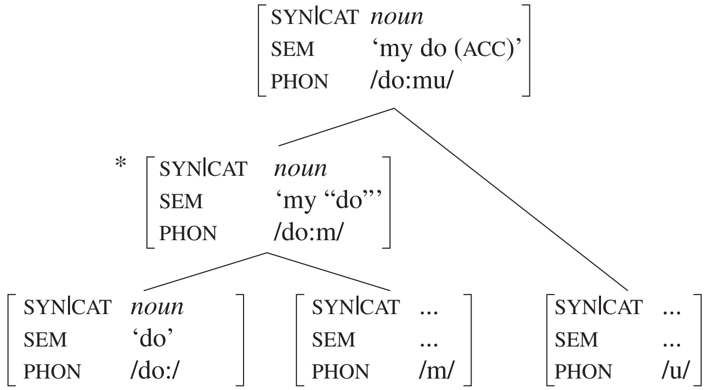
Example (22) shows the structure of the grammatical from *do:ju* 'do-ACC'. This form is grammatical because every node in the constituent structure satisfies all relevant grammatical requirements. In particular, there is no node that violates the disyllabic minimal size condition.

(22)



Example (23) illustrates the crucial form **do:-m-u* 'my "do"-ACC', which is ungrammatical even though it contains two syllables. This form is ungrammatical because not every node in the constituent structure satisfies every relevant grammatical requirement. In particular, the intermediate node **do:m* 'my "do"' violates the disyllabic minimal size requirement.

(23)



Finally, (24) shows the structure of the form *do* 'do'. This is the form whose grammaticality does not receive a satisfactory account in the paradigmatic approach. From a Sign-Based Morphology perspective, it is clear why this form is grammatical. There is no node that violates any grammatical condition. Note that there is no way a related morphological more complex form could have any effect on this word. This is because more complex forms, even if morphologically related, are not part of the representation of the simplex form *do*.

(24)

SYN|CAT	*noun*
SEM	'do'
PHON	/do:/

The same problem is present in Kenstowicz's analysis of Italian s-voicing I have discussed in Section 5. The main point of that analysis was that intervocalic s-voicing fails to apply in *asotʃale* because of identity constraints holding between this form and the related word *sotʃale*. The pair ⟨*sotʃale*, *asotʃale*⟩ satisfies identity better than the pair ⟨*sotʃale*, **azotʃale*⟩. However, it is also necessary to account for the fact that the pair ⟨**zotʃale*, *azotʃale*⟩, which satisfies identity just as well, is not attested. In general, why is it that morphologically simpler forms do not accommodate to constraints imposed on related more complex forms? A principled account of such inside-out effects is not possible within the paradigmatic approach. Proponents of this approach are forced into the unmotivated and arbitrary stipulation of 'the primacy of the base' just in order to encode this basic property of interleaving effects. In Sign-Based Morphology, the inside-out nature of interleaving effects is an automatic consequence of stating that some forms are morphologically simple, and morphologically complex forms may be derived from (or related to) them.

5.2 *Cyclic versus Noncyclic Effects*

The appeal of the paradigmatic approach, once one sets aside the problems noted above, lies in its ability to deal with interleaving effects. Correspondences between related words give rise to apparently cyclic phonological effects. However, the paradigmatic approach has no way to deal with *noncyclic* effects, the subject of the present section.

An excellent example of the contrast between cyclic and noncyclic phonology comes from the Turkish minimal size condition (Itô and Hankamer 1989; Orgun and Inkelas 1992; Inkelas and Orgun 1995a) that was introduced in Section 5.1, where it was shown that a word whose total size is two syllables may nonetheless be ungrammatical because it has a subconstituent that violates the disyllabic minimal size condition. Example (25) shows that the passive suffix as well as the possessive suffix may give rise to minimality violations.

(25)	a.	*monomorphemic forms*		b.	*suffixed forms (minimum 2 σ)*	
		je	'eat!'		*je-n	'eat-PASS'
		do:	'musical note do'		*do:-m	'do-1SGPOSS'
		jut	'swallow!'		jut-ul	'swallow-PASS'
		sol^j	'musical note sol'		sol^j-ym	'sol-1SGPOSS'

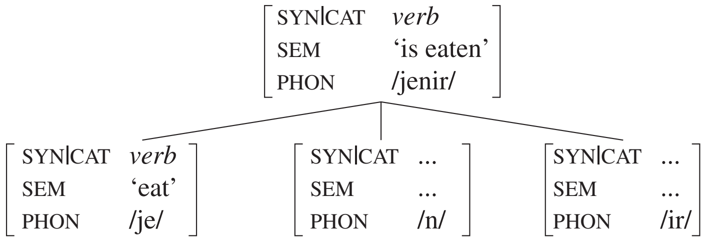
Although repair by adding more suffixes is not possible for possessive forms (26a), subminimal passive forms may be rendered grammatical by the addition of, for example, an aspect suffix, which brings the total size to two syllables (26b) (Orgun and Inkelas 1992; Inkelas and Orgun 1995a):

(26)	a.	*repair not possible*		b.	*repair possible*	
		*do:-m	'do-1SGPOSS'		*je-n	eat-PASS'
		*do:-m-u	'do-1SGPOSS-ACC'		je-n-ir	'eat-PASS-IMPRF'
		sol^j-ym-y	'sol-1SGPOSS-ACC'		jɯka-n-ɯr	'wash-PASS-IMPRF'

The forms in (26a) make sense in a cyclic approach, as we have seen in Section 5.1: the subconstituent N-POSS is subject to minimality, which it violates. The forms in (26b), however, call for a noncyclic approach: the whole word is grammatical if it is two syllables.

As in Orgun (1994a, 1995b, c,) I propose that the proper analysis of these forms requires allowing *n*-ary branching structures (where $n>2$). The form *je-n-ir* 'eat-PASSIVE-IMPRF′ has a ternary branching (i.e. 'flat') structure:

(27)



The constituent structure in (27) makes it clear why the form *je-n-ir* 'is eaten' is grammatical: There is no node here that violates the disyllabic minimal size condition.

The contrast between flat and branching structures, which allows Sign-Based Morphology to encode the difference between cyclic and noncyclic phonological effects, is not an ad-hoc stipulation. There is independent morphological evidence for the branching structures needed for the apparent cyclic enforcement of the minimal size condition in nominal forms (Orgun 1994b, 1995b, c, Inkelas and Orgun 1995a). The evidence comes from Suspended Affixation, described by Lewis (1967: 35) as a construction in which "one grammatical ending serves two or more parallel words."

An example is shown in (28a), where the nouns *sɯhhat* 'health' and *a:fijet* 'well being' are conjoined; the locative suffix, which has scope over both of them, is found only once at the end of the conjoined phrase. Example (28b) is similar. Further examples can be found in Lewis (1967), Underhill (1976), and Inkelas and Orgun (1995a).

(28) a. sɯhhat ve a:fijet-te
health and well-being-LOC
'in health and well-being'

b. halk-ɯn [adʒɯ ve sevintʃ-ler-i]
people-GEN sorrow and joy-PL-POSS
'the people's sorrows and joys'

Example (29) shows the null hypothesis for the structure of this construction. I assume that the constituent structure is as implied by the scope relations, with the locative suffix attached to the whole conjoined NP.

(29) [[sɯhhat ve a:fijet] te]

There are initially puzzling restrictions on the combinations of affixes that Suspended Affixation can target. As seen in (30a), it is possible to suspend *all* eligible affixes. Here, the plural suffix *-ler*, the possessive *-im*, and the accusative suffix *-i* are all suspended. Example (30b) shows that it is acceptable not to suspend any affixes at all. Here, all suffixes are realized on both conjuncts.

(30) a. All affixes suspended
[tebrik ve teʃekkyr]-ler-im-i
[congratulation and thank]-PL-1SGPOSS-ACC
'my congratulations and thanks (acc)'
b. No affixes suspended
tebrik-ler-im-i ve teʃekkyr-ler-im-i

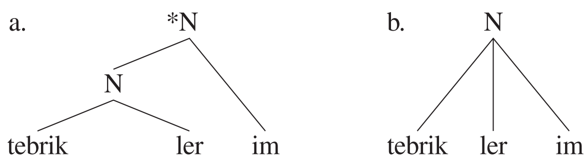
Example (31) shows the puzzling restrictions on Suspended Affixation. In (31a), we see that it is possible to suspend just the accusative suffix *-i* while realizing the plural and possessive suffixes on both conjuncts. Example (31b) shows that it is *not* possible to realize the plural suffix *-ler* on both conjuncts while suspending the possessive and accusative suffixes.

(31) *Suspension of some but not all affixes*
a. [tebrik-ler-im ve teʃekkyr-ler-im]-i
b. *[tebrik-ler ve teʃekkyr-ler]-im-i

Our task is to account for this inseparability of the plural and possessive suffixes in Suspended Affixation. That is, we need to find a formal account of the observation that the plural and possessive suffixes are either both realized on all conjuncts or both suspended.

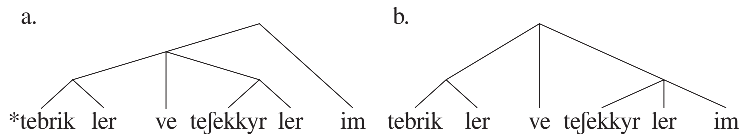
I offer an analysis of this seemingly strange restriction in terms of constituent structure. I claim that the plural and possessive suffixes form a flat (ternary branching) structure with the base they attach to, as in (32b), rather than a binary branching structure as in (32a).

(32) a. b.



The plural and possessive suffixes have to be sisters whenever they are both present (32b). Given that the plural and possessive suffixes form a ternary branching structure with the base they attach to, the pattern of suspension in (33) is ungrammatical. This example is similar to the one we have seen before in (31b), except that the accusative suffix is not involved here. (This further supports the position that the source of the problem is the configuration of the plural and possessive suffixes). There are two possible structures for this example. The first is shown in (33a). Here, the possessive suffix is attached to the conjoined NP, as it has scope over both conjuncts. This configuration violates the condition that the plural and possessive suffixes must be sisters whenever they both have scope over the same head. Therefore, this structure is ruled out. This leaves us with the possibility in (33b), which is structurally well formed. However, this structure does not give us the desired scope relations. In particular, the possessive suffix has scope over the second conjunct but not the first conjunct. Therefore, we explain the fact that the plural and possessive suffixes have to be suspended together, or not suspended at all.

(33) *[tebrik-ler ve teʃekkyr-ler]-im
[thank-PL and congratulation-PL]-1SGPOSS



In general, then, suffixes can be separated in Suspended Affixation only if they form a hierarchical structure. If they form a flat structure, they have to be suspended as a group, or not at all. See Orgun (1995b, c) for more details.

Suspended Affixation data show that the possessive and accusative suffixes come in a hierarchical structure. Sign-Based Morphology predicts cyclic effects whenever hierarchical structures are found. Indeed, the minimal size condition exhibits cyclic effects in possessed accusative forms (**do:-m-u* ‘my “do”-ACC’ is ungrammatical even though it contains two syllables, because the subconstituent **do:-m* ‘my “do”’ is subminimal).

Thus, Sign-Based Morphology not only accounts for both cyclic and noncyclic phonology, but also relates the contrast to independently needed morphological structure. Past derivational approaches (e.g. Kiparsky 1982a, Mohanan 1982; Halle and Vergnaud 1987; Halle and Kenstowicz 1991) had to stipulate the difference between cyclic and noncyclic phonology, a distinction not motivated in any way by the morphology.

6. Conclusions

The first conclusion of this study is that phonology-morphology interleaving is necessary if a principled account of certain types of phonology-morphology interaction is to be found.

The second conclusion is that there is nothing derivational about phonology-morphology interleaving. Thus, to take up the question that is the topic of this volume, is there then a derivational residue in phonology in this age of constraint based-theories? The answer depends solely on the nature of phonological theory. Past researchers (e.g. Lakoff 1993) have often failed to distinguish cyclic phonology from rule ordering within a cycle. The quest for a nonderivational theory of phonology, which should be limited in scope to eliminating rule ordering within a cycle, has erroneously been taken to entail endeavoring to eliminate cyclic phonology as well. As I have shown in this paper, any theory that utilizes constituent structures and feature percolation is able to derive interleaving effects from declarative constraints on static phrase structure configurations.

One question that must be answered is the following: Sign-Based Morphology is nonderivational, but is it empirically and theoretically superior to derivational approaches to interleaving (such as Lexical Phonology)? Is it superior to other nonderivational conceptions of the phonology-morphology interface? I contend that only Sign-Based Morphology has the virtue that cyclic and noncyclic effects follow from independently motivated morphological structure. Within past cyclic approaches, the cyclic-noncyclic distinction is stipulated. Current paradigmatic approaches have no way of addressing noncyclic effects at all.

Acknowledgments

I am grateful to Larry Hyman, Andreas Kathol, Jaye Padgett, Tom Shannon, Jean-Pierre Koenig, Dan Jurafsky, David Perlmutter, Sharon Inkelas, Gene Buckley, Laura Benua, Chris Golston, Cheryl Zoll, Susanne Riehemann, Farrell Ackerman, Adele Goldberg, Paul Kay, Charles Fillmore, Collin Baker, Jonathan Segal, Mark Ellison, Jim Scobbie, Markus Walther, Andy Dolbey, András Kornai, Chris Manning, David Odden for discussing this paper with me.

Derivationalism in Kikamba Vowel Hiatus Phenomena

R. Ruth Roberts-Kohno

1. Introduction

Recent developments in phonological theory have seen a shift from rule-based theories of languages to constraint-based theories with enriched representations. This is most clearly seen in the Optimality Theory framework of Prince and Smolensky (1993) and elsewhere. One of the central features which has distinguished Optimality Theory (OT) from derivational phonology is that derivational theories assume a chain of steps from the underlying form to the surface, with each step mediated by the application of a rule. Optimality Theory, on the other hand, has assumed a single step from the underlying to the surface form, mediated by the functions Gen and H-Eval. The Bantu language Kikamba has phonological phenomena of a type which challenge this view of the organization of the grammar. The goal of this paper is to present these phenomena, and to consider the theoretical implications of the data for a non-derivational model. In underlying and certain derived representations, Kikamba has empty skeletal positions (henceforth, empty Cs), which we will argue should be analyzed as root nodes unspecified for phonetic content. These root nodes play an active role up to a certain point in the grammar. After that point, they are deleted. We will show that a monostratal constraint-based approach cannot handle these facts.

A central issue in understanding the role of empty Cs is the fact that when two vowels come together through morpheme concatenation, they coalesce forming a single syllable. However, there are certain morphemes which do not coalesce with an adjacent vowel. The lack of syllable fusion in places where it would otherwise be expected to occur can be explained by the presence of an empty C. For example, there are numerous minimal pairs in Kikamba where the only difference is whether vowel hiatus is maintained or eliminated, as seen in (1).

(1)	a.	/ko-að-a/ infin-stem-infl	kwaàðà[1]	‘govern’
	b.	/ko-Cáð-a/	ko.a̋ða̋	‘shoot’

The lack of vowel coalescence in cases like (1b) can be explained by postulating an empty root node in just those places where vowel hiatus is maintained, in contrast to those vowel-vowel sequences where it is eliminated, as in (1a). Another type of evidence supporting the empty C is that the choice of excrescent consonant after the 1st singular Object prefix depends on the presence of the empty C, as will be demonstrated shortly.

While there is strong evidence for the empty C at certain stages of the grammar, other processes such as vowel shortening crucially apply as though the empty C is not present. An example is seen in (2b), as compared with (2a).

(2)	a.	/né-tó-a-kon-â/ t/a-1pl-t/a-stem-infl	nétwáakonâ	‘we just hit’
	b.	/né-tó-a-Ceβ-á	nétwá.eβá	‘we just paid’

Although an empty C is present in the verb stem ‘pay,’ the preceding long vowel shortens as though the empty C were not there, because, as we will demonstrate, the empty C has been deleted.

We argue that the empty C is present in the underlying representation, but is deleted later in the grammar. Moreover, this deletion must be extrinsically ordered with respect to other rules of Kikamba. Therefore, in order to handle the vowel hiatus and coalescence facts of Kikamba, the equivalent of the derivational devices of extrinsic ordering and deletion are needed. We demonstrate that, while a monostratal constraint-based account of Kikamba does not work, a multistratal account is able to handle the facts. While recent work in Correspondence Theory allows deletion, the levels that must be postulated weaken the basis of a constraint-based approach by allowing the postulation of levels not independently justified in the language.

Section 2 briefly reviews the evidence supporting the empty C analysis. We then review evidence that the empty C deletes later in the grammar in Section 3. Section 4 further supports the empty C analysis with instrumental phonetic evidence. Finally, we demonstrate that the only way to incorporate the devices

1. Kikamba is a four tone language. Tones are represented as follows: super-high = ″, high = ′, low = unmarked or -, and super-low = `.

of rule-ordering and deletion into a constraint-based theory such as OT is to allow the postulation of multiple levels, where such levels are not independently justified.

2. Evidence for the Presence of Empty Cs

2.1 *Vowel Hiatus Phenomena*

We will now consider some of the evidence for empty Cs. More extensive evidence can be found in Roberts-Kohno (1995, 1998). The first type of evidence is vowel hiatus. In Kikamba, when two vowels come together by morpheme concatenation, the underlying vowel sequence coalesces, forming a single surface syllable. This pattern of hiatus resolution is due to the fact that onsetless syllables are dispreferred in Kikamba, and they are avoided whenever possible. This is explained by the Onset Principle. Processes such as Glide formation and Mid vowel fusion function to eliminate vowel hiatus, and thereby eliminate violations of the onset constraint.

The first strategy Kikamba uses, Glide formation, is a process by which the first of two vowels becomes a glide, thereby losing its mora. The second vowel spreads to that mora, resulting in compensatory lengthening. Examples of Glide formation with compensatory lengthening in the infinitive form of the verb are seen in (3).

(3)	a.	/ko-ák-a/ infin-stem-infl	kwa̋a̋ka̋	‘build’
	b.	/ko-enok-a/	kweenokà	‘go home’
	c.	/ko-ít-a/	kwı̋ı̋ta̋	‘strangle’

However, sometimes a surface vowel-vowel sequence does not coalesce and the vowels remain in separate syllables. We propose that in such cases, the onset principle is not violated because these syllables *do* in fact have an onset. That onset is an empty root node, which is notated C and syllabifies as the onset of a syllable. In surface forms, a period (.) represents a syllable break. This notation is seen in (4).

(4) C = empty root node (.) = syllabification break

Let us look at an example of this contrast. In (5), we have a minimal pair with the verbs ‘like’ and ‘go’. In (5a), the first syllable /ko/ coalesces with the vowel of the verb stem, resulting in Glide formation and compensatory lengthening.

(5) *Glide formation and compensatory lengthening*

a.	/ko-ɛnd-a/	kwɛɛ̀ndà	'like'
	infin-stem-infl		
b.	/ko-Cɛnd-a/	ko.ɛɛ̀ndà	'go'

In (5b), however, vowel hiatus is maintained: the sequence [o + ɛ] does not coalesce. The failure of syllable coalescence in (5b) is explained under the hypothesis that there is an empty root node between the vowels which functions as the onset of the syllable [ɛ], thus eliminating the motivation for syllable fusion. This robust contrast can be demonstrated with numerous other near minimal pairs, as seen in (6).

(6) *Glide formation and compensatory lengthening with other verb roots*

a.	/ko-ín-a/	kwı̋ı̋na̋	'dance'
b.	/ko-Cind-a/	ko.iìndà	'submerge'
c.	/ko-ɛlɛkɛl-a/	kwɛɛlɛkɛlà	'face towards'
d.	/ko-Cɛk-a/	ko.ɛkà	'stop/leave'
e.	/ko-ák-a/	kwa̋a̋ka̋	'build'
f.	/ko-Calyool-a/	ko.alyoòlà	'translate'
g.	/ko-ɔ́t-a/	kwɔ̋ɔ̋ta̋	'warm self'
h.	/ko-Cɔ́ɔt-a/	ko.ɔ̋ɔ̀tà	'dream'
i.	/ko-ék-a/	kwe̋e̋ka̋	'do'
j.	/ko-Ceβ-a/	ko.eβà	'pay'

A second strategy Kikamba uses to eliminate vowel hiatus is Mid vowel fusion. Mid vowel fusion is a process which occurs if the first vowel is Low and the second vowel is non-High. When /a/ is followed by a Mid vowel, the resulting coalesced syllable is a long lower mid vowel, as demonstrated in (7a).

(7) *Mid vowel fusion*

a.	a + e, ɛ → ɛɛ	a + o, ɔ → ɔɔ	
b.	/ko-má-ɛnd-a/	komɛ̋ɛ̀ndà	'like them'
	infin-3pl-stem-infl		
c.	/ko-má-Cɛnd-eC-a/	komá.ɛndɛ.à	'go for them'

(7b, c) represent a near minimal pair. We see that Mid vowel fusion takes place in (7b), but hiatus is maintained in (7c). Interestingly, it is exactly the same set of verbs which do not allow hiatus to be resolved through Glide formation which also do not allow hiatus to be resolved through Mid vowel fusion. This pattern can also be seen in numerous other verb roots, as demonstrated in (8).

(8) *Mid vowel fusion with other verb roots*

a.	/ko-má-enz-eC-a/	koméénzɛ.à	'dig for them'
b.	/ko-má-Ceβ-a/	komá.eβà	'pay them'
c.	/ko-má-ɔ́s-eC-a/	komɔ́ɔ́sɛ.à	'take for them'
d.	/ko-má-Cɔ́ɔt-eC-a/	komá.ɔ́ɔtɛ.à	'dream for them'
e.	/ko-má-ókit-a/	komɔ́ɔ́kità	'fight them'
f.	/ko-má-Cóm-a/	komá.ő́mã́	'bite them'

The postulation of an underlying empty C in such stems accounts for all cases where vowel hiatus is maintained.

The hypothesis that there are phonetically unrealized root nodes is further supported by the fact that the maintenance of vowel hiatus also occurs in other environments. In other words, the empty C can be found in all classes of morphemes. For example, in (9) we find examples of infinitives which have the applied suffix /eC/. There is evidence that the applied suffix also contains an empty C. (9a) indicates that an empty C is present because there is no coalescence between the vowel of the applied suffix /eC/ and the final vowel /a/, as would otherwise be expected when two morphemes are concatenated.

(9) *Applied suffix /eC/*

a.	/ko-suuŋg-eC-a/ infin-stem-appl-infl	kosuuŋge.à	'guard for'
b.	/ko-βul-eC-a/	ko-βul-el-à	'go through pile for'
c.	/ko-lées-eC-a/	ko-lées-el-à	'climb mountain for'

Notice also that when the stem contains /l/, an /l/ appears in the applied suffix. This can be explained as the rightward spreading of [lateral] from /l/ to an empty C. Examples of this are seen in (9b, c). The feature [lateral] spreads and surfaces as /l/ just in case hiatus is maintained between two vowels. The empty root node is the only position the [lateral] feature is able to spread to, and thus further confirms the presence of the empty C.

2.2 *The 1st Singular Object prefix*

The second type of evidence for empty Cs comes from alternations involving the 1st singular Object prefix /-N-/. There is a clear distinction between V-initial and empty C-initial verb roots, which supports the existence of the empty root node.

In Kikamba, the 1st singular Object prefix is underlyingly a moraic nasal unspecified for place of articulation. In the data in (10), we see this prefix before a consonant-initial verb. Notice that the nasal assimilates to the place of articulation of the following consonant.

(10)		*UR with 1st sing. OP*	*1st sing. OP form*	*gloss*
	a.	/ko-N-βiindo-a/ infin-1sg-stem-infl	koo<u>mb</u>iindo.à	'wake me up'
	b.	/ko-N-tál-a/	koo<u>nd</u>a̋la̋	'count me'
	c.	/ko-N-kon-a/	koo<u>ŋg</u>onà	'hit me'
	d.	/ko-N-suuŋg-a/	koo<u>nz</u>uùŋgà	'guard me'

2.2.1 *Different allomorphs surface for V- and C-initial verbs*

However, in verb roots that phonetically begin with a vowel, an excrescent consonant surfaces after the prefix. This supports the existence of the empty C because a different excrescent consonant surfaces for verbs which begin with an empty C versus verbs which are truly vowel-initial. In the vowel-initial verbs in (11), we find the excrescent consonant [b].

(11) *Vowel-initial verbs surface with [b]*

	infinitive	*1st sing. OP form*	*gloss*
a.	kwa̋a̋ka̋	koomba̋ka̋	'strengthen me'
b.	kweènzà	koombenze.à	'dig for me'
c.	kwɛ̋ɛ̋nza̋	koombɛ̋nza̋	'shave me'
d.	kwɪ̋ɪ̋na̋	koombíne.à	'sing for me'
e.	kookità	koombokità	'fight me'
f.	kwɔ̋ɔ̋na̋	koombɔ̋na̋	'see me'
g.	kúúmelà	koombúmelà	'appear to me'

However, the excrescent consonant for the 1st singular Object prefix in empty C-initial verbs is different, as seen in (12). For such verbs, the excrescent consonant is [d].

(12) *Empty C-initial verbs surface with [d]*

	infinitive	*1st sing. OP form*	*gloss*
a.	ko.a.à	koonda.e.à	'divide for me'
b.	ko.eβà	koondeβà	'pay me'
c.	ko.ɛkà	koondɛkà	'leave me'
d.	ko.i.ità	koondi.ità	'treat me'
e.	ko.o.à	koondo.e.à	'buy for me'
f.	ko.ɔ.à	koondɔ.à	'bewitch me'
g.	ko.ʊ̋.a̋	koondú.e.à	'cook for me'

Notice that it is exactly those verbs which do *not* allow coalescence with a preceding vowel that insert [d] after the 1st singular Object prefix. And, it is those verbs which *always* coalesce with a preceding vowel that insert [b] after this prefix. The important point to realize is that, however the consonant gets inserted, whether by rule or constraint, the insertion must be sensitive to the distinction between empty C-initial verbs and vowel-initial verbs, thereby supporting the existence of the empty root node.

2.2.2 *Stem Vowel Length is Different in V- and C-initial Verbs*

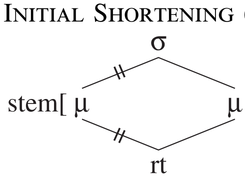
Furthermore, there is an alternation in the length of stem-initial vowels associated with selecting the 1st singular Object prefix, which supports the existence of the empty C. In (13), the stem vowel is underlyingly long. However, the stem vowel surfaces as short following the 1st singular Object prefix, as seen in (13b).

(13) *Vowel-initial verbs surface with a short vowel following the 1st sing OP /-N-/*

a.	*infinitive*		*gloss*
	/ko-óok-a/	kőòkà	'come'
	infin-stem-infl		
b.	*1st sing. OP form*		
	/ko-N-óok-eC-a/	koombó̲ke.à	'come for me'
	infin-1sg-stem-appl-infl		

Such data provide evidence for a rule of Initial shortening (IS), seen in (14). Initial shortening shortens a long onsetless syllable which is in verb stem-initial position.

(14) INITIAL SHORTENING (stem level)



(15) provides examples of empty C-initial verbs. Whereas a long vowel-initial verb surfaces with a short vowel after the 1st singular Object prefix, an empty C-initial verb retains its long vowel, as seen in (15b). This contrast provides evidence for the empty C at the stem level, since the empty C serves as the onset of the verb stem. Since the long vowel is not in stem-initial position, Initial shortening is blocked.

(15) *Empty C-initial verbs surface with a long vowel following the 1st sing. OP /-N-/*

a.	*infinitive*		*gloss*
	/ko-Cɔ́ɔt-a/	ko.ɔ̋ɔ̀tà	'dream'
	infin-stem-infl		
b.	*1st sing. OP form*		
	/ko-N-Cɔ́ɔt-eC-a/	koondɔ́ɔtɛ.à	'dream about me'
	infin-1sg-stem-appl-infl		

(16) provides more examples of stem-initial long vowels that follow the same pattern: if the verb begins with an empty C, Initial shortening cannot apply and the vowel surfaces as long.

(16) *Vowel-initial verbs: vowel surfaces as SHORT following the 1st sing OP /-N-/*

a.	kwaambatà	koombambatyà	'cause me to go up'
b.	kweènzà	koombenze.à	'dig for me'
c.	kwɛ̋ɛ̋nza̋	koombɛ̋nza̋	'shave me'
d.	kőőmba̋	koombőmba̋	'mold me'

Empty C-initial verbs: vowel surfaces as LONG following the 1st sing OP /-N-/

e.	ko.aandekà	koonaandeke.à	'write for me'
f.	ko.ɛɛ̀ndà	koonɛɛndɛ.à	'go for me'
g.	ko.ooŋgamyà	koonooŋgamyà	'stop me'

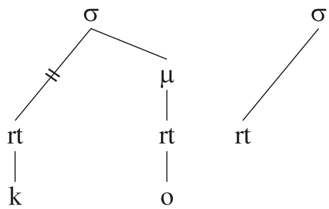
3. Evidence that Empty Cs Have Been Deleted

In addition to strong evidence supporting the existence of empty root nodes, there is also evidence that such root nodes are deleted after a certain point in the derivation. The evidence for this conclusion will now be considered.

3.1 *Prefix /k/-Deletion*

There is a rule of prefix /k/-deletion, seen in (17), which deletes /k/ in the prefix /ko/, just in case the following syllable has an onset.

(17) PREFIX-/k/ DELETION



post-vocalic, left-to-right, non-iterative

3.1.1 *[k] Deletes in the 2nd Singular Object Prefix /ko/*

In (18–20), we find examples with the 2nd singular Object prefix /ko/. In vowel-initial verbs, prefix /k/-deletion does not apply, as seen in (18). This motivates the presence of /k/ in this prefix.

(18) *Vowel-initial verbs surface with [k]*

	UR with 2sOP	*2sOP form*	*gloss*
a.	/ko-ko-ák-a/ infin-2sg-stem-infl	kokwa̋a̋ka̋	'build you-sg'
b.	/ko-ko-ɔ́n-a/	kokwɔ̋ɔ̋na̋	'see you-sg'

In (19), we find examples of phonetically-realized consonant-initial verb stems. Prefix /k/-deletion applies here, so underlying /ko/ shows up as [o].

(19) *Consonant-initial verbs surface without [k]*

	UR with 2sOP	*2sOP form*	*gloss*
a.	/ko-ko-tál-a/ infin-2sg-stem-infl	ko.ota̋la̋	'count you-sg'
b.	/ko-ko-βálok-i-a/	ko.oβáloča̋	'make you-sg fall'
c.	/to-káa-ko-tál-á/	toká.otálâ	'we will count you-sg'

Unexpectedly, verbs which begin with an empty C pattern with vowel-initial verbs, as seen in (20). Such verbs select [ko] as the 2nd singular Object prefix, rather than [o].

(20) *Empty C-initial verbs surface with [k]*

	UR with 2sOP	*2sOP form*	*gloss*
a.	/ko-ko-Calyool-a/ infin-2sg-stem-infl	koko.alyoòlà	'change you-sg'
b.	/ko-ko-Cɔ́ɔt-eC-a/	koko.ɔ́ɔtɛ.à	'dream for you-sg'

That is, for this phenomenon, the empty C-initial verbs behave as though they are vowel-initial phonologically, as well as phonetically. The empty C, in effect, cannot be present. Nevertheless, notice that vowel hiatus is still maintained before these empty C-initial verbs, showing that there is an underlying empty C in these stems.

3.1.2 *[k] Deletes in the Cl.15 Infinitive Prefix /ko/*

In the present progressive, the [k] of the infinitive prefix /ko/ will also delete in exactly the same environment. Examples are seen in (21). As expected, the infinitive prefix surfaces as /ko/ in vowel-initial verbs, as in (21a, b). Consonant-initial verbs take [o], as seen in (21c,d). And in (21e,f), empty C-initial verbs take [ko].

(21)	*Present progressive*		*/ t/a + SP + inf. ko + verb + fv /*
	V-initial verbs		
	a.	némákwɛ̋ɛ̋nza̋	'they are shaving'
	b.	némákwa̋a̋ka̋	'they are building'
	C-initial verbs		
	c.	némá.ota̋la̋	'they are counting'
	d.	némá.oβikà	'they are arriving'
	Empty C-initial verbs		
	e.	némáko.amokà	'they are waking up'
	f.	némáko.eβà	'they are paying'

3.1.3 *[k] Does* ***not*** *Delete in the Cl.15 Object Prefix /kó/*

Finally, note that Prefix /k/-deletion is a lexical rule. The class 15 Object prefix, which is also realized as /kó/, never undergoes Prefix /k/-deletion. (22) provides examples of the non-reducing class 15 Object prefix /kó/ in contrast with the reducing 2nd singular Object prefix /ko/.

(22)	*2nd sing. OP /ko/ vs. cl.15 OP /kó/*			
		UR	*surface form*	*gloss*
	a.	/ko-ko-titiC-a/	ko.otiti.à	'rub you-sg'
	b.	/ko-kó-titiC-a/	kokótiti.à	'rub it /cl.15'
	c.	/ko-ko-βák-a/	ko.oβa̋ka̋	'rub/smear on you-sg'
	d.	/ko-kó-βák-a/	kokóβa̋ka̋	'smear on it /cl.15'
	e.	/né-ne-ko-ko-tál-aC-a/	nééŋgo.otála.à	'I've counted you-sg'
	f.	/né-ne-ko-kó-tál-aC-a/	nééŋgokótála.à	'I've counted it /cl.15'
	g.	/ne-káa-ko-kon-á/	ŋgá.okonâ	'I will hit you-sg'
	h.	/ne-káa-kó-kon-á/	ŋgáakókonâ	'I will hit it /cl.15'

These facts can be easily handled if we assume that there is a rule deleting empty Cs and that deletion of the empty C is ordered before deletion of prefixal /k/. In other words, we are now looking at the part of the grammar which is ordered after empty C deletion.

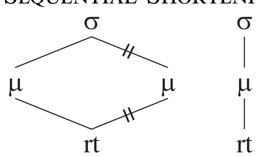
3.2 *Sequential Shortening*

The second case of transparent empty Cs can be seen by considering Sequential shortening (SS). In Kikamba, a long vowel can never be immediately followed by another vowel. Sequential shortening shortens a long vowel just in case it is followed by another vowel. Although the vowels in the vowel sequences of (23) never coalesce, showing that there is an empty C, the empty C is transparent to this vowel shortening process.

(23)	a.	/né-tó-a-tál-á/ t/a-1pl-t/a-stem-infl	nétwáatálâ	'we just counted'
	b.	/né-tó-a-Cóm-á/	nétwá.ómâ	'we just bit'
	c.	/né-né-a-sέεmb-aC-á/ t/a-1sg-t/a-stem-t/a-infl	nénáasέεmba.â	'I always run'
	d.	/né-né-a-Cεnd-aC-á/	néná.εnda.â	'I always go'

In (23a), /..tó + a../ coalesce by Glide formation with the expected compensatory lengthening of the second vowel and surface as [..twáa..]. In (23b), Glide formation and compensatory lengthening have also taken place, but the resulting surface vowel is short due to Sequential shortening. Another example of this alternation is seen in (23c-d). Sequential shortening is formulated in (24).

(24) Sequential Shortening



The data in (25) provide more examples of Sequential shortening. In all these cases, two vowels coalesce, and we therefore expect a long vowel to surface. However, the vowel is short due to Sequential shortening.

(25) *More examples of Sequential shortening*

a.	/ko-éC-a/	kwé.ά̋	'tell'
b.	/né-mó-éC-éet-έ/	némwé.éetε̂	'you-pl told'
c.	/né-né-éC-aC-á/	néné.a.â	'I always tell'
d.	/né-tó-éC-aC-á/	nétwé.a.â	'we always tell'

e.	/né-né-a-CeC-á/	néná.e.â	'I just cried'
f.	/né-tó-a-CoC-á/	nétwá.o.â	'we just bought'
g.	/ko-óCaC-a/	kó.a.à	'kill'

Again, under the assumption that there is a rule deleting empty Cs, Sequential shortening takes place at a point in the grammar after empty Cs have deleted.

3.3 *Initial Shortening in the Imperative*

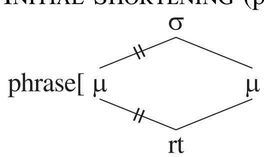
Further evidence for the deletion of the empty Cs comes from the imperative. The singular imperative is formed by taking the verb stem and adding the final vowel [a], as seen in (26).

(26)

	Verb stem	*Imperative*	*gloss*
a.	-tál	tálā	'count'
b.	-ŋɔlɔt	ŋɔlɔ́tā	'snore'
c.	-okit	okítā	'fight'

However, when a verb begins with a long vowel, that initial vowel shortens in the imperative. This indicates that Initial shortening is also applicable at the phrase level. Initial shortening is repeated in (27) (cf. (14) for the initial formulation):

(27) Initial Shortening (phrase level)



Examples of this shortening are given in (28a-f).

(28) *Vowel-initial verbs*

	infinitive		*imperative*	*gloss*
a.	/ko-óok-a/	kőòkà	óka	'come'
b.	/ko-ambat-a/	kwaambatà	ambáta	'go up'
c.	/ko-enz-a/	kweènzà	enzâ	'dig'
d.	/ko-ɛ́nz-a/	kwɛ̋ɛ̋nzɑ̋	ɛ́nza	'shave'
e.	/ko-ómb-a/	kőőmbɑ̋	ómba	'mold'
f.	/ko-ɔŋgel-a/	kwɔɔŋgɛlà	ɔŋgɛ́la	'increase'

In this data, we observe that the long vowel of the verb shortens only if the vowel is at the beginning of the verb stem. Based on the contrasting patterns of vowel shortening at the stem level associated with the selection of the 1st singular Object prefix, we might expect that long vowels in stems which begin with empty Cs would *not* shorten in the imperative. However, as seen in (29), the initial long vowel does shorten, providing evidence that the empty C has in fact been deleted.

(29) Empty C-initial verbs

	infinitive		*imperative*	*gloss*
a.	/ko-Cɔ́ɔt-a/	ko.ɔ̋ɔ̀tà	ɔ́ta	‘dream’
b.	/ko-Caasi-a/	ko.aàsyà	asyâ	‘pay dowry’
c.	/ko-Candek-a/	ko.aandekà	andéka	‘write’
d.	/ko-Cɛnd-a/	ko.ɛɛ̀ndà	ɛ́nda	‘go’

Therefore, phrasal Initial shortening also seems to ignore the empty C. This is the third case where the empty Cs have been deleted.

4. Phonetic Evidence for the Presence and Deletion of Empty Cs

Finally, phonetic data also support the existence of empty Cs, and help to determine when they are deleted. As is evident from many examples previously seen, a heterosyllabic sequence of identical short vowels is phonetically distinct from a single long vowel. This syllabification difference is correlated with duration measurements. Contrasting pairs are seen in (30):

(30)

a. Monosyllabic vowel sequences	b. Bisyllabic vowel sequences
/ko-má-að-a/ koma̋àðà ‘govern them’	/ko-má-Cað-a/ komá.aðà ‘shoot them’
/ko-ké-ék-a/ kokɛ̋ɛ̋ka̋ ‘possess $\text{him}_{\text{cl.7}}$’	/ko-ké-Ceβ-a/ koké.eβà ‘pay $\text{him}_{\text{cl.7}}$’

Duration measurements for monosyllabic long vowel sequences, as in Column (a), and bisyllabic vowel sequences as in Column (b) demonstrate that there is a significant difference in duration between these two categories. These duration differences are the direct result of syllabification contrasts. Empty Cs are able to account for these syllabification contrasts: such contrasts arise because empty Cs block syllable fusion. (31) shows the kinds of vowel sequences which were measured. Recall that if the two V’s are separated by a period (.), this indicates a syllabification break, and thus, the presence of an empty C. (31b, c) provide specific examples of the measured syllables, which are underlined.

(31) a.

Representation of vowel sequences	monosyllabic	bisyllabic
identical vowel sequences	V_iV_i	$V_i\ .V_i$
non-identical vowel sequences	V_iV_j	$V_i\ .V_j$

b. *identical vowels*

V_iV_i	/ko-tó-óm-i-a/	kotőőmya̋	'dry us'
$V_i\ .V_i$	/ko-tó-Cóm-a/	kotó.őma̋	'bite us'

c. *non-identical vowels*

V_iV_j	/ko-má-óm-i-a/	komɔ̋ɔ̋mya̋	'dry them'
$V_i\ .V_j$	/ko-má-Cóm-a/	komá.őma̋	'bite them'

Figure 1 presents the results of comparing the monosyllabic and bisyllabic vowel sequences. Notice that both when the vowels are the same, as in Figure 1a, or different, as in Figure 1b, the result is that a bisyllabic short vowel sequence has a significantly greater duration than a monosyllabic long vowel.

Figure 1. Durational contrast between a monosyllabic long vowel-VV and a bisyllabic vowel sequence-V.V

	duration	N		duration	N	p-value
a. V_iV_i	150 ms	194	$V_i.V_i$	214 ms	39	$p < .001$
b. V_iV_j	179 ms	25	$V_i.V_j$	211 ms	58	$p < .001$

Therefore, since empty Cs affect the syllabification of vowel sequences, we can use duration evidence to indirectly detect empty Cs.

Let us now turn to imperatives. Phonological evidence from phrasal Initial shortening has shown that shortening occurs in all imperatives, regardless of whether the verb begins with an empty C, or with a vowel. Therefore, the empty C must be deleted at the phrase level. Given that, there should be neutralization of vowel durations across a phrasal boundary for roots that underlyingly begin with a vowel versus those beginning with an empty C; this can be checked by seeing if the vowel sequence durations are the same across categories. In order to test this, syllable durations of post-lexical vowel-vowel sequences were measured using the imperative with a preceding subject. Examples are seen in (32); the measured syllable is underlined.

(32) *Phrasal examples*

a. *identical vowels*

$V_i\#V_i$ Mweema aða 'Mweema govern!'

$V_i\#CV_i$ Mweema aða 'Mweema shoot!'

b. *non-identical vowels*

$V_i\#V_j$ Mweema okita 'Mweema fight!'

$V_i\#CV_j$ Mweema oma 'Mweema bite!'

As seen in Figure 2, there is no difference in phrase level vowel duration between underlying vowel-initial verbs and empty C-initial verbs. The results are the same whether the vowels in the sequence are identical, as in Figure 2a, or different, as in Figure 2b. This supports the conclusion that empty Cs have been deleted at the word level.

Figure 2. Duration differences between phrasal V # V and V # CV

	duration	N		duration	N	p-value
a. $V_i\#V_i$	150 ms	25	$V_i\#CV_i$	163 ms	7	$p > .22$
b. $V_i\#V_j$	191 ms	19	$V_i\#CV_j$	181 ms	19	$p > .27$

Given that the durations of vowel sequences are the same at the phrase level, are they the same as or different from word-internal single-syllable durations? If they are the same as word-internal disyllabic short vowel sequences, this would suggest that there is *no* resyllabification at the phrase level. If they are the same as word-internal monosyllabic long vowels, this would constitute an argument *for* resyllabification at the phrase level.

Figure 3. Durational contrast between V#(C)V and VV

	duration	N		duration	N	p-value
a.V_i# $(C)V_i$	153 ms	32	V_iV_i	150 ms	194	$p > .686$
b.V_i# $(C)V_j$	186 ms	38	V_iV_j	179 ms	25	$p > .342$

Figure 3 indicates that the duration of a word-internal long vowel is the same as the duration of a post-lexical vowel sequence. In other words, there *is* resyllabification at the phrasal level. Moreover, a disyllabic word-internal vowel sequence is significantly *different* from a post-lexical vowel sequence, as seen in Figure 4.

Figure 4. Durational contrast between V#(C)V and V.V

	duration	N		duration	N	p-value
a.V_i # (C)V_i	153 ms	32	V_i .V_i	214 ms	39	$p < .001$
b.V_i # (C)V_j	186 ms	38	V_i .V_j	211 ms	58	$p < .001$

This information is summarized in (33).

(33) *Summary of phonetic evidence*
monosyllabic vs. bisyllabic word-internal vowel sequence:
VV (word-internal) and V.V (word-internal) are *different.*
word-internal vs. phrasal vowel duration:
VV (word-internal) and V#V (imperative) are *the same.*
V.V (word-internal) and V#V (imperative) are *different.*

The crucial point is that postlexical vowel sequences must fuse into one syllable. Duration evidence thus constitutes strong evidence that syllabification is distinctive at the word level as determined by the presence of empty Cs. And at the phrasal level, vowel segments re-syllabify as single syllables, demonstrating that the empty Cs have been deleted.

5. Theoretical Implications

The phonological and phonetic data provided above demonstrate that there is an empty root node present underlyingly, which affects word-level phonological rules, but deletes at a certain point in the grammar. This can be handled straightforwardly in a derivational approach, as will be demonstrated next. An OT analysis, however, is forced to deal with the question of deletion and extrinsic ordering.

5.1 *A Derivational Account of Empty Cs*

Let us first look at the derivational account of Kikamba. Three levels are needed to handle the data: the stem level, the word level, and the phrase level. (34) provides the order of the processes discussed in this paper. Notice that the split between those processes sensitive to the empty C versus those blind to it is handled very easily: processes sensitive to the empty C apply before empty C deletion and those blind to the empty C are ordered after deletion.

(34)	Stem	Syllabification/Vowel Coalescence
		Initial Shortening (stem)
	Word	Syllabification/Vowel Coalescence
		Excrescent consonant insertion (1sg.Object prefix /-N-/)
		Lateral Spread
		Empty C Deletion
		Prefix /k/-deletion (2sg.Object prefix, Infin. /ko/)
	Phrase	Syllabification/Vowel Coalescence
		Sequential Shortening
		Initial Shortening (phrasal)

We have already seen that the empty C is present at the stem level, but is deleted in the phrasal phonology. The presence of the empty C prevents stem level application of Initial shortening after the 1st singular Object prefix, as in (35a). But the empty C is ignored at the phrasal level, and therefore Initial shortening can apply to all phonetically vowel-initial verbs in the imperative, as seen in (35b).

(35)	a.	/ko-N-Cɔ́ɔt-eC-a/	[koondɔ́ɔtɛ.à]	'dream about me'
	b.	/Cɔ́ɔt-a/	[ɔ́ta]	'dream'

The interaction of empty C deletion with segmental rules allows us to pinpoint where the empty C deletes, namely, at the word level. This is seen in (36–37) with the segmental rules of excrescent consonant insertion, and prefix /k/-deletion. In (36a-b), the excrescent consonant associated with the 1st singular Object prefix /-N-/ is inserted prior to deletion of the empty C, which is precisely how the correct consonant is chosen. Insertion of the excrescent consonant takes place at the word level, not the stem level, since it is triggered by the pre-stem 1st singular Object prefix /-N-/. In contrast, prefixal /k/ deletes in (37a), but not in (37b), where the empty C has been deleted. This explains why prefix /k/ deletion is not applicable to the empty C-initial verb 'change' in (37b). After the empty C deletes in (37b), the initial syllable of the verb is onsetless, so Prefix-/k/ deletion may not apply. This rule too is a word-level lexical rule, as demonstrated above.

(36)	*1st SG OP form*		
		V-stem	*Empty C-stem*
		a. /ko-N-að-a/	b. /ko-N-Cɔ́ɔt-eC-a/
	Excrescent consonant insertion	ko-Nb-að-a	ko-Ndɔ́ɔt-eC-a

Empty C deletion	N/A	ko-ndɔ́ɔt-ɛ.-a
Prefix-/k/ deletion	N/A	N/A
	[koombaðà]	[koondɔ́ɔtɛ.à]
	'govern me'	'dream about me'

(37) *2nd SG OP form*

	C-stem	*Empty C-stem*
	a. /ko-ko-tál-a/	b. /ko-ko-Calyool-a/
Excrescent consonant insertion	N/A	N/A
Empty C deletion	N/A	koko.alyoola
Prefix-/k/ deletion	ko.otála	N/A
	[ko.ota̋la̋]	[koko.alyoòlà]
	'count you-sg'	'change you-sg'

Therefore, the segmental rules indicate that the empty C deletes in the middle of the word-level phonology.

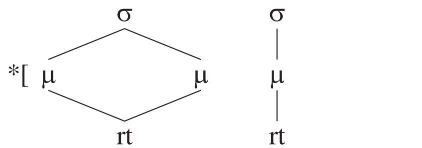
5.2 *An Optimality Account of Empty Cs*

We have seen how the devices of ordering and deletion can explain the range of phenomena related to empty Cs in Kikamba in a derivational account. We will now see that an attempt to handle this data in a version of the OT framework which is strictly monostratal and obeys containment runs into problems, indicating that such an approach will not work. However, by incorporating deletion as advocated in Correspondence Theory, and by adopting a multistratal account which emulates extrinsic ordering, the Kikamba data can be accounted for. Our focus here will be on the problems of Sequential shortening and Initial shortening.

5.2.1 *Sequential Shortening and Vowel Hiatus*

Let us first look at Sequential shortening. The OT solution to this phenomena must do two things at once. On the one hand, it must maintain the degree of phonological separateness that the empty C provides, so that the contrast between hiatus-maintenance and hiatus-resolution can be preserved. But, on the other hand, an OT account must also eliminate that contrast, so that shortening of long vowels before other vowels is not blocked by the empty C. Recall that Sequential shortening captures just that generalization: a long vowel may not precede another vowel. A formulation of this generalization as a constraint is seen in (38).

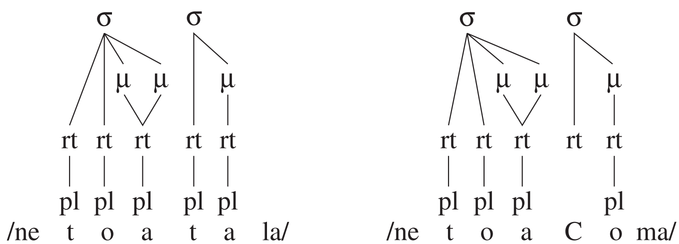
(38) *LONG VOWEL + ONSETLESS SYLLABLE



The question is: how can Sequential shortening be accomplished without deleting the empty C, and without introducing derivational levels. We will see that a monostratal constraint-based theory cannot do this. In a monostratal OT account, there is only one level, and therefore ordering cannot be implemented. Moreover, there is no deletion in the classical version of OT. Therefore, the two possibilities are that the empty C is either parsed or unparsed. We will see that either assumption is problematic.

Consider the possibility that an empty C *is* parsed. Compare the representations in (39):

(39)



a. [nétwáatálâ]
verb stem /-tál/ 'count'
'we just counted'

b. [nétwá.ómâ]
verb stem /-Cóm/ 'bite'
'we just bit'

In (39a), vowel coalescence does not take place between the tense/aspect prefix and the vowel of the verb stem 'count.' Since the tense/aspect marker /a/ is followed by the syllable /ta/, which has an onset, there is no motivation to fuse the syllables. Likewise in (39b), the verb stem 'bite' has an onset which is an empty C, and again there is no motivation for syllable fusion. While parsing the empty C explains why the vowel of 'bite' does not coalesce with a preceding vowel, we cannot explain why Sequential shortening occurs in (39b), but not (39a). With the parsed empty C as an onset, (39b) wouldn't in fact violate the constraint against a long vowel followed by a vowel, since the long vowel is not immediately followed by a vowel, and therefore, there is no motivation for Sequential shortening. Nonetheless, Sequential shortening occurs.

The alternative is that the empty C is *un*parsed, as in (40).

(40)	a.	/né-né-a-óm-aC-á/ [nénɔ́ɔ́ma.â] verb stem /-óm/ 'dry' 'I'm always dry'	b.	/né-né-a-⟨C⟩óm-aC-á/ [néná.óma.â] verb stem /-Cóm/ 'bite' 'I always bite'

If the empty C is not parsed, it is not problematic that Sequential shortening occurs in (40b). The long vowel /aa/, resulting from the coalescence of the subject prefix /ne/ and the tense/aspect prefix /a/, precedes an onsetless syllable, which is not allowed. Therefore in this context, the long /aa/ shortens. However, while Sequential shortening is explained, we are left wondering why coalescence doesn't occur in (40b) between the tense/aspect marker /a/ and the vowel of the verb stem 'bite.' In (40a), coalescence does occur with the verb stem, resulting in the bimoraic syllable [nɔɔ]. The problem is that leaving the empty C unparsed leaves the verb stem 'bite' with no onset, and thus undermines the whole explanation for the various contrasting syllable fusion effects of the language. With an unparsed empty C, it is a mystery that coalescence occurs with the verb stem in (40a), but is blocked in (40b).

To summarize, we see that a parsing paradox results in a monostratal version of OT. It appears that the empty root node needs to be parsed to explain vowel hiatus, but unparsed to explain Sequential shortening, which is not possible. We need a multistratal account to explain the fact that there is a level at which the empty C is present, and a level at which it is not. The level at which it is present chooses as the optimal candidate one in which the underlined vowels in (41a) are not coalesced into a single syllable. The level at which the empty C is not present chooses as the optimal candidate the one where an underlying long vowel surfaces as short before an onsetless syllable, as in (41b).

(41)			(a)		(b)
			output of 1st level		*output of 2nd level*
	/né-né-a-Cóm-aC-á/	→	nénáa óma.á	→	néná.óma.â
	/né-tó-a-Cóm-á/	→	nétwáa ómá	→	nétwá.ómâ

Sequential shortening represents our first example of the need for deletion, and the notion of before and after in an account of Kikamba empty C phenomena.

5.2.2 *Initial Shortening*

5.2.2.1 *A Monostratal Account* ***Without*** *Derivational Residue*

We will now examine problems with Initial shortening. A monostratal account of Initial shortening presents the same problem as seen for Sequential shortening: there is no way for the empty C to be simultaneously parsed and unparsed. Furthermore, a monostratal account of Initial shortening has the additional problem that the correct form of vowel-initial verbs cannot be generated.

First consider the generalization that we want to capture with Initial shortening. Syllable-initial long vowels are prohibited, so the initial mora of a long vowel is not parsed if the syllable is onsetless. This will be referred to as the NoLongVowel constraint, seen in (42):

(42) NoLongVowel (NLV)

*[σ μ μ

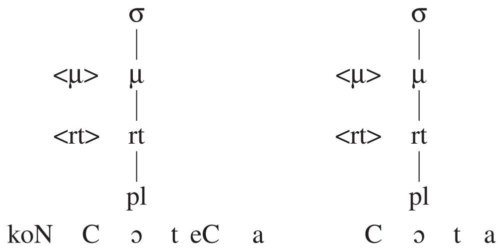
rt

The empty C initial verbs present a problem for a monostratal view because they have an onset at the stem level, but not at the phrase level. The paradox is that the empty C must be parsed to explain why the stem-initial vowel remains long after the 1st singular Object prefix in (43a). Parsing the empty C allows the verb to escape stem-level IS. But this empty C must also be unparsed to explain why the stem-initial vowel shortens at the word level in the imperative in (43b).

(43)	a.	/ko-N-Cɔ́ɔt-eC-a/ infin-1sg-stem-appl-infl	[koondɔ́ɔtɛ.à]	‘dream about me’
	b.	/Cɔ́ɔt-a/ stem-infl	[ɔ́ta]	‘dream!’

Suppose we leave the empty C unparsed. This would allow the long vowel to be shortened in the imperative. However, this would leave us unable to explain why the stem vowel is long when preceded by the 1st singular Object prefix, as demonstrated in (44).

(44)

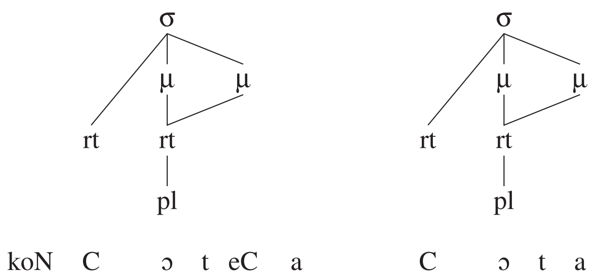


a. *[koondɔ́tɛ.à]
'dream about me'

b. [ɔ́ta]
'dream'

The second possibility is that the empty C is parsed. Exactly the same problem arises. In (45), we see that parsing the empty C does not predict that the imperative will surface with a short vowel, because with a parsed empty C, the initial vowel is not in an onsetless syllable, and therefore should not be subject to Initial shortening.

(45)



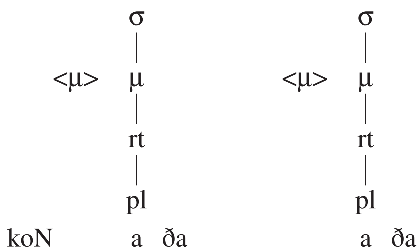
a. [koondɔ́ɔtɛ.à]
'dream about me'

b. *[ɔ́ɔta]
'dream'

The crucial observation is that the empty root node is parsed at the stem level, but is unparsed at the phrase level. A parsing paradox results for the imperative of empty C initial verbs. Since the same segment is both at the beginning of a stem and a word, it is impossible for it to be both parsed and unparsed simultaneously. Again, a monostratal account cannot explain this fact.

The second problem for a monostratal account is posed by vowel-initial verbs. Recall that Initial shortening is a reflection of the ban on long, onsetless syllables. Now consider the data from the long vowel-initial verb 'govern' in (46), where the long, onsetless syllable shortens in both (46a), after the 1st singular Object prefix, and in (46b), the imperative.

(46)



a. [koombaðà] 'govern me' b. [aðâ] 'govern!'

In both of these contexts the stem vowel surfaces as short. The problem is that, in (46a), the stem-initial syllable actually has an onset on the surface, which is provided by the excrescent consonant /b/. Therefore, in a monostratal account, the constraint prohibiting onsetless long vowels at the stem level becomes technically irrelevant to this form and there is no motivation for vowel shortening. Given these problems I conclude that, since a monostratal account in a version of OT with containment cannot employ deletion and is unable to handle the distinction between stem and phrase level, it cannot handle the facts of Kikamba.

5.2.2.2 *A Multistratal Account* **With** *Derivational Residue*

In order to handle the Kikamba data in a constraint-based theory, then, we need to assume that there can be more than one level in the grammar. Although the ideal would be for a nonderivational theory to be monostratal, we have seen that such an approach cannot handle the facts of Kikamba. A multistratal analysis allows the postulation of levels that can be ordered with respect to the deletion of elements at the beginning or end of a level. Therefore, following the McCarthy and Prince (1993b) analysis of Axininca Campa, as well as recent suggestions by Kenstowicz (1994) and Buckley (1994b), and as was argued by Duanmu (this volume), we will assume a multistratal version of OT. We need the empty C to be present, and will delete it later at the interface of levels.

We will focus on two issues. The first is the actual deletion of the empty C. It has been suggested by McCarthy and Prince (1993b) for Axininca Campa that unparsed elements are automatically deleted at the end of a given level if they remain unparsed (Prince and Smolensky 1993). In order to incorporate empty C deletion into a constraint-based account of Kikamba, we need to take this a step further and allow parsed elements without phonetic content to be deleted at some point. However, the difference for Kikamba is that empty Cs do not delete at the end of *every* level. We will see that they delete at the end of one, specified level. This type of deletion is impossible in the 1993 version of OT, which constrains Gen from deleting elements. However, there is no problem with deletion in recent versions of OT, which abandon the principle of containment. In Correspondence Theory, the constraint MAX generally requires all elements in the input to be present in the output. Competing with this constraint is a constraint we call *SILENCE, which prohibits segments that lack phonetic output, as seen in (47).

(47) MAX
Everything present in the input must be present in the output.
*SILENCE
Segments may not lack phonetic content.

In strata where MAX dominates *SILENCE, empty Cs will be preserved, and due to the ONS constraint, will be parsed as syllable onsets. See (48).

(48) *Level*$_i$*: Empty C is present at the stem level*
MAX, ONS » *SILENCE

/Cɔɔt-a/	MAX	ONS	*SILENCE
ɔta	*!	*	
☞Cɔɔta			*

In strata where *SILENCE dominates MAX, the syllabified empty C must be removed to satisfy *SILENCE, as seen in (49).

(49) *Level*$_j$*: Empty C is not present at the phrase level*
*SILENCE » MAX, ONS

σ σ /|\ /| /Cɔɔ t-a/	*SILENCE	MAX	ONS
☞ɔta		*	*
Cɔɔta	*!		

By reranking the constraints at different levels, we are able to get deletion of the empty C in a constraint-based account.

The other piece of derivational residue is ordering itself, which forces the adoption of a multistratal approach. Let us first review the ordering of the rules needed in a derivational account in (50).

(50) *Rules*

Level 1 – stem	Initial Shortening of onsetless syllables
Level 2 – word	Excrescent consonant insertion (insert /d/ if there is an empty C, otherwise, insert /b/)
	Empty C deletion
	Prefix-/k/ Deletion (2sOP, infinitive prefix)
Level 3 – phrase	Initial Shortening of onsetless syllables

Having levels allows the possibility that an element can be present and parsed at one level and can be deleted at a later level. As seen in (48,49), this is accomplished by different rankings of MAX and *SILENCE. The prediction of this approach is that, throughout a given level, a deleted segment behaves consistently as if it is present or as if it is not present. However, if we look again at (50), we see that empty C deletion is in the middle of the word level. Therefore, a constraint-based approach is forced to bifurcate the word level into two separate word levels in order to encode the distinction between phenomena which respect the empty C versus those which ignore the empty C. This is represented in (51).

(51)

STEM	ONS, MAX » *SILENCE
WORD LEVEL 1	ONS, MAX » *SILENCE
WORD LEVEL 2	*SILENCE » ONS, MAX
PHRASE	*SILENCE » ONS, MAX

The purpose of positing these two levels is to get the effect of extrinsic rule ordering: this is the only place the empty C can delete to generate the correct forms. In one level, MAX is ranked higher than *SILENCE; in the next level where there is deletion, MAX is ranked lower.

This split in the word level does not correspond to a natural morphological split. In fact, it forces different Object prefixes to be assigned to different levels, since the 1st singular Object prefix phonology takes place in WORD LEVEL1 where empty Cs are retained, but the reduction of the 2nd singular Object prefix /ko/ takes place in WORD LEVEL 2, where empty Cs have been deleted. Since the sole purpose of positing these levels is to get the effect that empty C deletion is ordered with respect to other principles of the language, we see that a constraint-based theory is forced to posit levels that are not independently justified in the language.

6. Conclusion

Kikamba provides both phonological and phonetic evidence for empty root nodes. These empty underlying nodes sometimes behave as though they are present and sometimes behave as though they are not. This data can be easily handled in a derivational account, as shown above: the empty C deletes after some rules, and before others.

Kikamba data indicates that ordering and deletion are crucial to accounting for the vowel hiatus facts. Correspondence Theory allows there to be deletion of the empty C, but still leaves us with the result that levels are postulated for the sole purpose of providing a level for the empty C to delete. If we allow levels that are not independently justified in order to obtain the correct rankings of constraints at different points in the grammar, then the ordering requirements of Kikamba can be incorporated into a constraint-based theory. However, it is critical to point out that allowing a proliferation of levels such as this constitutes a weakening of the theory. The potential remains that there could be as many levels as phonological rules, which would make a constraint-based theory look more and more like a notational variant of a derivational account. It is clear that these data raise questions about how ordering can be incorporated into a multistratal constraint-based theory; it further raises the question of how we can constrain a constraint-based account from incorporating otherwise unnecessary levels.

Acknowledgments

I would like to thank Beatrice Mulala, a speaker of the Machakos dialect of Kikamba, from whom the data for this paper was gathered. This research was in part supported by NSF grant SBR-9421362.

References

Papers marked [ROA] can be found at the Rutgers Optimality Archive on the Internet: http://ruccs.rutgers.edu/roa/

Abu-Salim, Issam M. 1980. Epenthesis and Geminate Consonants in Palestinian Arabic. *Studies in the Linguistic Sciences* 10.2: 1–11.

Alderete, John. 1995a. Winnebago Accent and Dorsey's Law. In: Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.) *Papers in Optimality Theory. University of Massachusetts Occasional Papers in Linguistics*.

———. 1995b. Faithfulness to Prosodic Heads. Manuscript, University of Massachusetts, Amherst.

———. 1996. Prosodic Faithfulness in Cupeño. Manuscript, University of Massachusetts, Amherst.

Al-Mozainy, Hamza Q., Robert Bley-Vroman and John McCarthy. 1985. Stress Shift and Metrical Structure. *LI* 16: 135–144.

Anderson, John. 1987. The Limits of Linearity. In: J. Anderson and J. Durand (eds.) *Explorations in Dependency Phonology*. Dordrecht: Foris.

Anderson, S. 1985. *Phonology in the Twentieth Century.* Chicago: University of Chicago Press.

———. 1992. *A-morphous Morphology*. Cambridge: Cambridge University Press.

Archangeli, Diana. 1991. Syllabification and Prosodic Templates in Yawelmani. *NLLT* 9: 231–283.

Archangeli, Diana, and Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambridge, Massachusetts: MIT Press.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, Mass: The MIT Press.

———. 1994. *Morphology by Itself*. Cambridge, Mass: The MIT Press.

Ashton, E. O. 1944. *Swahili Grammar (Including Intonation).* Burnt Hill, Harlow: Longman Group.

Becker, T. 1993. Back-Formation, Cross-Formation and 'Bracketing Paradoxes' in Paradigmatic Morphology. In: Geert Booij and Jaap van Marle (eds.) *Yearbook of Morphology*. Dordrecht: Kluwer, 1–26.

Beckman, Jill. 1995. Shona Height Harmony: Markedness and Positional Identity. In: Jill Beckman, Suzanne Urbanczyk and Laura Walsh Dickey (eds.) *Papers in Optimality Theory. University of Massachusetts Occasional Papers in Linguistics* 18. Amherst Graduate Linguistic Student Association.

———. 1998. Positional Faithfulness. PhD Dissertation, University of Massachusetts, Amherst.

Beckman, Mary E., Maria G. Swora, Jane Rauschenberg and Kenneth DeJong. 1990. Stress Shift, Stress Clash, and Polysyllabic Shortening in Prosodically Annotated Discourse. *Proceedings of the 1990 International Conference on Spoken Language Processing* 1:5–80.

Benua, Laura. 1995. Identity Effects in Morphological Truncation. In: Jill Beckman, Suzanne Urbanczyk and Laura Walsh Dickey (eds.) *Papers in Optimality Theory. University of Massachusetts Occasional Papers in Linguistics* 18. Amherst, Massachusetts: Graduate Linguistic Student Association.

———. 1997. Transderivational Identity: Phonological Relations between Words. PhD Dissertation, University of Massachusetts, Amherst.

Benua, Laura, and John McCarthy. In progress. Prosodic Faithfulness and Prosodic Circumscription. Manuscript, University of Maryland, College Park, and University of Massachusetts at Amherst.

Berko, Jean. 1958. The Child's Learning of English Morphology. *Word* 14:150–177.

Bird, Steven. 1990. Constraint-Based Phonology. PhD Dissertation, University of Edinburgh.

———. 1995. *Computational Phonology*. Cambridge, UK: Cambridge University Press.

Birk, D.B.W. 1976. *The MalakMalak Language, Daly River (Western Arnhem Land)*, (Pacific Linguistics B45) Canberra: Australian National University.

Blevins, Juliette. 1995. The Syllable in Phonological Theory. In: John Goldsmith (ed.), 206–244.

Bochner, H. 1993. *Simplicity in Generative Morphology*. Berlin: Mouton de Gruyter.

Boersma, Paul. 1998. Functional Phonology; Formalizing the Interactions between Articulatory and Perceptual Drives. PhD Dissertation, University of Amsterdam.

Booij, Geert. 1997. Non-Derivational Phonology Meets Lexical Phonology. In: I. Roca (ed.) *Derivations and Constraints in Phonology.* Oxford: Clarendon Press.

Bradshaw, Mary. 1995. Tone on Verbs in Suma. In: A. Akinlabi (ed.) *Theoretical Approaches to African Linguistics.* Trenton: Africa World Press, 255–272.

Brame, Michael. 1974. The Cycle in Phonology: Stress in Palestinian, Maltese, and Spanish. *LI* 5:39–60.

Broselow, Ellen. 1982. On the Interaction of Stress and Epenthesis. *Glossa* 16: 115–132.

———. 1992. Parametric Variation in Arabic Dialect Phonology. In: E. Broselow, M. Eid and J. McCarthy (eds.) *Perspectives on Arabic Linguistics IV. Papers from the Fourth Annual Symposium on Arabic Linguistics*. Amsterdam-Philadelphia: John Benjamins.

Browman, C.P. 1978. Tip of the Tongue and Slip of the Ear. Implications for Language Processing. *UCLA Working Papers in Phonetics* 42: 1–149.

Brown, R., and D. McNeill. 1966. The "Tip of the Tongue" Phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5: 325–337.

Buckley, Eugene. 1991. Kashaya Closed-Syllable Shortening and Prosodic Syllabification. *WCCFL* 10: 65–74.

———. 1994a. *Theoretical Aspects of Kashaya Phonology and Morphology*. Stanford: CSLI Publications.

———. 1994b. Persistent and Cumulative Extrametricality in Kashaya. *NLLT* 12: 423–464.

———. 1995a. Constraint Domains in Kashaya. *WCCFL* 14:47–61.

———. 1995b. Constraint Domains in Optimality Theory. Generative Linguists of the Old World, Workshop on Constraints in Phonology. Tromsø, Norway, June 3.

———. 1996. Levels vs. Domains: The Case of Kashaya Vowel Length. BLS 22, 36–45.

———. 1997. Optimal Iambs in Kashaya. Rivista di Linguistica 9, 1–44.

Burzio, Luigi. 1994. *Principles of English Stress*. Cambridge, UK: Cambridge University Press.

Butterworth, B. 1989. Lexical Access in Speech Production. In W. Marslen-Wilson (ed.) *Lexical Representation and Process.* Cambridge, MA: MIT Press, 108–135.

Cassimjee, F. 1995. Isixhoso Tonology: An Optimal Domains Theory Analysis. Manuscript, University of Illinois.

Chao, Yuen-Ren. 1930. A System of Tone Letters. *Le maître phonétique* 45:24–27.

———. 1968. *A Grammar of Spoken Chinese.* Berkeley, California: University of California Press.

Chen, Matthew Y. 1986. The Paradox of Tianjin Tone Sandhi. *Proceedings of the Chicago Linguistic Society Meeting* 22:98–114.

———. 1987. Introductory Remarks to a Symposium on Tianjin Tone Sandhi. *Journal of Chinese Linguistics* 15:203–227

———. In press. *Tone Sandhi*.

Cheng, Chin-Chuan. 1973. A Synchronic Phonology of Mandarin Chinese. Monographs on Linguistic Analysis no. 4. The Hague: Mouton.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English.* New York: Harper and Row.

Chomsky, Noam, Morris Halle and Frank Lukoff. 1956. On Accent and Juncture in English. In: *For Roman Jakobson.* The Hague: Mouton.

Clements, G.N. 1976. Vowel Harmony in Nonlinear Phonology: An Autosegmental Model. Manuscript. Distributed by the Indiana University Linguistics Club 1980.

———. 1985. The Geometry of Phonological Features. *Phonology Yearbook* 2: 220–252.

———. 1990. The Role of the Sonority Cycle in Core Syllabification. In: J. Kingston and M. Beckman (eds.) *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech.* Cambridge, UK: Cambridge University Press, 283–333.

———. 1991. Place of Articulation in Consonants and Vowels: A Unified Approach. *Working Papers of the Cornell Phonetics Laboratory* 5:124–152. Ithaca, New York: Cornell University.

Clifton, C., L. Frazier and K. Rayner (eds). 1994. *Perspectives on Sentence Processing*. Hillsdale, NJ: Erlbaum.

Cohn, Abigail. 1989. Stress in Indonesian and Bracketing Paradoxes. *NLLT* 7:167–216.

Cohn, Abigail, and John McCarthy. 1994. Alignment and Parallelism in Indonesian Phonology. Manuscript, Cornell University and University of Massachusetts, Amherst.

Cole, Jennifer 1990. Arguing for the Phonological Cycle: A Critical Review. In: D. Meyer, S. Tomioka and L. Zidani-Eroglu (eds.) *Proceedings of the Formal Linguistics Society of Midamerica*. Madison: Linguistics Student Organization, University of Wisconsin, 51–67.

———. 1995. The Cycle in Phonology. In: J. Goldsmith (ed.) *The Handbook of Phonological Theory.* Cambridge, UK: Basil Blackwell. *Blackwell Handbooks in Linguistics* 1.

Cole, Jennifer, and John Coleman. 1992. No Need for Cyclicity within Generative Grammar. In: J. Denton, G.Chan and C. Canakis (eds.) *CLS* 28: *Parasession on the Cycle in Linguistic Theory*. Chicago: CLS, 36–50.

Coleman, John. 1991. Phonological Representations. PhD Dissertation, University of York.

———. 1995. Declarative Lexical Phonology. In: J. Durand and F. Katamba (eds.) *Frontiers of Phonology: Atoms, Structures, Derivations*. London: Longman, 333–383.

Cooper, W.E., and S.J. Eady. 1986. Metrical Phonology in Speech Production. *Journal of Memory and Language* 25:369–384.

Crompton, A. 1982. Syllables and Segments in Speech Production. In: A. Cutler (ed.) *Slips of the Tongue and Language Production.* Berlin: Mouton.

Crowhurst, Megan. 1991. Minimality and Foot Structure in Metrical Phonology and Prosodic Morphology. PhD Dissertation, University of Arizona, Tucson. Distributed by Indiana University Linguistics Club, 1993.

Crowhurst, Megan, and Mark S. Hewitt. 1995. Directional Footing, Degeneracy, and Alignment. *NELS* 25, 47–61.

Cutler, Anne. 1986. Phonological Structure in Speech Recognition. *Phonology Yearbook* 3: 161–178.

Cutler, Anne, and Norris, Dennis. 1988. The Role of Strong Syllables in Segmentation for Lexical Access. *Journal of Experimental Psychology: Human Perception and Performance* 14.1.

Dai, John X.L. 1992. Chinese Morphology and its Interface with the Syntax. PhD Dissertation, Ohio State University, Columbus.

Davis, Stuart. 1995. Some Matters Regarding the Derivational Residue. Paper presented at the Conference on the Derivational Residue in Phonology, Tilburg October 5–7, 1995.

Dikken, Marcel den, and Harry van der Hulst. 1988. Segmental Hierarchitecture. In: H. van der Hulst and N. Smith (eds.) *Features, Segmental Structure and Harmony Processes. Volume I.* Dordrecht: Foris, 1–78.

Dixon, Robert M. W. 1977. *A Grammar of Yidin*^{y}. Cambridge, UK: Cambridge University Press.

Dobrin, Lise M. 1993. On the Prosodic Integrity of Underlying Forms: In Support of a Word-Based Lexicon. Paper presented at LASSO XXII.

Dolbey, A. 1996. Output Optimization and Cyclic Allomorph Selection. *WCCFL.*

Downing, Laura. 1990. Local and Metrical Tone Shift in Nguni. *Studies in African Linguistics.*

Dresher, Elan, and Aditi Lahiri. 1991. The Germanic Foot: Metrical Coherence in Old English. *LI* 22:251–286.

Duanmu, San. 1993. Rime Length, Stress, and Association Domains. *Journal of East Asian Linguistics* 2:1–44.

———. 1995. Metrical and Tonal Phonology of Compounds in Two Chinese Dialects. *Language* 71.2:225–259.

———. 1997a. Recursive Constraint Evaluation in Optimality Theory: Evidence from Cyclic Compounds in Shanghai. *NLLT* 15:465–507.

———. 1997b. Wordhood in Chinese. In: Jerome Packard (ed.) *New Approaches to Chinese Word Formation: Morphology, Phonology and the Lexicon in Modern and Ancient Chinese*. Berlin: Mouton de Gruyter, 135–196.

Ernout, A. 1953. *Morphologie historique du latin*. Librairie C. Klincksieck, Paris.

Fan, Jiyan. 1958. Xing-ming zuhe jian 'de' zi de yufa zuoyong [The Grammatical Function of 'de' Adjective-Noun Constructions]. *Zhongguo Yuwen* 1958.5:213–217.

Farwaneh, Samira. 1995. Directional Effects in Arabic Dialect Syllable Structure. PhD Dissertation, University of Utah.

Fay, D., and A. Cutler. 1977. Malapropisms and the Structure of the Mental Lexicon. *LI* 8:505–520.

Ferguson, Charles. 1976. "Short A" in Philadelphia English. In M.E. Smith (ed.), *Studies in Linguistics: In Honor of George L. Trager.* The Hague: Mouton, 250–274.

Fillmore, C., and P. Kay. in progress. Construction Grammar. Manuscript, UC Berkeley.

Fleisch, H. 1963. Le parler arabe de Kfar-ṣghāb. *Bulletin d'Etudes Orientales, Institut Francais de Damas* 18:95–125.

Flemming, Edward. 1993. The Role of Metrical Structure in Segmental Rules. MA thesis, UCLA.

———. 1995. Auditory Representations in Phonology. PhD Dissertation, UCLA.

Foley, W.A. 1991. *The Yimas Language of New Guinea.* Stanford: Stanford University Press.

Fromkin, V.A. (ed.) 1973. *Speech Errors as Linguistic Evidence*. The Hague: Mouton.

——— . 1985. Evidence in Linguistics. In: Siew-Yue Killingley (ed.) *Linguistics and Linguistic Evidence*. Newcastle upon Tyne: Grefatt & Grevatt, 18–38.

———. 1987. The Lexicon: Evidence from Acquired Dyslexia. *Language* 63.1:1–22.

Fujimura, O., and J.B. Lovins. 1978. Syllables as Concatenative Units. In: A. Bell and J. B. Hooper (eds.) *Syllables and Sounds.* Amsterdam: North-Holland.

Fukui, Naoki. 1996. On the Nature of Economy in Language. *Cognitive Studies* 3:51–71.

Gafos, A. 1996. The Articulatory Basis of Locality in Phonology. PhD Dissertation, Johns Hopkins University.

Gardiner, J. M., F. I. Craik and F. A. Bleasdale. 1973. Retrieval Difficulty and Subsequent Recall. *Memory and Cognition 1*: 213–216.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum and Ivan Sag. 1985. *Generalized Phrase-Structure Grammar*. Cambridge, MA.: Harvard University Press.

Gnanadesikan, Amalia. 1995. Markedness and Faithfulness Constraints in Child Phonology. Manuscript, University of Massachusetts, Amherst. Available as ROA-67.

Goldsmith, John. 1976. Autosegmental Phonology. PhD Dissertation, MIT.

———. 1989. Licensing, Inalterability and Harmonic Rule Application. In: C. Wiltshire, R. Graczyk and B. Music (eds.) *Papers from CLS 25.* Chicago: CLS, 145–156.

———. 1990. *Autosegmental and Metrical Phonology*. Oxford: Blackwell.

———. 1993a. Harmonic Phonology. In: Goldsmith (ed.) 1993b, 21–60.

———. (ed.) 1993b. *The Last Phonological Rule*. Chicago: Chicago University Press.

Golston, Chris. 1991. Minimal Word, Minimal Affix. *NELS 21.*

———. 1991. Both Lexicons. PhD Dissertation, UCLA.

———. 1995. Representation as Pure Markedness. Manuscript, Heinrich Heine University. [Available as ‘Direct Optimality Theory’, ROA–71].

———. 1996a. Direct Optimality Theory: Representation as Pure Markedness. *Language* 72:713–748.

———. 1996b. Prosodic Constraints on Roots and Words. In: Ursula Kleinhenz (ed.) *Studia Grammatica: Interfaces in Phonology*. Berlin: Akademie Verlag. 172–193.

———. 1998. Feature Geometry is Syllable Geometry. *Annual Meeting of the Linguistics Society of America,* New York City.

Golston, Chris. and Wolfgang Kehrein. 1998. Mazatec Onsets and Nuclei. *International Journal of American Linguistics* 64.4:311–337.

———. 1999. A Prosodic Theory of Laryngeal Contrasts. Paper presented at the *Holland Institute of Linguistics Phonology Conference 4.*

Green, Thomas. 1995. The Stress Window in Piraha: A Reanalysis of Rhythm in Optimality Theory. Manuscript, MIT.

Green, Thomas, and Michael Kenstowicz. 1995. The Lapse Constraint. *FLSM* 6.

Hale, Kenneth. 1973. Deep-Surface Canonical Disparities in Relation to Analysis and Change. In: Thomas Sebeok (ed.) *Current Trends in Linguistics* 11:401–458.

Hall, Tracy A. 1994. Extrasyllabicity and Resyllabification. In: R. Wiese (ed.) *Recent Developments in Lexical Phonology*. Düsseldorf: Heinrich-Heine Universität, 151–66.

Halle, Morris, and William Idsardi. 1994. A Reanalysis of Indonesian Stress. *MIT Working Papers in Linguistics* 21:1–9.

Halle, Morris, and Michael Kenstowicz. 1991. The Free Element Condition and Cyclic versus Noncyclic Stress. *LI* 22:457–501.

Halle, Morris, and Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge, MA.: MIT Press.

———. 1987b. Stress and the Cycle. *LI* 18:45–84.

Hargus, S. 1993. Modeling the Phonology-Morphology Interface. In: S. Hargus and E. Kaisse (eds.) *Studies in Lexical Phonology.* San Diego, Cal: Academic Press, 45–74.

Harris, James. 1970. A Note on Spanish Plural Formation. *Language* 46: 928–930.

———. 1977. Remarks on Diphthongization in Spanish. *Lingua* 41: 261–305.

Hayes, Bruce. 1980. A Metrical Theory of Stress Rules. PhD Dissertation, MIT, Cambridge, Massachusetts. Published 1985, New York: Garland Press.

———. 1982. Metrical Structure as the Organizing Principle of Yidin^y Phonology. In: H. van der Hulst and N. Smith (eds.) *The Structure of Phonological Representations, Part I*. Dordrecht: Foris Publications.

———. 1985. Iambic and Trochaic Rhythm in Stress Rules. *BLS* 11:429–446.

———. 1989. Compensatory Lengthening in Moraic Phonology. *LI* 20:253–306.

———. 1995a. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

———. 1995b. On What to Teach the Undergraduates: Some Changing Orthodoxies in Phonological Theory. In: Ik-Hwan Lee (ed.) *Linguistics in the Morning Calm 3*. Seoul: Hanshin, 59–77.

———. 1999. Phonetically-Driven Phonology: The Role of Optimality Theory and Inductive Grounding. In: Michael Darnelly, Edith Moravcik, Michael Noonan, Frederich Newmeyer, and Kathleen Wheathy (eds.) *Functionalism and Formalism in Linguistics. Volume I: General Papers.* Amsterdam: John Benjamins, 243–285.

Hewitt, Mark. 1994. Deconstructing Foot Binarity in Koniag Alutiiq. Manuscript, University of British Columbia. Available as ROA 12.

Hirst, D. 1985. Linearization and the Single Segment Hypothesis. In: J. Guéron and H.G. Obenhauer (eds.) *Grammatical Representations*. Dordrecht: Foris, 87–100.

Hockett, C. 1954. Two Models of Grammatical Description. *Word* 10:210–231.

Hooper, Joan. 1976. *An Introduction to Natural Generative Phonology*. New York: Academic Press.

Hulst, Harry van der. 1984. *Syllable Structure and Stress in Dutch*. Dordrecht: Foris

———. 1990. The Book of Segments. Manuscript, Leiden University

———. 1994. Radical CV Phonology: The Locational Gesture. *UCL Working Papers in Linguistics* 6, 439–477.

———. 1995. Radical CV Phonology: The Categorial Gesture. In: J. Durand and F. Katamba (eds.) *Frontiers of Phonology,* 80–116. Essex: Longman.

———. 1996. Radical CV Phonology: The Segment — Syllable Connection. In: J. Durand and B. Laks (eds.) *Current trends in Phonology: Models and methods*. Vol 1. CNRS/ESRI Paris X, 333–363.

———. 1999. Features, Segments and Syllables in Radical CV Phonology. In: J. Rennison (ed.) *Phonologica 1996: Syllables!?* The Hague: Holland Academic Graphics.

———. In prep. Principles of Radical CV Phonology. Manuscript, Leiden University.

Hulst, Harry van der and Marian Klamer. 1996. The Uneven Trochee. To appear in *HIL Proceedings* 2. The Hague: Holland Academic Graphics

Hulst, Harry van der, and Norval Smith. 1988. The Variety of Pitch Accent Systems. In: Harry van der Hulst and Norval Smith (eds.) *Autosegmental Studies in Pitch Accent.* Dordrecht: Foris.

Hume, Elisabeth and David Odden 1996. Reconsidering [Consonantal]. *Phonology* 13: 345–376.

Hung, Tony. 1987. Tianjin Tone Sandhi: Towards a Unified Approach. *Journal of Chinese Linguistics* 15:274–305.

Hyman, Larry M., and Ngessimo Mutaka. 1990. Syllable and Morpheme Constraints in Kinande reduplication. *Phonology* 7:73–119.

Hyman, Larry M., and Russell G. Schuh 1974. Universals of Tone Rules. *LI* 5:81–115.

Ikawa, Hajime. 1995. On Stress Assignment, Vowel-Lengthening, and Epenthetic Vowels in Mohawk and Selayarese: Some Theoretical Implications. Manuscript, University of California, Irvine.

Inkelas, Sharon. 1990. *Prosodic Constituency in the Lexicon*. New York: Garland.

———. 1993a. Nimboran Position Class Morphology. *NLLT* 11:559–624.

———. 1993b. Deriving Cyclicity. In: S. Hargus and E. Kaisse (eds.) *Studies in Lexical Phonology.* San Diego, Cal: Academic Press, 75–110.

———. 1994a. Lexicon Optimization and Underspecification Theory. *NELS* 24.

———. 1994b. Exceptional Stress-Attracting Suffixes in Turkish. In: R. Kager (ed.) *Proceedings of a Workshop on Prosodic Morphology*. Berlin: Mouton de Gruyter.

Inkelas, Sharon, and Orhan Orgun. 1995a. Level Ordering and Economy in the Lexical Phonology of Turkish. *Language* 763–93.

———. 1995b. Level (Non)Ordering in Recursive Morphology: Evidence from Turkish. To appear in Lapointe, ed.

Irshied, Omar, and Michael Kenstowicz. 1984. Some Phonological Rules of Bani-Hassan Arabic: A Bedouin Dialect. *Studies in the Linguistic Sciences* 14.1:109–147.

Itô, Junko. 1986. Syllable Theory in Prosodic Phonology. PhD Dissertation, University of Massachusetts, Amherst.

———. 1989. A Prosodic Theory of Epenthesis. *NLLT* 7:217–59.

Itô, J., and J. Hankamer. 1989. Notes on Monosyllabism in Turkish. *PASC 1*: 61–69.

Itô, Junko, Yoshihisa Kitagawa and Armin Mester. 1996. Prosodic Faithfulness and Correspondence: Evidence from a Japanese Argot. (Appeared in JEAL 5 [Also: ROA 146]

Itô, Junko, and Armin Mester. 1995. Japanese Phonology. In: John Goldsmith (ed.) *The Handbook of Phonological Theory*. Oxford: Basil Blackwell, 817–838.

Iverson, G.K., and J.C. Salmons. 1995. Aspiration and Laryngeal Representation in Germanic. *Phonology* 12.3:369–396.

Jensen, S. 1994. Is ʔ an Element?: Towards a Non-segmental Phonology. *Working Papers in Linguistics and Phonetics,* 4, 71–78. Department of Linguistics, SOAS, UK.

Jespersen, Otto. 1904. *Lehrbuch der Phonetik*. Leipzig and Berlin.
Johnson, D. 1972. *Formal Aspects of Phonological Description*. The Hague: Mouton.
Jones, Daniel, and Dennis Ward. 1969. *The Phonetics of Russian.* Cambridge, UK: Cambridge University Press.
Kager, René. 1993. Alternatives to the Iambic-Trochaic Law. *NLLT* 11:381–432.
———. 1994. Ternary Stress in Alignment Theory. Manuscript, Utrecht University.
———. 1997. Rhythmic Vowel Deletion in Optimality Theory. In: I. Roca (ed.) *Derivations and Constraints in Phonology*. Oxford: Oxford University Press, 463–499
———. Forthcoming. Stem Stress and Peak Correspondence in Dutch. To appear in J. Dekkers, F. van der Leeuw and J. van de Weijer (eds.) *Optimality Theory: Phonology, Syntax and Acquisition.* Oxford: Oxford University Press.
Kaisse, Ellen, and Sharon Hargus. 1993. *Studies in Lexical Phonology*. San Diego: Academic Press.
Kaplan, Ronald M., and Martin Kay. 1981. Phonological Rules and Finite State Transducers. *Linguistic Society of America Meeting Handbook*. Fifty-Sixth Annual Meeting, December 27–30, 1991. New York: Linguistic Society of America.
———. 1994. Regular Models of Phonological Systems. *Computational Linguistics* 20:3.
Karttunen, Lauri 1993. Finite-State Constraints. In: John Goldsmith (ed.) 1993b, 173–194.
Kay, Martin. 1983. *Unification Grammar*. Technical report. Palo Alto, CA: Xerox.
Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud. 1985. The Internal Structure of Phonological Elements: A Theory of Charm and Government. *Phonology Yearbook* 2: 303–326.
———. 1990. Constituent Structure and Government in Phonology. *Phonology* 7.2: 193–232.
Kayne, Richard S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.
———. 1994. *The Antisymmetry of Syntax.* Cambridge, Massachusetts: MIT Press.
Kehrein, W. 1998. Phonology without Affricates. Manuscript, Universität Marburg.
Kelly, M.H., and J.K. Bock. 1988. Stress in Time. *Journal of Experimental Psychology: Human Perception and Performance* 14:389–403.
Kennedy, Chris. 1994. Morphological Alignment and Head Projection. In: J. Merchant et al. (eds.) *Phonology at Santa Cruz* 3:47–64.
Kenstowicz, Michael. 1983. Parametric Variation and Accent in the Arabic Dialects. *CLS* 19: 205–213.
———. 1994. Cyclic vs. Noncyclic Constraint Evaluation. *MIT Working Papers in Linguistics* 21:9–37. (Also appeared with the same title in *Phonology* 12:397–436.)
———. 1996. Base-Identity and Uniform Exponence: Alternatives to Cyclicity. In: J. Durand and B. Laks (eds.) *Current Trends in Phonology: Models and Methods*. Paris: CNRS, 363–393. Also appeared 1996 as European Studies Research Institute, University of Salford Publications. [Also: ROA 103].
Kenstowicz, Michael, and Kamal Abdul-Karim. 1980. Cyclic Stress in Levantine Arabic. *Studies in the Linguistic Sciences* 10.2:55–76.

Kenstowicz, Michael, and Charles Kisseberth. 1977. *Topics in Phonological Theory*. New York: Academic Press.

———. 1979. *Generative Phonology: Description and Theory.* New York: Academic Press.

Keyser, S.J., and K.N. Stevens. 1995. Feature Geometry and the Vocal Tract. *Phonology* 11.2:207–236.

Kiparsky, Paul. 1982a. Lexical Morphology and Phonology. In: I.-S. Yang (ed.) *Linguistics in the Morning Calm*. Seoul: Hanshin, Linguistics Society of Korea, 3–91.

———. 1982b. From Cyclic Phonology to Lexical Phonology. In: H. van der Hulst and N. Smith (eds.) *The Structure of Phonological Representations*, *Part 2*. Dordrecht: Foris.

———. 1983. Word Formation and the Lexicon. In: F. Ingeman (ed.) *Proceedings of the Mid-America Linguistics Conference*. Lawrence: University of Kansas, 3–29.

———. 1985. Some Consequences of Lexical Phonology. *Phonology* 2:85–138.

———. 1988. Phonological Change. In: F. Newmeyer (ed.), *Linguistics: The Cambridge Survey.* Vol. I. Cambridge, UK: Cambridge University Press, 363–415.

———. 1993. Blocking in Non-Derived Environments. In: S. Hargus and E. Kaisse (eds.) *Studies in Lexical Phonology* (Phonetics and Phonology 4). San Diego: Academic Press, 277–313.

———. 1995a. The Phonological Basis of Sound Change. In: J. Goldsmith (ed.), 640–670.

———. 1995b. *Livonian Stød*. Handout of presentation at workshop on Diachrony and Prosodic Theory, University of Nijmegen, September 6, 1995.

———. 1998. Paradigm Effects and Opacity. Manuscript, Stanford.

Kirchner, Robert. 1993. Optimizing Yidiɲ Phonology. Manuscript, UCLA.

———. 1995. Going the Distance: Synchronic Chain Shifts in Optimality Theory. Manuscript, UCLA [also available as ROA-66].

Kisseberth, Charles. 1970. On the Functional Unity of Phonological Rules. *LI* 1:291–306.

Koenig, J.-P., and D. Jurafsky. 1994. Type Underspecification and On-Line Type Construction in the Lexicon. *WCCFL* 13.

Kohn, S.E. et al. 1987. Lexical Retrieval: The Tip of the Tongue Phenomenon. *Applied Psycholinguistics* 8:245–266.

Koriat, A., and I. Lieblich. 1974. What Does a Person in a "TOT" State Know that a Person in a "Don't Know" State Doesn't Know. *Memory & Cognition* 2:647–655.

———. 1977. A Study of Memory Pointers. *Acta Psychologica* 41:151–164.

Koskenniemi, Kimmo. 1983. *Two-level Morphology: a General Computational Model for Word-form Recognition and Production*. Publication no. 11, Department of General Linguistics. University of Helsinki.

Krause, S. 1979. Topics in Chuckchee Phonology and Morphology. PhD Dissertation, University of Illinois at Urbana-Champaign.

Labov, W. 1981. Resolving the Neogrammarian Controversy. *Language* 57:267–308.

———. 1993. *Principles of Language Change.* Oxford: Basil Blackwell.

Lakoff, George. 1993. Cognitive Phonology. In: John Goldsmith (ed.) 1993b, 117–145.

Lapointe, S. (ed.) 1996. *Proceedings of the Davis Conference on Morphology and its Interactions with Phonology and Syntax*. Stanford: CSLI Publications.

Leben, W. 1973. Suprasegmentals. PhD Dissertation, MIT.

Levelt, W.J.M. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Levin, Juliette. 1985. A Metrical Theory of Syllabicity. PhD Dissertation, MIT.

Li, Xingjian, and Sixun Liu. 1985. Tianjin fangyan de liandu biandiao [Tone Sandhi in the Tianjin Dialect]. *Zhongguo Yuwen* 1:76–80.

Lynch, John D. 1978. A *Grammar of Lenakel.* Pacific Linguistics B55. Canberra: Australian National University.

Marantz, Alec. 1982. Re: Reduplication. *LI* 13:435–482.

McCarthy, John. 1979a. Formal Problems in Semitic Phonology and Morphology. PhD dissertation, MIT.

———. 1979. On Stress and Syllabification. *LI* 10:443–465.

———. 1980. The Role of the Evaluation Metric in the Acquisition of Phonology. In: C.L. Baker and John McCarthy (eds.) *The Logical Problem of Language Acquisition.* Cambridge, Massachusetts: MIT Press.

———. 1981. A Prosodic Theory of Nonconcatenative Morphology. *LI* 12:373–418.

———. 1988. Feature Geometry and Dependency: A Review. *Phonetica* 45:84–108.

———. 1993. The Parallel Advantage: Containment, Consistency and Alignment. Paper presented at the Rutgers Optimality Workshop, October 1993, Rutgers University.

———. 1994. On Coronal 'Transparency'. *TREND II.* Stanford University.

———. 1996a. Remarks on Phonological Opacity in Optimality Theory. In: J. Lecarme, J. Lowenstamm and U. Shlonsky (eds.) *Proceedings of the Second Colloquium on Afro-Asiatic Linguistics* (Sofia Antipolis, June 1994).

———. 1996b. Faithfulness in Prosodic Morphology and Phonology: Rotuman Revisited. To appear in *NLLT*.

———. 1997. Process-Specific Constraints in Optimality Theory. *LI* 28:231–251.

———. 1998. Sympathy and Phonological Opacity. Manuscript: University of Massachusetts, Amherst.

———. To appear. Faithfulness and Prosodic Circumscription. To appear in Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds.) *Optimality Theory: Phonology, Syntax and Acquisition.* Oxford: Oxford University Press.

McCarthy, John, and Alan Prince. 1986. Prosodic Morphology. Manuscript, University of Massachusetts, Amherst and Brandeis University.

———. 1990. Foot and Word in Prosodic Morphology: The Arabic Broken Plural. *NLLT* 8:209–282.

———. 1993a. Generalized Alignment. In: Geert Booij and Jaap van Marle (eds.) *Yearbook of Morphology* 1993. Dordrecht: Kluwer, 79–153.

———. 1993b. Prosodic Morphology I: Constraint Interaction and Satisfaction. Manuscript, University of Massachussetts and Rutgers University.

———. 1994. Prosodic Morphology: An Overview. Talks presented at the PTS/HIL Workshop on Prosodic Morphology, University of Utrecht.

———. 1995. Faithfulness and Reduplicative Identity. In: Jill Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.) *Papers in Optimality Theory. University of Massachusetts Occasional Papers in Linguistics* 18:249–384. [Also: ROA 60].

Mester, Armin, and Junko Itô. 1989. Feature Predictability and Underspecification. *Language* 65:258–93.

Mester, Armin, and Jaye Padgett. 1993. Directional Syllabification in Generalized Alignment. ROW-1.

Meyers, Antje S. 1990. The Time Course of Phonological Encoding in Language Production: The Encoding of Successive Syllables of a Word. *Journal of Meaning and Language* 29:524–45.

———. 1991. The Time Course of Phonological Encoding in Language Production: Phonological Encoding Inside a Syllable. *Journal of Meaning and Language* 30:69–89.

Michelson, Karen. 1981. Stress, Epenthesis and Syllable Structure in Mohawk. In: N. Clements (ed.) *Harvard Studies in Phonology* Vol. II, pt.2.

———. 1988. *A Comparative Study of Lake Iroquoian Accent.* Dordrecht: Kluwer.

Mithun, Marianne, and Hasan Basri. 1985. The Phonology of Selayarese. *Oceanic Linguistics* 25.1–2:210–254.

Mohanan, K.P. 1982. Lexical Phonology. PhD Dissertation, MIT.

———. 1986. *The Theory of Lexical Phonology.* Dordrecht: Reidel.

Myers, Scott. 1987. Vowel Shortening in English. *NLLT* 5.4:485–518.

———. 1994. OCP effects in Optimality Theory. Manuscript. [Available as ROA-6]

Nash, David. 1979.Yidin^{y} Stress: A Metrical Account. *NELS 9/CUNYForum* 7–8: 112–130.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.

Noske, Roland. 1992. A Theory of Syllabification and Segmental Alternation; With Studies on the Phonology of French, German, Tonkawa and Yawelmani. PhD Dissertation, Tilburg University.

Odden, David. 1993. Interaction between Modules in Lexical Phonology. In: Kaisse and Hargus (eds.) 111–144.

———. 1996. Patterns of Reduplication in Kikerewe. *OSU Working Papers in Linguistics*, 111–148.

———. 1998. Principles of Tone Assignment in Tanzanian Yao. In: Larry Hyman and Charles Kisseberth (eds.) *Theoretical Aspects of Bantu Tone.* Stanford: CSLI Publications, 265–314.

Öhman, Sven. 1966. Coarticulation in VCV Utterances: Spectrographic Measurements. *Journal of the Acoustical Society of America* 39:151–168.

Orgun, Orhan. 1994a. A Declarative Theory of Phonology-Morphology Interleaving. Handout Phonology Workshop. University of California, Berkeley.

———. 1994b. Monotonic Cyclicity and Optimality Theory. *NELS* 24. 461–474.

———.1994c. Monotonic Cyclicity and Suspended Affixation. Colloquium. Edinburgh University.

———. 1995a. Flat vs. Branching Morphological Structures: The Case of Suspended Affixation. *BLS* 21.

———. 1995b. Reference to Internal Morphological Structure in Sign-Based Morphology. Manuscript, UC Berkeley.

———. 1995c. Response to Stump (1995). Manuscript, UC Berkeley.

———. 1995d. Suspended Affixation: A New Look at the Phonology-Morphology Interface. In: M. Bierwisch (ed.) *Proceedings of the Conference on Interfaces in Phonology*.

———. 1996. Sign-Based Morphology: A Declarative Theory of Phonology-Morphology Interleaving. PhD Dissertation, University of California, San Diego.

Orgun, Orhan, and Sharon Inkelas. 1992. Turkish Prosodic Minimality. Handout Sixth International Conference on Turkish Linguistics. Anadolu University, Eskiehir, Turkey.

Orgun, Orhan, J.-P. Koenig and D. Jurafsky. 1996. Constraint-based Morphology. 3rd International Conference on HPSG. Marseilles, France.

Oswalt, Robert L. 1961. A Kashaya Grammar (Southwestern Pomo). PhD Dissertation, UC Berkeley.

Padgett, Jaye. 1991. Stricture in Feature Geometry. PhD Dissertation, University of Massachusetts, Amherst.

———. 1995. Review of John Goldsmith (ed.) 1993b. *Phonology* 12:147–155.

Paradis, Carole. 1988. On Constraints and Repair Strategies. *The Linguistic Review* 6: 71–97.

Pater, Joe. 1996. Minimal Violation and Phonological Development. *Language Acquisition* 6:201–253.

———. In preparation. On the Nonuniformity of Weight-to-stress and Stress Preservation Effects in English. Manuscript, McGill University. [ROA-107].

Pesetsky, David. 1979. Russian Morphology and Lexical Theory. Manuscript, MIT.

Piggott, G.L. 1995. Epenthesis and Syllable Weight. *NLLT* 13.2:283–326.

Poletto, R. 1998. Constraints on Tonal Association in Olusiama: An Optimality Theoretic Account. In: Larry Hyman and Charles Kisseberth (eds.) *Theoretical Aspects of Bantu Tone.* Stanford: CSLAI, 331–364.

Polgárdi, Krisztina. 1998. Vowel Harmony; An Account in Terms of Government and Optimality. PhD Dissertation, Leiden University.

Pollard, C., and I. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Stanford: CSLI Publications.

Polomé, Edgar C. 1967. *Swahili Language Handbook.* Washington D.C.: Center for Applied Linguistics.

Potter, Brian. 1994. Serial Optimality in Mohawk Prosody. *Proceedings of CLS* 30, University of Chicago.

Priller, J., and E. Mittenecker. 1988. Experiment zum Unterschied von “Wort auf der Zunge” und “Gefühl des Wissens”. *Zeitschrift für experimentelle und angewandte Psychologie* 35:129–146.

Prince, Alan. 1983. Relating to the Grid. *LI* 14.1:19–100.

———. 1990. Quantitative Consequences of Rhythmic Organization. *Proceedings of CLS; Parasession on the Syllable in Phonetics and Phonology*. Chicago: Chicago Linguistic Society, 355–98.

———. 1993. Minimal Violation. Paper presented at ROW, Rutgers University.

Prince, Alan, and Paul Smolensky. 1991. Optimality. Paper presented at the Fourth Arizona Phonology Conference, Tucson, April 5–6.

———. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Manuscript, Dept. of Linguistics, Rutgers University, New Brunswick, NJ and Program in Cognitive Science, Johns Hopkins University, Baltimore.

Pritchett, B.L. 1992. *Grammatical Competence and Parsing Performance*. Chicago: University of Chicago Press.

Pulleyblank, D. 1983. Tone in Lexical Phonology. PhD Dissertation, MIT.

———. 1988.. Underspecification, the Feature Hierarchy and Tiv Vowels. *Phonology* 5, 299–326.

Reason, J., and D. Lucas. 1984. Using Cognitive Diaries to Investigate Naturally Occurring Memory Blocks. In: J. E. Harris and P. E. Morris (eds.) *Everyday Memory, Actions and Absent-Mindedness.* London: Academic Press.

Rice, K. 1988. Continuant Voicing in Slave. In: M. Hammond and M. Noonan (eds.) *Theoretical Morphology*. San Diego: Academic Press, 371–88.

———. 1989. *A Grammar of Slave*. Berlin: Mouton de Gruyter.

Riehemann, S. 1994. Morphology and the Hierarchical Lexicon. Manuscript. Stanford.

Roberts-Kohno, R. Ruth. 1995. Vowel Coalescence and Hiatus in Kikamba. In: Akinbiyi Akinlabi (ed.) *Theoretical Approaches to African Linguistics*. Trenton: Africa World Press.

———. 1998. Empty Root Nodes in Kikamba: Conflicting Evidence and Theoretical Implications. In: Ian Maddicson and Thomas J. Hinnebush (eds.) *Language History and Linguistic Description in Africa*. Trenton: Africa World Press.

Roca, Iggy. (ed.). 1997. *Derivations and Constraints in Phonology.* Oxford: Clarendon Press.

Rosenthall, Sam. 1994. Vowel/Glide Alternations in a Theory of Constraint Interaction. PhD Dissertation, University of Massachusetts, Amherst.

Rubin, D.C. 1975. Within Word Structure in the Tip-of-the-Tongue Phenomenon. *Journal of Verbal Learning and Verbal Behavior* 14:392–397.

Sagey, Elisabeth. 1986. *The Representation of Features and Relations in Non-Linear Phonology*. New York: Garland.

Schane, Sanford. 1968. *French Phonology and Morphology*. Cambridge, Massachusetts: MIT Press.

Schuh, Russell. 1972. Rule Inversion in Chadic. *Studies in African Linguistics* 3:379–397.

Scobbie, James 1991. Attribute Value Phonology. PhD Dissertation, University of Edinburgh.

———. 1993. Constraint Violation and Conflict from the Perspective of Declarative Phonology. *Canadian Journal of Linguistics* 38: 155–68.

Selkirk, Elisabeth. 1981. Epenthesis and Degenerate Syllables in Cairene Arabic. In: H. Borer and J. Aoun (eds.) *Theoretical Issues in the Grammar of Semitic Languages*. Cambridge: MIT Working Papers in Linguistics, 111–140.

———. 1984. On the Major Class Features and Syllable Theory. In: Mark Aronoff and Richard Oehrle (eds.) *Language Sound Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students.* Cambridge, Massachusetts: MIT Press.

Selkirk, Elisabeth, and Tong Shen. 1990. Prosodic Domains in Shanghai Chinese. In Sharon Inkelas and Draga Zec (eds.) *The Phonology-Syntax Connection*. Stanford, Cal: CSLI/Stanford University,. Distributed by University of Chicago Press, 313–337.

Shaw, Patricia. 1976. Theoretical Issues in Dakota Phonology and Morphology. PhD Dissertation, University of Toronto.

———. 1985. Modularism and Substantive Constraints in Dakota Lexical Phonology. *Phonology Yearbook* 2:173–202.

Shieber, S. 1986. *Introduction to Unification-Based Approaches to Grammar*. Stanford: CSLI Publications.

Shih, C-L. 1986. The Prosodic Domain of Tone Sandhi in Chinese. PhD Dissertation, UC San Diego.

Sibanda, Galen. 1997. Reduplication of Imbricated Verb Forms in Ndebele. Paper presented at the 28th Annual Conference on African Linguistics. Cornell University, July 11–13.

Sietsema, Brian. 1988. Reduplication in Dakota. *CLS* 24:337–352.

Sievers, E. 1881. *Grundzüge der Phonetik*. Leipzig: Breitkopf and Hartel.

Smolensky, Paul. 1993. Harmony, Markedness and Phonological Activity. ROA 87.

———. 1996. The Initial State and 'Richness of the Base' in Optimality Theory. http://www.cogsci.jhu.edu/TechReports. [Also: ROA-154].

Sproat, R. 1985. On Deriving the Lexicon. PhD Dissertation, MIT.

———. 1992. *Morphology and Computation*. Cambridge, Mass: MIT Press.

Sproat, Richard, and Chilin Shih. 1991. The Cross-Linguistic Distribution of Adjective Ordering Restrictions. In: Carol Georgopoulos and Roberta Ishihara (eds.) *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda*. Dordrecht: Kluwer Academic Publishers.

Stampe, David. 1973. A Dissertation on Natural Phonology. PhD Dissertation, University of Chicago. Distributed 1979 by Indiana University Linguistics Club, Bloomington.

Steriade, Donca. 1982. Greek Prosodies and the Nature of Syllabification. PhD Dissertation, MIT.

———. 1988. Reduplication and Syllable Transfer in Sanskrit and Elsewhere. *Phonology* 5:73–155.

———. 1992. Closure, Release and Nasal Contours. In: M. Huffman and R. Krakow (eds.) *Nasality*. San Diego: Academic Press.
———. 1993. Segments, Contours and Clusters. In: A. Crochetière, J.-C. Boulanger and C. Ouellon (eds.) *Proceedings of the XVth International Congress of Linguists*. Sainte-Foy: Les presses de l'université Laval.
———. 1994a. Complex Onsets as Single Segments: The Mazateco Pattern. In: J. Cole and C. Kisseberth (eds.) *Perspectives in Phonology,* 201–291
———. 1994b. Class notes from UCLA seminar on Lexical Phonology. UCLA.
———. 1995. Underspecification and Markedness. In: John Goldsmith (ed.).
———. 1996. Paradigm Uniformity and the Phonetics-Phonology Boundary. Paper presented at the 5th Conference in Laboratory Phonology. Evanston, Illinois.
Stump, G. 1996. Comments on Inkelas and Orgun (1995b). In Lapointe (ed.) 1996.
Tan, Fu. 1987. Tone Sandhi in the Tianjin Dialect. *Journal of Chinese Linguistics* 15:228–246.
Tanenhaus, Michael K., and John C. Trueswell. 1995. Sentence Comprehension. To appear in Joanne Miller and Peter Eimas (eds), *Handbook of Perception and Cognition*, volume 11. San Diego: Academic Press.
Tesar, Bruce B., and Paul Smolensky. 1993. The Learnability of Optimality Theory: An Algorithm and Some Basic Complexity Results. ROA 2.
———. 1996. Learnability in Optimality Theory. ROA 110.
Trommelen, Mieke 1983. *The Syllable in Dutch*. Dordrecht: Foris.
Urbanczyk, Suzanne. 1995. Double Reduplications in Parallel. In: J. Beckman, S. Urbanczyk and L. Walsh (eds.) *Papers in Optimality Theory. University of Massachusetts Occasional Papers in Linguistics* 18. Amherst, Massachusetts: Graduate Linguistic Student Association, 499–531.
Vennemann, Theo. 1972. Rule Inversion. *Lingua* 29:209–242.
Xu, Baohua, Zhenzhu Tang, Rujie You, Nairong Qian, Rujie Shi and Yaming Shen. 1988. *Shanghai Shiqü fangyan zhi* [Urban Shanghai Dialects]. Shanghai: Shanghai Jiaoyu Chubanshe.
Yarmey, A. D. 1973. I Recognize Your Face but I Can't Remember Your Name: Further Evidence on the Tip-of-the-Tongue Phenomenon. Memory and *Cognition* 1, 287–290.
Yip, Moira. 1989. Contour Tones. *Phonology* 6:149–174.
———. 1996. Feet, Tonal Reduction and Speech Rate at the Word and Phrase Level in Chinese. To appear in: Harry van der Hulst, René Kager and Wim Zonneveld (eds.) *Phrasal Phonology*. Dordrecht: Foris.
Zec, Draga. 1994. Footed Tones and Tonal Feet: Rhythmic Constituency in a Pitch Accent Language. Manuscript, Cornell University.
Zhang, Hongming. 1992. Topics in Chinese Phrasal Phonology. PhD Dissertation, University of California, San Diego.

Zimmer, K. 1975. Some Thoughts on Likely Phonologies for Non-Ideal Speakers. In: R. Grossman et al. (eds.) *Papers from the Parasession on Functionalism*. Chicago: Chicago Linguistics Society, 556–567.

Zubritskaya, Katya. 1995. The Categorial and Variable Phonology of Russian. PhD Dissertation, University of Pennsylvania.

Language Index

Name Index

Subject Index

In the series *Linguistik Aktuell/Linguistics Today* the following titles have been published thus far or are scheduled for publication:

15 **ROHRBACHER, Bernhard Wolfgang:** Morphology-Driven Syntax. A theory of V to I raising and pro-drop. 1999. viii, 296 pp.
16 **LIU, Feng-Hsi:** Scope and Specificity. 1997. viii, 187 pp.
17 **BEERMAN, Dorothee, David LEBLANC and Henk van RIEMSDIJK (eds.):** Rightward Movement. 1997. vi, 410 pp.
18 **ALEXIADOU, Artemis:** Adverb Placement. A case study in antisymmetric syntax. .. 1997. x, 256 pp.
19 **JOSEFSSON, Gunlög:** Minimal Words in a Minimal Syntax. Word formation in Swedish. 1998. ix, 199 pp.
20 **LAENZLINGER, Christopher:** Comparative Studies in Word Order Variation. Adverbs, pronouns, and clause structure in Romance and Germanic. 1998. x, 371 pp.
21 **KLEIN, Henny:** Adverbs of Degree in Dutch and Related Languages. 1998. x, 232 pp.
22 **ALEXIADOU, Artemis and Chris WILDER (eds.):** Possessors, Predicates and Movement in the Determiner Phrase. 1998. vi, 388 pp.
23 **GIANNAKIDOU, Anastasia:** Polarity Sensitivity as (Non)Veridical Dependency. 1998. xvi, 282 pp.
24 **REBUSCHI, Georges and Laurice TULLER (eds.):** The Grammar of Focus. 1999. vi, 366 pp.
25 **FELSER, Claudia:** Verbal Complement Clauses. A minimalist study of direct perception constructions. 1999. xiv, 278 pp.
26 **ACKEMA, Peter:** Issues in Morphosyntax. 1999. viii, 310 pp.
27 **RŮŽIČKA, Rudolf:** Control in Grammar and Pragmatics. A cross-linguistic study. 1999. x, 206 pp.
28 **HERMANS, Ben and Marc van OOSTENDORP (eds.):** The Derivational Residue in Phonological Optimality Theory. 2000. viii, 322 pp.
29 **MIYAMOTO, Tadao:** The Light Verb Construction in Japanese. The role of the verbal noun. 2000. xiv, 232 pp.
30 **BEUKEMA, Frits and Marcel den DIKKEN (eds.):** Clitic Phenomena in European Languages. 2000. x, 324 pp.
31 **SVENONIUS, Peter (ed.):** The Derivation of VO and OV. 2000. vi, 372 pp.
32 **ALEXIADOU, Artemis, Paul LAW, Andre MEINUNGER and Chris WILDER (eds.):** The Syntax of Relative Clauses. 2000. vi, 397 pp.
33 **PUSKÁS, Genoveva:** Word Order in Hungarian. The syntax of Ā-positions. 2000. xvi, 398 pp.
34 **REULAND, Eric J. (ed.):** Arguments and Case. Explaining Burzio's Generalization. 2000. xii, 255 pp.
35 **HRÓARSDÓTTIR, Thorbjörg:** Word Order Change in Icelandic. From OV to VO. 2001. xiv, 385 pp.
36 **GERLACH, Birgit and Janet GRIJZENHOUT (eds.):** Clitics in Phonology, Morphology and Syntax. 2001. xii, 441 pp.
37 **LUTZ, Uli, Gereon MÜLLER and Arnim von STECHOW (eds.):** Wh-Scope Marking. 2000. vi, 483 pp.
38 **MEINUNGER, Andre:** Syntactic Aspects of Topic and Comment. 2000. xii, 247 pp.
39 **GELDEREN, Elly van:** A History of English Reflexive Pronouns. Person, *Self*, and Interpretability. 2000. xiv, 279 pp.
40 **HOEKSEMA, Jack, Hotze RULLMANN, Víctor SÁNCHEZ-VALENCIA and Ton van der WOUDEN (eds.):** Perspectives on Negation and Polarity Items. 2001. xii, 368 pp.
41 **ZELLER, Jochen:** Particle Verbs and Local Domains. 2001. xii, 325 pp.
42 **ALEXIADOU, Artemis:** Functional Structure in Nominals. Nominalization and ergativity. 2001. x, 233 pp.
43 **FEATHERSTON, Sam:** Empty Categories in Sentence Processing. 2001. xvi, 279 pp.
44 **TAYLAN, Eser Erguvanlı (ed.):** The Verb in Turkish. 2002. xviii, 267 pp.
45 **ABRAHAM, Werner and C. Jan-Wouter ZWART (eds.):** Issues in Formal German(ic) Typology. 2002. xviii, 336 pp.

46 **PANAGIOTIDIS, Phoevos:** Pronouns, Clitics and Empty Nouns. 'Pronominality' and licensing in syntax. 2002. x, 214 pp.

47 **BARBIERS, Sjef, Frits BEUKEMA and Wim van der WURFF (eds.):** Modality and its Interaction with the Verbal System. 2002. x, 290 pp.

48 **ALEXIADOU, Artemis, Elena ANAGNOSTOPOULOU, Sjef BARBIERS and Hans-Martin GÄRTNER (eds.):** Dimensions of Movement. From features to remnants. 2002. vi, 345 pp.

49 **ALEXIADOU, Artemis (ed.):** Theoretical Approaches to Universals. 2002. viii, 319 pp.

50 **STEINBACH, Markus:** Middle Voice. A comparative study in the syntax-semantics interface of German. 2002. xii, 340 pp.

51 **GERLACH, Birgit:** Clitics between Syntax and Lexicon. 2002. xii, 282 pp.

52 **SIMON, Horst J. and Heike WIESE (eds.):** Pronouns – Grammar and Representation. 2002. xii, 294 pp.

53 **ZWART, C. Jan-Wouter and Werner ABRAHAM (eds.):** Studies in Comparative Germanic Syntax. Proceedings from the 15th Workshop on Comparative Germanic Syntax (Groningen, May 26–27, 2000). 2002. xiv, 407 pp.

54 **BAPTISTA, Marlyse:** The Syntax of Cape Verdean Creole. The Sotavento varieties. 2003. xxii, 294 pp. (incl. CD-rom).

55 **COENE, Martine and Yves D'HULST (eds.):** From NP to DP. Volume 1: The syntax and semantics of noun phrases. 2003. vi, 362 pp.

56 **COENE, Martine and Yves D'HULST (eds.):** From NP to DP. Volume 2: The expression of possession in noun phrases. 2003. x, 295 pp.

57 **DI SCIULLO, Anna Maria (ed.):** Asymmetry in Grammar. Volume 1: Syntax and semantics. 2003. vi, 405 pp.

58 **DI SCIULLO, Anna Maria (ed.):** Asymmetry in Grammar. Volume 2: Morphology, phonology, acquisition. 2003. vi, 309 pp.

59 **DEHÉ, Nicole:** Particle Verbs in English. Syntax, information structure and intonation. 2002. xii, 305 pp.

60 **TRIPS, Carola:** From OV to VO in Early Middle English. 2002. xiv, 359 pp.

61 **SCHWABE, Kerstin and Susanne WINKLER (eds.):** The Interfaces. Deriving and interpreting omitted structures. 2003. vi, 403 pp.

62 **CARNIE, Andrew, Heidi HARLEY and MaryAnn WILLIE (eds.):** Formal Approaches to Function in Grammar. In honor of Eloise Jelinek. 2003. xii, 378 pp.

63 **BOECKX, Cedric:** Islands and Chains. Resumption as stranding. 2003. xii, 224 pp.

64 **BOECKX, Cedric and Kleanthes K. GROHMANN (eds.):** Multiple *Wh*-Fronting. 2003. x, 292 pp.

65 **MANNINEN, Satu Helena:** Small Phrase Layers. A study of Finnish Manner Adverbials. 2003. xii, 275 pp.

66 **GROHMANN, Kleanthes K.:** Prolific Domains. On the Anti-Locality of movement dependencies. 2003. xvi, 372 pp.

67 **MIŠESKA TOMIĆ, Olga (ed.):** Balkan Syntax and Semantics. 2004. xvi, 499 pp.

68 **BREUL, Carsten:** Focus Structure in Generative Grammar. An integrated syntactic, semantic and intonational approach. 2004. x, 432 pp.

69 **KISS, Katalin É. and Henk van RIEMSDIJK (eds.):** Verb Clusters. A study of Hungarian, German and Dutch. 2004. vi, 514 pp.

70 **AUSTIN, Jennifer R., Stefan ENGELBERG and Gisa RAUH (eds.):** Adverbials. The interplay between meaning, context, and syntactic structure. 2004. x, 346 pp.

71 **GELDEREN, Elly van:** Grammaticalization as Economy. 2004. xvi, 320 pp.

72 **FUSS, Eric and Carola TRIPS (eds.):** Diachronic Clues to Synchronic Grammar. 2004. viii, 226 pp.

73 **CARNIE, Andrew, Heidi HARLEY and Sheila Ann DOOLEY-COLLBERG (eds.):** Verb First. On the syntax of verb initial languages. *Forthcoming*

74 **HEGGIE, Lorie and Fernando ORDONEZ (eds.):** Theoretical Perspectives on Clitic and Affix Combinations. *Expected Winter 04-05*