



Royal Netherlands Academy of Arts and Sciences (KNAW) KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN

Optimality Theory and Derivational Effects

van Oostendorp, M.; Hermans, B.J.H.

published in

The Derivational Residue in Phonological Optimality Theory
1999

document version

Publisher's PDF, also known as Version of record

[Link to publication in KNAW Research Portal](#)

citation for published version (APA)

van Oostendorp, M., & Hermans, B. J. H. (1999). Optimality Theory and Derivational Effects. In B. Hermans, & M. van Oostendorp (Eds.), *The Derivational Residue in Phonological Optimality Theory* John Benjamins Publishing.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

pure@knaw.nl

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.

Series Editor

Werner Abraham
Germanistisch Instituut
Rijksuniversiteit Groningen
Oude Kijk in 't Jatstraat 26
9712 EK Groningen
The Netherlands
E-mail: Abraham@let.rug.nl

Advisory Editorial Board

Guglielmo Cinque (University of Venice)
Günther Grewendorf (J.W. Goethe-University, Frankfurt)
Liliane Haegeman (University of Lille, France)
Hubert Haider (University of Salzburg)
Christer Platzack (University of Lund)
Ian Roberts (University of Stuttgart)
Ken Safir (Rutgers University, New Brunswick NJ)
Höskuldur Thráinsson (University of Iceland, Reykjavik)
Lisa deMena Travis (McGill University)
Sten Vikner (University of Stuttgart)
C. Jan-Wouter Zwart (University of Groningen)

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



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

© 1999 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

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

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 indissoluble 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 and 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):

$$(1) \quad \begin{array}{ll} \text{a.} & \text{Gen (In}_k\text{)} \quad \rightarrow \quad \text{Out}_{\text{inter}} = \{\text{Out}_1, \text{Out}_2, \dots\} \\ \text{b.} & \text{H-Eval (Out}_{\text{inter}}\text{)} \quad \rightarrow \quad \text{Out}_{\text{real}} \end{array}$$

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 $\text{Out}_{\text{inter}}$ 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 syllabification 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 and 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 under-application 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 successfully 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 of 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.

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 (/fhihm/) 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 /fhihm/ 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, undervived 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 [wag]?" and replies "[wagz].""

Although the classic concept of cyclicity seems to implement the idea behind the Wug test to some extent by providing a way to derive [wagz] from [wag], 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' — gindanu-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 to 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.