

Stratal OT: A synopsis and FAQs

Paul Kiparsky
Stanford University

1 Parallel, Transderivational, and Stratal OT

1.1 Theoretical choices

By modeling phonology as a system of ranked violable constraints, Optimality Theory (OT) succeeded in bringing substantive universals and typological generalizations to bear on the analysis of individual phonological systems, and uncovered important generalizations that escaped classical generative phonology, such as top-down effects and the emergence of the unmarked, to name just two (Prince and Smolensky 1993, 2004). Another fundamental principle of classic OT, that all constraints are evaluated in parallel on output representations, initially contributed much to the theory's conceptual appeal, but it soon became clear that the price for maintaining it is prohibitive. In order to account for phonology/phonology and phonology/morphology interactions under parallelism, numerous new computationally and learning-theoretically intractable constraint types had to be devised — Output-Output constraints, Paradigm Uniformity constraints, Base-Reduplication constraints, Sympathy constraints, Precedence constraints, among others. They vastly expanded the factorial typology and, separately or in any combination, failed to do the empirical job they were intended for.

Stratal OT returns to a pristine version of OT which countenances only markedness and faithfulness constraints. Instead of exploding the constraint typology, it deals with phonology/phonology and phonology/morphology interactions by organizing the grammar into strata (levels) analogous to those posited in Lexical Phonology and Morphology (LPM). Each stratum is a classic “pure” parallel OT

system, but the strata interface serially. Since the constraints at each stratum are limited to the well-understood markedness and faithfulness families of constraints regimented by Correspondence Theory, Stratal OT retains the major results of OT about factorial typology, and is formally clean like classical OT.

The specific arguments for Stratal OT fall into two types. The first is that Stratal OT is the best solution to the undergeneration problems of OT phonology, collectively known as the “derivational residue” (Roca 1997). They involve two common kinds of phenomena and a number of more exotic ones. The common kinds are OPACITY, the unexpected non-interaction of phonological processes, and CYCLICITY, the inheritance of phonological properties from bases to derivatives, also known as PARADIGMATIC TRANSFER EFFECTS, or SYNCHRONIC ANALOGY. These phenomena are briefly defined and illustrated below, and the analytic issues relating to them are explored at length in the chapters that follow. Their common feature is that they are on the face of it incompatible with parallel constraint evaluation, the central principle of OT. For the express purpose of dealing with opacity and paradigmatic effects, many types of constraints have been proposed which are not Markedness constraints and Input/Output Faithfulness constraints of the well-understood sort formalized in OT Correspondence Theory (McCarthy and Prince 1995). They include Sympathy constraints, PREC constraints (in OT-CC), Turbidity, Targeted constraints, Paradigm Uniformity constraints, and Output/Output Constraints. Their common feature is that they refer not just to the form under evaluation, but either to the steps by which has been derived (the CHAIN that maps inputs to outputs), or to some other input or output or derivation. Borrowing terms from early generative grammar, I will refer to the former as DERIVATIONAL CONSTRAINTS, and to the latter as TRANSDERIVATIONAL CONSTRAINTS, and to the enriched versions of OT that incorporate them are DERIVATIONAL OT and TRANSDERIVATIONAL OT, respectively. Derivational and transderivational constraints are too powerful in some ways and too weak in others, and that they lead to massive loss of generalizations. Stratal OT countenances no derivational or transderivational constraints, only the standard kinds of Markedness constraints and Input/Output Faithfulness constraints. Instead, it relies on level-ordering and principled (rather than process-specific) cyclic constraint evaluation. It yields a better understanding of opacity and cyclicity, capturing the range of occurring opacity and cyclic effects more accurately than any version of transderivational OT.

The second and perhaps more important type of argument is that Stratal OT limits the *overgeneration* of OT phonology. It contributes to the explanatory

goals of phonological theory by narrowing the typological space of constraints and constraint systems, by predicting the interactions between morphology and phonology, and by formally characterizing a lexical level of representation, whose linguistic significance is attested by convergent synchronic and diachronic evidence. This leads to new insights about sound change (Bermúdez-Otero, Kiparsky 2014a, 2014b), analogical change (Kiparsky to appear), and loanword adaptation (Kiparsky to appear), and solves classic OT's problems in dealing with phenomena such as compensatory lengthening (Kiparsky 2011). As always, arguments from explanatory adequacy are dependent on in-depth analyses, which cannot be adequately reproduced in the scope of this article.

In addition to its empirical superiority, Stratal OT is conceptually more attractive than transderivational OT because it recaptures some of the original simplicity of OT, by eliminating the special apparatus needed for handling phenomena that resist straightforward parallel OT. In this respect, Stratal OT completes the original OT program of eliminating such stipulative aspects as extrinsic rule ordering and derives the interaction and non-interaction of constraints from first principles.

Stratal OT is not LPM dressed up in OT costume. It is neither a graft of LPM onto OT, nor a graft of OT onto LPM. Nor is it some kind of compromise between them. It is more like a happy marriage. It combines the mutually compatible aspects of both theories, which complement each other because they deal with different things. LPM is primarily about the phonology-morphology interface, with consequences for interactions among phonological processes. Until the advent of constraint-based theories it was implemented in rule-based format by default.¹ But it is in no way intrinsically a rule-based theory. OT, on the other hand, is primarily about constraint interaction; its core ideas are that constraints are ranked and violable, and that violations are minimal. Parallelism has been a deep and fruitful guiding principle behind the development of OT, and giving it up, even in the limited and regimented way proposed here, is a serious move. But as has been repeatedly stressed by OT researchers, serial constraint evaluation is in principle perfectly compatible with the OT approach, and various types of it have been proposed and continue to be proposed by OT phonologists. Deriving serial effects from the stratal interface retains the desirable results of OT, including a restrictive factorial typology due to the integration of naturalness and markedness into phonological descriptions. Furthermore, while it is true that many cases of opacity and cyclicity can be treated well in ordered rule theory, that is by no

¹If constraints were sometimes invoked in LPM, it was in a merely empirical and ad hoc way, with no attempt to resolve the formal issues that arise when constraints and rules are mixed.

means true of all of them. Some provide quite compelling evidence in favor of Stratal OT over any ordered rule theory including LPM. The same is true for an even larger proportion of the second, explanatory type of evidence for Stratal OT.

1.2 Outline of Stratal OT

At the most general level, I will adopt the tenets in (1).

- (1) a. *Modularity*: Grammar is organized into components that interface via their input and output representations.
- b. *Optimality Theory*: Grammars are constituted by systems of ranked violable constraints.

Assumption (1a) is common ground in linguistics. For example, almost all researchers treat phonology and syntax as separate grammatical subsystems. Assumption (1b) is currently shared by a majority of phonologists, and by a substantial minority of syntacticians and semanticists. I will assume that syntax and semantics are constraint-based, just as phonology is. Nothing depends critically on that assumption, though it would be surprising if the components of grammar differed profoundly in their basic organizing principles.

The specific instantiation of this framework that I will be exploring, STRATAL OT, extends modularity within phonology and morphology.

- (2) a. *Stratification*: phonology and morphology are organized into STRATA (also known as LEVELS), each constituting a parallel constraint system.
- b. *Level-ordering*: each of the cross-categorical domains *stem*, *word*, *phrase* corresponds to a morphosyntactic and phonological stratum.
- c. *Cyclicity*: Stems and words must satisfy the applicable stem and word constraints at every stage.

These points are not unique to Stratal OT, but their combination is. (2a) has been assumed and defended in studies of morphology and vocabulary layering. The idea is that individual morphemes, classes of morphemes, morphological constructions such as reduplication, and vocabulary strata may be associated with

their own constraint rankings, or “cophonologies” (Inkelas, Orgun, & Zoll 1997, Itô and Mester 1995). A part of (2b) is sometimes implicitly or explicitly adopted in mainstream OT work in that lexical and postlexical phonology are treated as separate constraint systems, with the output of the former providing inputs to the latter. (2c) is a generalization of the generative phonology’s phonological cycle (Chomsky and Halle 1968) from the stem level, to which LPM assumed it was restricted, to the word level (Borowsky 1993, Harris & Kaisse 1999). The principle that morphology and phonology operate in tandem is of course reminiscent of “rule-to-rule” interpretation in Montague semantics.

The more specific claims concern the stratification of the lexicon, and the nature of the OT constraints.

(3) *Stratal OT*

- a. *Ranking*: The strata may differ in constraint ranking.
- b. *Correspondence theory*: Each stratum is a “pure” OT system comprising Input/Output constraints and markedness constraints; there are no transderivational constraints such as Output-Output constraints, Paradigm Uniformity constraints, Base-Reduplication constraints, Sympathy constraints, Precedence constraints, etc.

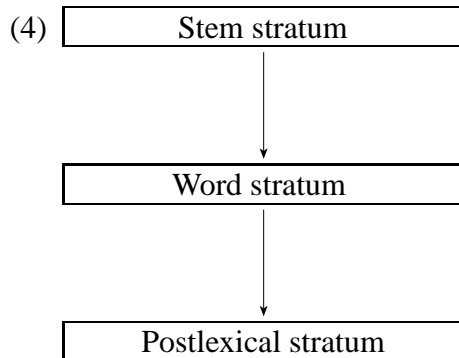
By Stratal OT I mean a theory which subscribes to (3) as well as to (1) and (2).

Stratal OT in turn can be implemented in several possible ways. Applying the theory requires commitment to a specific implementation, sometimes even beyond a point that can be empirically justified at present.

The version of Stratal OT explored here adopts Lexical Phonology and Morphology’s three hierarchically ordered strata (levels): stems, words, and phrases/sentences.² The stem phonology corresponds to Lexical Phonology’s level 1 and the word phonology corresponds to Lexical Phonology’s level 2. Together the two are traditionally called the lexical phonology, and I’ll use this term, but without LPM’s theoretical baggage. What is important is that each constitutes a distinct parallel constraint system, and that they interface serially. Stems must satisfy the stem

²The terms *stratum* and *level* are interchangeable in the literature (except in conventional combinations such as Stratal OT and level-ordering) and will both be used here. *Level* was the original term, launched in Allen’s 1978 study of English morphology, and used in early Lexical Phonology (Pesetsky 1979, Mohanan 1982, Kiparsky 1982). Halle and Mohanan 1985 introduced *stratum* in order to avoid the potential confusion with a level of representation.

constraints, and provide the input to the word system, which in turn provides the input to the syntax and postlexical phonology. The relation between each pair of adjacent levels is formally the same as the familiar input/output correspondence relation of standard OT. There is no direct correspondence between the stem phonology and the postlexical phonology.



Therefore the ranking of faithfulness and markedness constraints at each stratum jointly determine what properties of the input will be retained in the output. “Cyclic” retention of properties of bases in derivatives is an input/output faithfulness effect, and opacity is dealt with is by ranking constraints differently at different levels.

The reason the three strata *stem*, *word*, and *phrase* are likely to be universal is that these are the three universal cross-categorical morphosyntactic units. A theory that conflates stem and word phonology, leaving only a lexical and postlexical stratum, would still have some of the advantages of Stratal OT as presented here, but would not be able to account for the full range of cyclic effects and opacity. Such a theory would therefore still need transderivational constraints, the elimination of which I take to be the major result of Stratal OT. It would also be unsuited for treating phonological conditioning of morphology and mutual phonology/morphology dependencies in prosodic morphology, as will be shown in part III.

Conversely, a theory that enriches the set of universal strata, perhaps by further articulating the postlexical phonology into a phrase level and an utterance level, or which permits additional language-specific strata, would retain the essential results obtained here. Suggestive evidence for such a richer postlexical stratification comes from studies by Kaisse 1985, 1990, Kiparsky 1985, Clark 1990, McHugh

1990, Mutaka 1994, Koontz-Garboden 2003, Pak 2005, Pak & Friesner 2006. Cyclicity and opacity *within* postlexical phonology would not only be consistent with such a postlexical phonology, but predicted by the theory. Additional lexical strata have been argued for as well, most convincingly in languages with exceptionally rich morphologies, such as Odden 1996 for Kimatuumbi and Jaker 2011 for Dogrib. In each case, the proposed extra strata, whether universal or language-specific, conform to the phonological and morphological properties predicted by Stratal OT principles.

At a still more specific level, I will be arguing for a particular instantiation of (3), and exploring its empirical and theoretical consequences. Two hypotheses in particular are important.

The first hypothesis concerns the possible differences in ranking between the levels.

- (5) a. *Default*: All strata have the same ranking of phonological constraints.
- b. *Stratum-specific ranking*: The constraint system of stratum $n+1$ may differ in ranking from constraint system of stratum n by promotion of one or more constraints to undominated status.

The import is that if a constraint is ranked differently at the word-level than at the stem level, it is undominated at the word level, and if a constraint is ranked differently at the postlexical level than at the word level, it is undominated at the postlexical level.

Secondly, an assumption about morphology. Affixes are specified for whether they must attach to (that is, whether they select) a Stem or a Word, and whether the resulting form is a Stem or a Word. Affixes are therefore of the following basic types:

- (6) a. Stem-to-stem affixes: [[X]_{Stem} + Affix]_{Stem}
- b. Stem-to-word affixes: [[X]_{Stem} + Affix]_{Word}
- c. Word-to-word affixes: [[X]_{Word} + Affix]_{Word}

In addition, we will also allow for selectionally underspecified affixes, which go both on stems and on words. A weakness of LPM was that it did not explicitly separate the category that the affix selects for from the category that it forms. Giegerich's (1999) theory of stem-driven level-ordering recognizes this distinction and is adopted here.

Although the levels and the affixal categories in (6) are assumed to be universally available, the allocation of morphemes to them is not universally predictable, and not all languages necessarily instantiate all types of affixes. For example, inflectional endings are attached mostly to words in English and Hindi, and to stems in Yokuts, Finnish, and Greek. And some languages have no inflectional endings at all. The choice of inflectional stratum has various morphological and phonological consequences that have attracted the attention of typologists since Humboldt.

(1)-(6) is the theory in a nutshell, though there is much more to it, and many alternative paths awaiting exploration.

To repeat: these proposals are falsifiable at different levels. For example, the discovery that the theory requires an additional stratum (such as the one mentioned above) would not falsify Stratal OT. On the contrary, if the new stratum was well-defined and further sharpened the empirical coverage of the theory, it would confirm it, while leading to a different instantiation of it. (3) would then make a new set of predictions about domains, constraint interactions, and so on. These would then provide additional empirical tests of Stratal OT, potentially conforming or falsifying it.

1.3 The “derivational residue”

The “derivational residue” of OT is the class of generalizations that can be described by ordered rules but, apparently, not by ranked constraints. It consists of (i) opaque relations between phonological processes, traditionally handled by stipulative rule ordering, (ii) cyclic inheritance of phonological properties from bases to derivatives, dealt with in ordered rule theory by application of rules from innermost domains outwards, and (iii) certain types of phonology/morphology interactions. Opacity appears from the parallelist perspective as OVERAPPLICATION and UNDERAPPLICATION, and is so referred to in the the extensive and inventive OT subliteration devoted to dealing with it under parallelism. Cyclicity, also known as synchronic analogy, has provoked almost as many innovative proposals in OT phonology, albeit they have tended to remain programmatic and informal. The problematic phonology/morphology interactions partly overlap with those that face non-interactionist approaches to morphology in general (Scheer 2011).

While the derivational residue is widely acknowledged as a problem, the range of responses to it is quite diverse. Some phonologists take it as conclusively re-

futing the idea that constraints are evaluated in parallel on output representations, and hence as sufficient reason for rejecting OT outright. Some even advocate a return to the unconstrained rule ordering of pre-OT days, which allowed opacity and paradigmatic transfer effects to be dealt with all too easily by rule ordering. The findings reported below confirm that this would be an ill-advised retreat and that the insights of OT are worth retaining. Although sequentially ordered rules generally serve well as a descriptive tool, they are the wrong basis for phonological theory because they have an excess of expressive power, which compromises the explanatory goals. Classical OT has just the opposite problem of insufficient expressive power. In this respect Stratal OT occupies an intermediate theoretical space, which accommodates the derivational residue, and indeed explains much of it, without giving up the descriptive and explanatory gains of classical OT.

Other phonologists propose to deal with the derivational residue by extending OT's constraint repertoire with new types of constraints. Since parallelism is a conceptually attractive core tenet of OT, sound method requires trying to save it in the face of recalcitrant data. The fact is that in struggling with the derivational residue OT has been forced into a gradual retreat from parallelism from its very beginning in the 90s. The first wave of devices designed to save it were transderivational constraints such as Sympathy and O/O constraints, which refer to other outputs that are generated or could be generated by the constraint system. Sympathy constraints require Faithfulness to designated losing candidates, and Output/Output constraints and Paradigm Uniformity constraints require Faithfulness to paradigmatically related forms. They reconstruct the ordering of processes and cyclicity within a formally parallelist theory. More recently interest has shifted to derivational constraints, such as OT-CC with its PREC(EDENCE) constraints, which impose an order on faithfulness violations in a derivation, and Harmonic Serialism. OT-CC is essentially derivational; its only vestigial parallelist feature is that the derivational chain is subjected to a single evaluation. With Harmonic Serialism the abandonment of parallelism is complete. At least to the Ordinary Working Phonologist it look like stipulative rule ordering all over again, only with constraint ranking dictating the order of application.

Transderivational constraints undermine three of OT's central goals: formalization, learnability, and a restrictive factorial typology. Tellingly, most mathematical and computational works on OT phonology ignore transderivational constraints. As far as I know there are no learnability results for them. Basic tools such as OT-Soft (Hayes, Tesar and Zuraw 2003), the Praat OT workbench (Boersma and Weenink 2007), OT-Help (Staubs et al. 2010), and PyPhon (Riggle, Bane &

Bowman 2011) are not very useful for them, since they assume that you can determine whether a form violates a constraint just by inspecting it. Since some such constraints are required in actual descriptive practice by any non-Stratal version of OT that deals with real phonologies (as opposed to toy examples used to illustrate theoretical points), this is a painful lacuna. But there is a good reason for it. Potts & Pullum's (2002) point that these constraint types can't be handled in a formal reconstruction of OT remains unanswered: "output-output correspondence and 'intercandidate' Sympathy are revealed to be problematic: it is unclear that any reasonable class of structures can reconstruct their proponents' intentions." The reason is that whether they are violated depends on rankings of other constraints, or on the existence of other outputs (real, potential, or fictitious, depending on the theory). Sympathy constraints require faithfulness to fictitious candidate which is selected by a designated fictitious constraint ranking. Output/Output constraints require faithfulness to the base, whose shape itself is determined by on the input and the ranking of constraints including other Output/Output constraints. It is not possible to compute factorial typologies for systems that include constraints with this formally refractory property, and the standard learning algorithms cannot handle them.

Derivational constraints such as PREC constraints, on the other hand, are not necessarily fatal to OT. Their properties are still unsettled and several versions of OT-CC are on the market, but it is clear that they are better behaved than Sympathy and O/O constraints, though a far cry from the original simple OT of containment or correspondence theory. Harmonic Serialism appears formally rather clean. The main inadequacy of derivational constraints is on the empirical side. Incorporating either transderivational or derivational constraints leads to systematic typological overgeneration, due to various types of unattested constraint interactions that they give rise to, as bizarre as anything that can be concocted with rule ordering. Many of the attractive results of OT are lost again — even the syllable typology that was its most persuasive initial achievement. At the same time, we shall see that they do not suffice for many types of well-documented phonological phenomena.

Stratal OT takes a very different approach to the problem. It treats seriality not as a stipulated relation between specific processes or specific constraints, but as the predictable result of the intrinsic relation between modules.

The basic idea is not new. OT research on syntax, semantics, and phonology happily continues to assume these well-established components of grammar even though a strict construal of parallelism would deny their existence. Within phonology, limited stratification has been around from the earliest work in OT in the form

of a division between lexical and postlexical phonology as separate constraint systems. It is adopted explicitly in such works as McCarthy & Prince (1995), Cohn & McCarthy (1994), Potter (1994), Hale and Kisseck (1998), Clements (1997), Itô & Mester (1999, 2002, 2003a, 2003b), Kim 2003, Padgett (2003), and Coetzee & Pater 2008, to name just a few.

A much larger group of OT studies has adopted such a two-stratum model unofficially, by simply equating the “output” of the phonology with words in citation form, that is, as pronounced in isolation, without any sentence-level phonology except for prepausal effects (even when these neutralize contrastive features that condition postlexical processes). Because of inexplicitness it is not always clear whether this is a tacit endorsement of two-stratum phonology, some kind of deliberate idealization, or simply negligence. Two strata are also implicitly acknowledged when it is stipulated that Output/Output constraints apply to “free forms”. Other researchers explicitly deny all stratification (McCarthy 2007, Steriade 1999, Burzio 2002a, 2002b), a stance which is certainly more principled and consistent with OT’s leading idea that constraints are evaluated in parallel.

That a truly explanatory theory requires going beyond the lexical/postlexical division and integrating ideas from Lexical Phonology and Morphology into OT has been suggested by a number of researchers. Booij (1996, 1997) puts it forward as a solution to opacity, making the important point that phonological generalizations that hold transparently in the lexical phonology are often made opaque at the sentence level, while the reverse never occurs, a core prediction of Stratal OT. Sign-based phonology (Orgun 1996) attempts to synthesize these approaches using ideas from construction grammar and formal tools of HPSG.

Rubach (1997, 2000, 2003a, 2003b) gives evidence from Polish for two lexical levels and one postlexical level, differing in constraint ranking. Bermúdez-Otero (1999, 2001, 2006, 2007, 2009, 2010, 2011, forthcoming) and Bermúdez-Otero and Hogg (2003) argue for essentially the same three-level stratification into stem, word, and sentence phonology, characterized by partly distinct constraint systems which are parallel but which interface in serial fashion. Bermúdez-Otero has contributed to this research program with numerous important studies which bring in-depth evidence from English, Spanish, and Catalan to bear on it, often with an important diachronic dimension as well.

This three-level stem/word/postlexical articulation of phonology and morphology has also been proposed for Kikamba (Roberts-Kohn 1998), Hebrew (Koontz-Garboden 2001), Auca (Kim 2003), Portuguese (File-Muriel 2004), Russian (Blu-

menfeld 2003, Gribanova 2008), Finnish (Kiparsky 2003, Anttila 2006), British English (Collie 2007), Tundra Nenets (Kavitskaya and Staroverov 2008), Nu-uchahnulth (Stonham 2007, 2008), Oroch (Tolskaya 2008), Korean (Yun 2008, to appear), Chamorro (Kaplan 2008: 168-186), Singapore English (Anttila, Fong, Benus, and Nycz 2008), Japanese (Sasaki 2008) and Dogrib (Jaker 2011).

Fundamentally, though, the theoretical divide runs between strict parallelism and modularization. Even the evidence for a modular distinction between syntax and phonology already undermines OT's aspirations of a fully parallel architecture for grammar. The recognition of a postlexical/lexical distinction undermines it for phonology. Once parallelism is abandoned, the question whether the lexical phonology includes a stem level distinct from a word level ceases to be a matter of first principles and becomes an empirical question about the implementation of Stratal OT in phonology. This book, therefore, can be read as an extended argument both for modularity and stratified phonology in general, and for the specific LPM-inspired version of stratified OT phonology that seems, at least for the present, to be the best instantiation of Stratal OT.

That stratified OT is formally well-behaved, unlike transderivational OT, follows from two mathematical results. First, Karttunen has shown that a classical OT system can be translated into a finite state transducer. Second, Kaplan has shown that a cascade of finite state transducers can be compiled into an equivalent single finite state transducer.

The initial arguments for Stratal OT came from the so-called derivational residue, primarily from opaque constraint interactions and "cyclic" effects. In a range of complex and well-documented cases, the analysis, when developed in sufficient depth, demonstrates the superiority of Stratal OT to both rule ordering and transderivational OT alternatives. But the descriptive virtues are only part of the story, the lesser part I think. Stratal OT does not just provide tools for "handling" opacity and paradigmatic effects, but, in many interesting cases, *predicts and explains the conditions under which they occur*. It does not re-import stipulative rule ordering and cyclicity into OT in a new guise. It is as remote from traditional derivational approaches employing unconstrained rule ordering as it is from parallel OT with stipulative Sympathy and Output/Output constraints. Unlike these alternatives, which fail to relate opacity and paradigmatic effects either to each other, to morphology, or to anything else in the grammar, Stratal OT derives them in a principled way from the organization of the grammar, specifically from the interaction of phonology and morphology in a stratified grammar and lexicon.

“Derivational residue” is therefore a misnomer. “Modular” might be more felicitous than “derivational”, and it is in any case not a “residue”. Actually the phenomena in question have always stood at the very core of theorizing about such issues as naturalness and abstractness, the phonology/morphology, phonology/syntax, and phonology/phonetics interfaces, sound change and analogical change. It is only from the strictly parallelist perspective that they can be portrayed as gratuitous overapplication or underapplication of constraints, and marginalized as mere nuisances for the theory. Stratal OT shows that the constraints in such cases neither overapply nor underapply, but rather take effect where predicted in contexts which are systematically masked in the output by other constraint applications, as dictated by the organization of the grammar. Seen in this light, the “residue” turns out to be a powerful probe into the organization of grammar.

Much OT theorizing about opacity and paradigm effects is based on idealized data involving a few phonological processes. Stratal OT by its very nature raises the stakes, and the case for it must be built on richer empirical ground. It cannot be convincingly supported or refuted solely by fragmentary analyses based on data the size of introductory phonology problems. In principle, motivating the strata and morphology-phonology interactions of a language requires working out its entire phonology and morphology. Because Stratal OT tightly relates phonology to morphology and restricts the ways in which processes can interact, it can make rich projections from sparse data. From the viewpoint of understanding language acquisition, this is a desirable feature of Stratal OT, for such projections are precisely what facilitates the learner’s task. From the viewpoint of the linguist faced with testing the predictions, it requires, at a minimum, enough morphological data to independently justify the predicted assignment of constraint rankings to their respective strata. Indeed, the evidence could potentially come from anywhere in the grammar. If the theory is right, there should be multiple convergent evidence for it in the grammar, but unearthing it may require a nontrivial investment of analytic effort.

In order to capitalize on that I have tried to embed my arguments in fuller phonological and morphological analyses than is usual. This meant concentrating on well-documented languages and presenting enough of their phonology and morphology. The relevant generalizations as I understand them have been stated at least verbally if not formally, and enough data is given to enable readers to develop their own counterproposals. I encourage the reader to rework the analyses in their favorite phonological theory.

In-depth empirical comparisons of Stratal OT with classic OT and transderiva-

tional OT require reasonably complete phonological and morphological analyses. Unlike previous theoretical innovations in phonology, OT has produced few of them. Metrical phonology (Lieberman & Prince 1977), autosegmental phonology (Goldsmith 1976, 1990), and Lexical Phonology (Kiparsky 1982, Mohanan 1985) were immediately adopted by working phonologists who wanted to analyze complex morphophonological systems and to write comprehensive phonologies of languages. Enthusiasm of OT, if anything even more immediate, was aroused primarily by its formal interest and conceptual attractiveness, particularly its integration of naturalness and typology into phonological theory, and by the simple and principled architecture of grammar it offered. It was not matched by a commensurate amount of in-depth analytic work. It turned out to be difficult to write comprehensive phonologies in classic OT and in transderivational OT. That is why ordered rule theory, in spite of having been marginalized in the theoretical discussion, continues to thrive in descriptive grammars. I believe that Stratal OT will serve descriptive grammarians better than other versions of OT, just as it will serve them better than LPM.

Still, even toy analyses are extremely useful for clarifying and probing conceptual issues, as long as they are not mistaken for empirical evidence. They bring out the consequences of different versions of OT for the analysis of problematic constraint interactions and phonology/morphology interactions, and lead to new predictions that invite deeper analysis.

Important evidence comes from considerations of learnability, naturalness, typological restrictiveness, and generalizations about language change. The problems of diachronic and “synchronic” linguistics are in my view fundamentally the same. In particular, the problem of phonologization in sound change is a special case of the problem of opaque constraint interaction, and that it receives the same solution in Stratal OT (Kiparsky 2014).

2 Opacity

2.1 Opacity in rule-based phonology

Opacity as non-interaction Derivational opacity was originally defined as a relation between a rule and the output of the grammar that it is a part of.³ A rule is opaque to the extent that there are output forms that look like they should have undergone it but did not, or output forms that that look like they should not have undergone it, but did. Opacity is not simply a matter of non-feeding and non-bleeding rule order, or of underapplication and overapplication.⁴ This can be illustrated by structure-building operations, such as stress assignment. In Arabic, final syllables are stressed just in case they are superheavy. Geminate consonants are regularly shortened in word-final position, so that underlying superheavy /-VCC/ is realized as -VC. In Bedouin dialects, a word-final -VC that is degeminated from /-VCC/ is treated like any other final -VC for purposes of stress, in accord with its output form. In Mesopotamian and Syrian-Palestinian dialects, though /-VCC/ behaves as a superheavy syllable in spite of degemination:

- (7) a. Bedouin: /yi-midd/ [yímid] ‘he spreads, extends’ *transparent*
b. Syrian: /bi-mədd/ [bimód] ‘id.’ *opaque*

In Bedouin, stress is transparent because it is assigned on the basis of the output form. In Syrian, stress is opaque because it is assigned on the basis of the input form. Here the terms “(non-)feeding” and “(non-)bleeding”, or for that matter “overapplication” and “underapplication”, are not appropriate. Both stress and degemination apply in both dialects; it is just that stress is assigned to a different syllable depending on whether it applies before degemination (the opaque case) or after degemination (the transparent case).⁵

The one generalization which covers *all* cases is that the transparent case is the one that *maximizes the interaction of processes*: in other words, an environment-changing process precedes an environment-dependent process. For example, in

³It should not be confused with the impermeability of “neutral” segments to vowel harmony and other long-distance spreading, also called opacity.

⁴See Bakovic 2011 for the full story.

⁵An exactly analogous example is Palestinian Arabic *fihim-na* ‘our understanding’ (Kiparsky 2000), versus Mesopotamian *fihim-na*, with respectively opaque and transparent interaction of stress and epenthesis.

the transparent derivation (7a), degemination, which changes syllable weight, takes effect before stress, which is sensitive to syllable weight. A unifying characterization of the opacity asymmetry is implicit behind the construction of Pāṇini's grammar of Sanskrit, and was made explicit by later Sanskrit grammarians. In modern terms it can be formulated as follows (Joshi and Kiparsky 1977, Kiparsky 1982):

- (8) Let $C(\phi)$ denote the output resulting from the application of C to an input ϕ (where C is a rule, or a system of rules or constraints that yields a well-defined output for any input). Then B is opaque in $A(B(\phi))$ iff $A(B(\phi)) = A, B(\phi) \neq B(A(\phi))$.

As the example of degemination and stress in (9) illustrates, opaque interaction involves minimizing the interaction of processes:

- (9) a. Degemination(Stress(*yimidd*)) = *yimíd* (opaque stress, Bedouin)
 b. Stress, Degemination(*yimidd*) = *yimíd*
 c. Stress(Degemination(*yimidd*)) = *yímid* (transparent stress, Syrian)

(9) correctly defines (9a) *yimíd* as the opaque form, because (9a) *yimíd* = (9b) *yimíd* \neq (9c) *yímid*. It also extends correctly to the familiar (non-)feeding and (non-)bleeding cases.

Rule ordering theories reveal no theoretically significant distinction between opaque and transparent rule ordering. Yet there is overwhelming evidence from acquisition, change, and processing for the unmarked status of transparency, that is, of *maximal rule interaction*.

Sometimes (8) does not establish an asymmetry between processes. If, for all A and B , $A(B(\phi)) = A, B(\phi) = B(A(\phi))$, then, obviously, A and B do not interact and according to (8) there is no opacity either way. More interesting are cases where $A(B(\phi)) \neq A, B(\phi) \neq B(A(\phi))$. Here both processes deprive each other of the chance of taking effect (mutual bleeding). A hypothetical example, schematized from an actual one from ancient Greek, is the following. Suppose that an input /pater-óon/ is potentially subject to two processes, one shifting the accent to a presuffixal light syllable (*patéroon*), the other deleting an unaccented vowel in a light syllable (*patróon*). Only one of the processes can possibly apply. Simultaneous application would be ill-defined or impossible here, since the presuffixal

vowel cannot be both accented and deleted at the same time. Therefore, although the two processes are crucially ordered, (8) again defines no opacity either way. In a constraint system, this case corresponds to the situation where different rankings of markedness constraints yield different outputs.

Overapplication and underapplication Because ordering is not available in parallel constraint systems, opacity is a problem for OT and for any constraint-based approach. Characterizing opacity formally in an OT constraint system is a tricky matter, because constraints (unlike rules) do not correspond to “processes” in any direct way. When constraints are defined on output representations, as in OT, then the problem in constructing a constraint system corresponding to (??) is that A is not realized as B even though a constraint forces it to. And the problem in constructing a constraint system corresponding to (??) is that A is realized as B even though no constraint forces it to. In other words, non-feeding orders appear as unmotivated markedness constraint violations, and non-bleeding orders appear as unmotivated faithfulness constraint violations. These are the situations that the OT literature respectively refers to as UNDERAPPLICATION and UNDERAPPLICATION of constraints.

One virtue of OT phonology is that the privileged character of transparency as defined by (9) is captured automatically, simply because constraint satisfaction is evaluated on outputs. This is not true for phonology/morphology interactions in parallel OT, whereas Stratal OT does extend correctly to those cases as well. In any case, all versions of OT correctly characterize transparency as the default for phonology-internal constraint interactions. For example, the transparent output *yimid* from /yimidd/ in Syrian Arabic comes for free. Deriving the opaque *yimíd* of Bedouin Arabic requires a bit of extra work. This is good, because there is a lot of evidence that, other things being equal, opacity constitutes an increment of complexity for learners. It is on the question what that extra work is that the theories diverge. In classical OT there really is no phonological solution to the Bedouin pattern (9a). In transderivational OT, it requires either activating a Sympathy constraint which transfers the stress of an imaginary output that does *not* undergo degemination to the actual output that does, or else a PREC constraint which imposes a derivation equivalent to the one posited by the rule ordering analysis. In Stratal OT, it requires learning that stress is lexical and final degemination is postlexical — a conclusion which could be reached by several independent routes, including the respective domains of the processes and their interaction with other phonological process and with morphology.

To summarize: opacity comes about when processes which *could* interact *don't* interact. This is the idea that (8) formalizes. It correctly characterizes non-bleeding and non-feeding interactions as opaque, but it extends also to interactions such as those in (9), where the taxonomy of feeding and bleeding, or overapplication and underapplication, is not helpful. Classical parallel OT predicts exactly the transparent interactions, as defined by (8). The problem is that it cannot deal with the others, while transderivational OT deals with them in the wrong way. The failure of ordering theories to privilege transparency is as damaging at the explanatory level as parallel OT's failure to countenance opacity is at the descriptive level.

2.2 Opacity in transderivational OT

Opaque interactions of phonological processes occur in practically any phonological system. A parallel OT theory in which all faithfulness constraints are based on I/O correspondence is incompatible with opaque constraint interaction. In the face of this dilemma, there are three a priori reasonable stances to choose from.

Denial The most radical one, harking back to a view that arose in 1970's rule-based phonology under the heading of Natural Generative Grammar, and ultimately going back to Saussure, is to treat opacity as an artifact of the sequential character of sound change. This view holds that opacity created by successive sound changes is not synchronically apprehended by language learners and language users in terms of processual non-interaction, but in terms of lexical exceptions or morphological restrictions on processes (see Sanders 2002, 2003 for an extended defense of this position in an OT framework, cf. the morphological solution to Lardil final -V deletion mentioned above). In this radically surface-oriented view of phonology, opaque generalizations are only granted a historical reality. For example, words with underlying final geminates in Bedouin Arabic (such as (7a)) would be lexical exceptions to regular stress and would instead be assigned final stress.

Assimilating opacity to exceptionality and morphological conditioning does not merely complicate descriptions and lead to loss of generalizations. It demands new descriptive resources. It is true that genuine exceptions and morphologized processes exist in any case, and require arbitrary lexical and morphological restrictions on the applicability of constraints. However, using them to handle opacity

would require greatly extending what they can do. All current proposals amount to indexing morphemes or lexemes for a constraint or a constraint ranking (Inkelas, Orgun, and Zoll 1997, Itô and Mester 1999b, Anttila 2002, Pater 2010). Any such mechanism will fail to extend to opacity properly, for in the general case opacity cannot be tied to any particular morpheme or lexeme. A simple example of contextual neutrality in vowel harmony will make this clear.

The Seto dialect of Estonian has pervasive front/back vowel harmony (Kiparsky and Pajusalu, 2003). Non-initial /i/ is neutral (though /i/, which occurs only in initial syllables, triggers back harmony). Consonants, including palatalized consonants, are also neutral. *õ* in (10) is phonemically a back mid vowel, realized in unstressed position as a schwa.

- (10) a. *opp:a-ji-lõ* ‘to teachers’ *rebäs-i-le* ‘to foxes’
 b. *klibisõ-ma* ‘to rattle’ *libise-mä* ‘to flutter’
 c. *naáa-tta-nuq* ‘joked’ (Pp.) *näli-ttä-nüq* ‘starved’ (Pp.)

Distinctively palatalized consonants and the glide /j/ tend to cause some degree of fronting in a following vowel (Hagu 1999:6). This fronting is gradient and normally does not result in complete merger, e.g. *ísura* [íʃura] ‘boy’. It is purely local and does not interact with the harmonic pattern of the rest of the word; regular back harmony always resumes in the next syllable. The gradient fronting is an allophonic process, governed by constraints of a separate postlexical constraint system which applies to the output of the lexical phonology, perhaps a matter of coarticulatory phonetics. However, fronted ə, the unstressed allophone of the back vowel /õ/ (stressed [ɤ]), does merge fully with *e* e.g. *naáse-lõ* [naáselõ] ‘woman’ (Allative). The result is a true front vowel in the middle of a back-vowel word, rendering vowel harmony opaque. But it is quite impossible to tie this effect to any morpheme or morphological category. It is not a property of the Allative ending, nor of the noun ‘woman’. Both these morphemes trigger, undergo, and transmit harmony with absolute regularity. It is only the *phonological* conjunction of a palatalized consonant with an immediately following mid vowel that triggers the opacity.

A second reason for rejecting the treatment of opacity by exception features or lexical/morphological restrictions on constraints is evidence of the productively phonological character of opaque generalizations (see McCarthy 2007 for a summary). Many such instances will be found in the analyses below, beginning with Finnish in Ch. 2. Moreover, linguistic change shows that opaque phonological

generalizations play a role in change in a way that would be incomprehensible if they were not apprehended and acquired as phonological by learners.

Transderivational constraints The obvious way to hold on to maintain parallelism in OT is by adding new types of constraints. This program has been pursued especially by J. McCarthy and his collaborators. McCarthy (1999, 2003c, 2007) advocates SYMPATHY constraints, designed expressly for dealing with opaque constraint interactions. They impose faithfulness between the output and a candidate selected by another faithfulness constraint. While formally constructed within parallelist bounds, Sympathy allows OT to capture some effects of SPE-style ordering theory. However, it more restricted than ordered rules in some respects, and less restricted in others, in both cases to its detriment, as we shall see below.

A more recent approach to opacity is OT with Candidate Chains (OT-CC), which has a very different architecture than classical OT and is in effect a derivational theory. In previous versions of OT, GEN generates a universal candidate set from which EVAL selects the most harmonic in accord with the language's constraint ranking. In OT-CC, GEN and EVAL collaborate to generate a CHAIN, which links input to output in a derivation-like series of minimal steps determined by the language's constraint system. The initial link in the chain is the most harmonic faithful parse of the input. Every successive link in the chain must be minimally less faithful than the immediately preceding one, i.e. it must have all of its unfaithful mappings plus one (GRADUALNESS), and it must be more harmonic than the preceding with respect to the language's constraint hierarchy (HARMONIC IMPROVEMENT). (In the conceptually similar Harmonic serialism approach, the output of each step becomes the new input to the next round). Opacity is then handled by PREC(EDENCE) constraints which impose an order on the faithfulness infractions in a chain. PREC constraints are not faithfulness constraints (unlike Sympathy constraints), but well-formedness conditions on derivations.

OT-CC shares much of classical rule ordering theory's excessive richness, and thereby represents a step backward compared even to LPM, let alone to Stratal OT. The gradualness requirement it imposes on derivations is too strong and actually gives up some of the nicest results of OT phonology (Kiparsky, to appear).

I conclude that so far, all treatments of opacity in parallel OT with sufficient descriptive coverage have depended on a special transderivational mechanism of considerable power, which is in practice harnessed to a rather limited purpose and

which does not serve that limited purpose very well.

2.3 Opacity in Stratal OT

Stratal OT changes the other prong of the original OT program. It keeps a restricted constraint inventory (only Markedness and I/O Faithfulness constraints) and introduces a regimented seriality between the stem, word, and postlexical strata. Morphology provides independent evidence that these strata and their serial relationship are part of UG. Very crudely put, Stratal OT recasts the traditional Lexical Phonology view of the organization of morphology and the lexicon into OT terms by treating each level as a parallel constraint system with transparently interacting I/O faithfulness and markedness constraints. Opacity then results from the masking of the constraint system of one domain by the constraint system of a more inclusive domain. Naturally, this marriage of OT and LPM requires rethinking many things from the ground up, and traditional assumptions on both sides may have to be modified or discarded.

One conceptual advantage of Stratal OT is that it uses only independently motivated theoretical devices to deal with opacity. The assumption that stems, words, and phrases are each governed by their own constraint systems is needed anyway to define regularities over the elements in each domain (stem structure, word structure, sandhi), and the assumption that these constraint systems are serially related is needed anyway to deal with the paradigmatic effects by which larger constituents reflect the form of the smaller constituents they contain (so-called CYCLICITY). Opacity, then, is a side effect of domain stratification.

Stratal OT's treatment of opacity has independently testable phonological and morphological consequences. In many cases it predicts what appear to be correct generalizations that transderivational OT cannot even express. In particular, it derives systematic predictions about constraint interaction from morphology and from the transitivity of opacity/transparency relations.

The interaction of processes is determined by the intrinsic relation of the levels. If constraint C_1 is active (visible) at level n , and constraint C_2 is active at level m , then:

- a. If $n = m$, then C_1 and C_2 must interact transparently (they are mutually visible)

- b. If $n < m$, then C_1 is visible to C_2 and C_2 is not visible to C_1
- c. If $n > m$, then C_2 is visible to C_1 and C_1 is not visible to C_2

The hypothesis, then, is that all opaque constraint interactions and “cyclic” effects arise this way. I propose to show that combining OT and Lexical Phonology and Morphology this way retains the insights of both, and in particular restores the possibility of a restrictive and well-defined constraint inventory such as was originally envisaged in OT. And I show that their combination affords new understanding beyond what either of them do separately.

Unlike Sympathy and OT-CC, Stratal OT limits the depth of opaque constraint interaction: in a system with three levels, there are maximally two degrees of opacity. If we further articulate the postlexical phonology by distinguishing a separate phrasal domain within it, we increase the possible depth to three, with additional specific commitments about the domains of the constraints involved. Whether three or four, the number of domains is in any case small, and, more importantly, determined by much other evidence that has nothing to do with opacity. Parallel OT, in contrast, has as many Sympathy constraints as there are faithfulness constraints (as yet no-one knows how large that number is), or some similarly large number of PREC constraints in the OT-CC version, and their ranking is determined by nothing except the opacity that they are supposed to account for. Therefore it can reconstruct the effect of rule ordering of arbitrary depth. This makes a difference for learnability, factorial typology, and computational implementation.

The disparity in depth of permitted opacity is impressive enough in terms of quantity, but the qualitative differences are more significant. Sympathy and PREC constraints are blind to the functional relationships between processes and of the domains in which they are active. In fact, as we shall see in Ch. ??, they predict the possibility of truly weird unattested kinds of constraint interactions. Surprisingly, in spite of their excessive power, transderivational constraints are also too weak, in that identify processes by the faithfulness violations they incur, and hence cannot distinguish between similar processes at different strata (see Kiparsky 2000 for brief discussion). The connection that Stratal OT establishes between phonological domains and derivational opacity gives it a handle on such phenomena, yet contributes to the OT typological program by ruling out a large class of impossible phonological systems.

OT innovations, paradigmatic constraints have pre-OT antecedents, primarily in a vast body of

3 FAQs about Stratal OT answered

This section addresses a collection of general criticisms that have been aimed at Stratal OT, some of a general nature and others concerning particularly the form of Stratal OT advocated here.

3.1 Opaque postlexical processes?

Optional processes Stratal OT predicts transparent constraint interaction within a level, setting apart the inherent opacity that is predicted by cyclicity, by constraint domination (as in standard OT), and finally (if the approach to chain shifts suggested in Kiparsky 2011 is right) by super-optimality. Selkirk’s study of English and French sandhi offers striking confirmation from postlexical phonology, and Clark’s LPM phonology of Igbo explicitly noted that level ordering eliminated all opaque rule ordering from the phonology (1990:97).

Against this, McCarthy (2007) cites Donegan and Stampe’s (1979) discussion of fortitions and lenition in English fast speech, which in their view are governed by natural, universal, and innate “processes” rather than “learned rules”. The processes they cite include both uncontroversially phonological ones, such as American English flapping (tapping), and coarticulation phenomena whose phonological status might be questioned, such as the progressive nasalization in *signal* [sɪgnɪl] → [sɪgñ̩l]. Unlike Selkirk’s sandhi material, this data raises issues about the demarcation of phonology and phonetics. Donegan and Stampe argue for the phonological character of their processes, primarily on the grounds that they may apply in counterfeeding order in certain dialects and styles, and that some processes must apply counterbleeding order. If they are right, the opaque cases would be a problem for classic OT. Most of them would be problems for Stratal OT as well, for the majority of Stampean processes (insofar as they are part of phonology and not phonetic implementation) are postlexical, and constraint interaction within that level should be transparent.

Among the processes that Donegan and Stampe consider are these four:⁶

- (11) a. Regressive nasalization: *plant* [plænt] → [pl̩̃ænt]

⁶The formulations are from Donegan and Stampe. I think the parentheses mean that the process is favored under the parenthesized conditions but can apply also outside of them.

- b. Elision of nasals before homorganic (tautosyllabic) (voiceless) consonants: [pl̥æ̃t] (with regressive nasalization)
- c. Flapping of intervocalic tautosyllabic apical stops: *that apple* [ðætæp] → [ðæræp]
- d. Progressive nasalization of (tautosyllabic) sonorants in unstressed syllables after nasalized segments: *signal* [sɪgnɪl] → [sɪgnɪ̃l]

Since all of them are optional, they generate five increasingly reduced and informal pronunciations of *plant it*:

- (12) a. [plæntɪt] (lexical representation)
- b. [pl̥æ̃ntɪt] (regressive nasalization)
- c. [pl̥æ̃tɪt] (regressive nasalization, elision of nasals)
- d. [pl̥æ̃rɪt] (regressive nasalization, elision of nasals, flapping)
- e. [pl̥æ̃rɪ̃t] (regressive nasalization, elision of nasals, flapping, progressive nasalization)

These are optional (variable) processes which apply preferentially if not exclusively in fast speech and informal style. Establishing an optional counterfeeding relation between optional processes is pretty challenging: how can we tell whether the underapplication of a process is due to an ordering restriction or just to the option of not applying it? To exclude the latter possibility, Donegan and Stampe argue that certain counterfeeding rule orderings are obligatory or near-obligatory at particular stylistic levels. They cite “speakers who, though they regularly flap basically intervocalic [t] as in *pat it*, do not flap the derivatively intervocalic [t] of [pl̥æ̃tɪt] *plant it*.” For these speakers, then, elision of nasals could potentially supply new inputs to flapping, but doesn’t (it “counterfeeds” it). For their own speech, they report variation between “sequenced application ([pl̥æ̃rɪ̃t]) in informal styles varying with nonsequenced ([pl̥æ̃tɪt]) in formal styles, while basically intervocalic [t]’s are almost invariably flapped.” In this variety, the processes would apply optionally in feeding order in informal registers and almost always in counterfeeding (nonfeeding) order in formal registers. Why would the option be usual in *plant it* in the same register in which it is rare in *pat it*, if both result from the same system where the processes are interacting transparently (modeled by feeding order in Donegan and Stampe’s theory)? Therefore — so goes the argument — we really have counterfeeding and not just the non-application option.

This argument for counterfeeding order is questionable on two grounds. First, constraint-based theories of variation (such as Anttila's) already predict that flapping in *plant it* is going to be substantially rarer than flapping in *pat it*. Suppose that the probability of vowel nasalization in *plant* is n , that the probability of eliding n after the nasal vowel is m , and that the probability of flapping in *pat it* is k . Since flapping in *plant it* can only take place if the [n] is elided (it is incompatible with its coronal closure), which in turn can only take place if the vowel is nasalized (*plant* can't be pronounced as **plat*), the simplest model of variation would say that the probability of flapping in *plant it* is only $n \times m \times k$: if, for example each of the three processes process applies in half the possible cases, then *pat it* will have 50% flapping and *plant it* will have 12.5% flapping. Of course matters are more complex since different processes may have different probabilities, and the probabilities may not be independent. But the point is that the kinds of disparities in the frequency of applying optional rules on which Donegan and Stampe base their argument arise naturally in consequence of the feeding dependencies among optional rules.

A second point is that defining a given stylistic level involves more than specifying a constraint ranking (or rule ordering). Thinking of styles as grammars is certainly an oversimplification. It is well known, for example, that frequent words tend to be reduced more often, and to a greater degree, than rare words. The obvious reason is that frequent words can be guessed more easily than rare words, and that speakers intuitively know this and take advantage of it to save on articulatory effort. (In rare words a corresponding reduction would not be judged cost-effective because the potential cost of misunderstanding and/or the risk of having to repeat the utterance would outweigh whatever articulatory economies it might achieve.)

This approach to the stylistic stratification of speech variants correlates the informality of a fast speech pronunciation with its distance from the formal pronunciation. This distance can be measured by the number of its optional faithfulness violations relative to the input to the postlexical phonology (which we are assuming is the output of the lexical phonology). If, as Donegan and Stampe put it (1979: 147), there is a tradeoff between clarity and expenditure of phonetic effort, then the relative informality of a variant might be assessed by how it negotiates that tradeoff. On the ranking of alternative output pronunciations established by (12), [p̃lærit], with four faithfulness violations, and [p̃læ̃rit], with six faithfulness violations, are more informal than [plærit], which has just one.⁷ Therefore they

⁷The details of how we count faithfulness violations are not important; the relative formality

might be unacceptable in some stylistic registers where [pæɪrt] is acceptable, in spite of the fact that all three are produced by the same transparent constraint ranking (allowing for optionality).

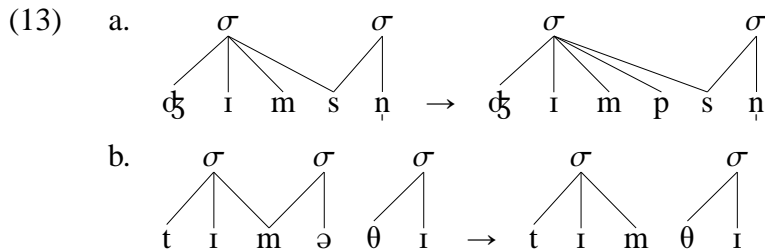
The main lesson is that opacity should be investigated with *obligatory* processes, where Stratal OT excludes even the semblance of level-internal counterfeeding or counterbleeding that can result from optionality in the way just described.

Obligatory processes It turns out that there is not a single clear case of opaque interaction among postlexical processes in Donegan and Stampe’s material.

As an example of counterbleeding among processes, Donegan and Stampe mention that (12a) regressive nasalization is never bled by (12a) nasal elision. For example, *can’t* may be pronounced as [kæ̃nt] or as [kæ̃t] but not as *[kæt]. In hindsight, of course, this is not a counterbleeding relation at all. Rather, it is only the nasalization of the vowel that enables the deletion of the nasal (it feeds it, in a sense). MAX(nas) dictates that the nasal may be displaced but not entirely lost; the contrast between *can’t* and *cat* must be preserved.

Most of Donegan and Stampe’s examples of counterfeeding and counterbleeding involve syllable structure, sometimes implicitly. In every case, once this covert intermediary is taken into account, it turns out that the opacity goes away. For example, they note that casual speech syncope of the medial vowel in words like *sinister*; *Timothy* does not feed the “stop insertion” process seen in *Jimson* [ɕjɪmpsn̩], *sense* [sɛns] → [sɛnts] (pronounced like *cents*). *Sinister* cannot be *[sɪntstɪ] and *Timothy* cannot be *[tɪmpθɪ]. In hindsight, this is not a counterbleeding relation either (even assuming, for the sake of the argument, that the intrusive stop is inserted by a phonological process at all). The intrusive stop appears in tautosyllabic nasal+fricative sequences. In elided *sin’ster*, *Tim’thy*, the nasal+fricative sequences are not tautosyllabic, for elision of the medial vowel does not lead to ambisyllabicity, because a dominant faithfulness constraint blocks resyllabification. Therefore elision does not create the required context for intrusive stop insertion. Assuming ambisyllabicity, the syllabic contrast between *Jimson* and *Tim’thy* is as follows.

of variants comes out the same on any reasonable measure.



Donegan and Stampe take the position that feeding and nonbleeding order are unmarked (at least for processes in their sense). They state that apparent cases where nonfeeding and bleeding order are unmarked really reflect other principles, namely that rules precede processes (with which I obviously agree, *mutatis mutandis*), and that fortitions precede lenitions. The first of these two principles, moreover, is supposed to prevail over the second.

They illustrate the claim that fortitions precede lenitions by the familiar example of *kisses*, *knitted*, where epenthesis (a fortition process, by their assumptions) bleeds assimilation (a lenition process). This is just the type of case which is dealt with cleanly in classic OT and in Stratal OT (“counterfactual feeding”).

As an instance of the workings of their precedence clause, Donegan and Stampe cite the fact that the shortening of [i:] to [ɛ] in *obscenity*, *dreamt* feeds the “southern” raising of [ɛ] to [ɪ] before nasals: *obsc[ɪ]nity*, *dr[ɪ]mt*. Their point is that this is apparently a case of a lenition feeding a fortition, but in reality it is a rule feeding a process. The Stratal OT account is in the same spirit: shortening in *obscenity*, *dreamt* applies at the stem level, while dialectal raising of [ɛ] to [ɪ] before nasals applies at the word level, so that shortening necessarily feeds raising.

On the other hand, the raising of [ɛ] to [ɪ] before nasals precedes all other processes, e.g. *lemme* ‘let me’ (obligatory counterfeeding) and [sɪt] *sent* (an obligatory counterbleeding relationship among obligatory processes). The Stratal OT account is that dialectal raising of [ɛ] to [ɪ] before nasals applies at the word level, and assimilation in *lemme* and nasal elision in *sent* are postlexical (the latter perhaps more properly ascribed to phonetic implementation).

I conclude that the empirical findings of careful studies of postlexical phonology are readily accounted for within Stratal OT. They are consistent with the thesis that constraints interact transparently within a level. Therefore we don’t want theories that make opacity available in such cases, let alone theories like OT-CC or Sympathy, which allow opaque constraint interactions of arbitrary depth.

3.2 Intra-stratum opacity?

3.2.1 McCarthy's claim

McCarthy (2007: 40) claims that “the Lexical Phonology program never sought to eliminate within-stratum rule ordering, including opaque ordering”, because “this hypothesis was self-evidently wrong”. On the contrary, the question whether level-ordering allows the elimination of extrinsic rule ordering was very much a subject of debate in LPM, and detailed empirical arguments were put forward both pro and con. Rubach (1984a) noted that his analysis of Polish phonology required extrinsic ordering. The early discussion of the issue was clouded by unclarity about the status of the word level, and by the mistaken assumption made initially by some phonologists (myself included) that rules are, in the normal case, confined to a single stratum. As the LPM framework developed, the apparent need for extrinsic ordering shrank. Thus, Clark (1990:97) pointed out that level ordering eliminated all opaque rule ordering in her LPM analysis of Igbo phonology.

Significantly, most of the putative cases of extrinsic ordering cited in the Lexical Phonology literature involved relations between rules *within the lexical stratum*. There is also one example in the literature of postlexical rules feeding or bleeding lexical rules, Ondarroa (Ondarru) Basque accentuation and vowel deletion (Hualde (1996). In Kiparsky (to appear) I show that even the seemingly most challenging cases, such as Icelandic, Russian, Tigrinya, and Basque, are readily analyzable in Stratal OT, once stem and word phonology are distinguished. A single example, chosen for its extreme simplicity, will have to suffice here: Catalan nasal assimilation.

3.2.2 Catalan

In Catalan, nasals assimilate in place to a following stop. In word-final position, the triggering stop is itself deleted. E.g. /pont/ *pont* ‘bridge’ is pronounced [pɔn], not *[pɔnt] (cf. *pontet* [pun'tɛt] ‘little bridge’). The deletion of word-final stops takes place even if the next word begins with a vowel: *pont antic* ‘old bridge’ is [pɔ.nən.'tik], not [pɔn.tən.'tik]. This implies that deletion takes effect in the lexical phonology. Assimilation is also lexical, since it must take effect before cluster simplification. This is a counterbleeding, opaque relation.

(14)	Underlying	bɛn-k	Underlying
	Place assimilation	bɛŋ-k	
	Cluster simplification	bɛŋ	
	Surface	'bɛŋ	
		'I sell'	

From the observation that both Place assimilation and Cluster simplification are lexical, as in my Lexical Phonology analysis (Kiparsky 1985), McCarthy (2007: 40) infers that they belong to the same stratum, and that their opaque relation therefore falsifies Stratal OT. Actually there are two lexical strata, the stem and the word. Cluster simplification takes effect only at the word level, for stem-final clusters appear intact unless they are word-final, as noted above. Nasal assimilation, on the other hand, takes effect only at the stem level; there are no instances where it must apply at the word level. Its contexts are limited to morpheme-internal nasals+stop clusters, and a small number of stem-level affixes. The *-k* which triggers assimilation in /bɛŋk/ is a stem formative, if it is a suffix at all. It has no discernible meaning, and appears apparently unpredictably in various forms of a certain class of verbs (deCesaris 1986, Viaplana 2005)⁸. But if nasal assimilation applies at the stem level, it *cannot* be bled by any process which applies only at the word level. Once the morphology is understood, the opaque relation of the phonological processes in question is not only unproblematic for Stratal OT, it is *predicted* by it. Indeed, it would be a problem for Stratal OT if the facts were different.

McCarthy mentions only dialects where deletion of word-final stops after nasals is obligatory in all contexts. Actually the conditions are more complex. A word-final velar stop after a nasal is optionally retained in prevocalic and prepausal position (Wheeler 2005: 223). E.g. *en tinc cinc* [ɛn.'iŋ.'iŋ.'siŋk] 'I have five of them', *blanc i negre* ['blaɛn.'ki.'nɛg.rə] 'back and white' (*ibid.*). In these dialects, then, deletion of final velar stops in complex codas is optional at the word level. But it is obligatory in the postlexical phonology, where resyllabification across word boundaries bleeds it (a transparent interaction within a stratum, as expected). Nasal assimilation remains a stem-level process, so the explanation for the observed opacity holds a fortiori in these dialects too. To repeat, these Catalan data are fully compatible with Stratal OT and indeed strongly support it.

⁸In fact, some analyses treat the forms in *-k* as “velarized” allomorphs of roots

3.3 Is Stratal OT too powerful?

Does Stratal OT harbor monsters? In the preceding sections I addressed the claim that Stratal OT is too restrictive in correlating constraint rankings with phonological domains. In particular, it is sometimes claimed that it predicts non-occurring constraint interactions, or fails to predict occurring interactions. These objections turned out to lack empirical support. I now turn to a purported weakness of the opposite kind: that Stratal OT is too rich because it allows arbitrary differences between strata. This objection was first expressed by McCarthy (2000) and has since been often repeated. Itô and Mester (2002) say:

“... unlimited freedom of demotion is clearly too powerful since there is nothing to prevent unwanted combinations, such as the lexical phonology of Dutch paired with the postlexical phonology of Indonesian, or the lexical phonology of Hindi with the postlexical phonology of English, etc.”

This popular argument has never been spelled out in detail. It is purely rhetorical. If we consider what these combinations would look like we see that they would be perfectly possible natural languages. Hindi with the postlexical phonology of English would be no different in principle from Hindi with an English accent, or Hindi after some sound changes. In spite of the changes it would be a fully functional language and certainly could not be excluded on grounds of UG.

Here is what would happen to Hindi if the output of its lexical phonology were subject to English postlexical phonology (call the result Hinglish). Hinglish would have the same word stress as Hindi because English word stress is assigned in the lexical phonology and protected by faithfulness in the postlexical phonology. Sentence prosody (intonation, phrasing, and sentence stress), however, would be as in English. There is no incompatibility whatever between Hindi word stress and English sentence prosody as far as I can tell, so this should be a viable language.

The most noticeable peculiarity of Hinglish would be that a number of phonological contrasts of Hindi would be neutralized in it. Aspiration would be neutralized, and phonetic aspirated and unaspirated stops would appear in their English distribution. The opposition between dental and retroflex coronals would be neutralized. Gemination would be neutralized, at least foot-internally.

- (15) a. [sita:r] → [sit^hɑ:ɹ] ‘sitar’

- b. [g^hi:] → [gi:y] ‘ghee’
- c. [tikəʈ] → [t^hɪkɪt] ‘ticket’
- d. [t^hərmɑ:mi:təʀ] → [t^hɜ:ma:mi:yrəʃ] ‘thermometer’

These neutralizations would lead to the restructuring of underlying representations and of the lexical phonology. Therefore, the actual Hindi lexicon would not coexist with the actual English postlexical phonology. However, no formal restriction against combining the Hindi lexicon with the English postlexical phonology is needed in Stratal OT, any more than classic OT needs to prohibit underlying Hindi contrasts from the English lexicon. Lexicon Optimization guarantees that unrealized contrasts are not acquired.

The upshot is that Hindi with English postlexical phonology would not be very different from Hindi with a heavy English accent. It would have a lot more more homonyms than Hindi but it would not be in violation of any UG principles or even grossly dysfunctional. This should not be surprising because in real life, the postlexical phonologies of languages are subject to constant change by the process known as sound change, and languages manage just fine.

What if we instead took just the Hindi stem level and grafted English word and postlexical phonology onto it? I see no reason to believe that it would be an impossible language either. The sentence prosody would still be English-like, and essentially the same irrecoverable neutralizations would take effect, forcing massive simplification of the lexicon. No new contrasts would arise.

What about an extreme case, such as a language that has completely different stress systems at different levels? It could exist if both systems were detectable from the output robustly enough to allow learners to reconstruct them. Suppose the stem phonology has left-to-right feet with associated segmental and prosodic effects (e.g. lengthening of the strong beat of each foot), and the word phonology superimposes quantity-insensitive right-to-left trochees on them. Stratal OT predicts that such languages can exist, and in fact they do. Jarawara appears to be such a language (Dixon 2004). Dogrib has left to right iambic stress at early levels and right to left trochaic stress at later levels (Jaker 2012).⁹

In sum, rather than calling Stratal OT into question, such disparities between levels provide further support for it.

⁹Auca has strictly trochaic stress assignment at the stem level and tolerates clash only in the word phonology (Kim 2003).

The relative richness of Stratal OT and parallel OT Regarding the relative richness of the two theories, it should be clear by now that transderivational OT, with or without the additional constraint types that we will be discussing (namely O/O Correspondence constraints, Paradigm Uniformity constraints, B/R Correspondence constraints, and Optimal Paradigms constraints) is neither more restricted nor less restricted than Stratal OT; it is just different. As we have already shown, each of the theories has expressive capacity that the other lacks.

The empirical side of the argument is that the differences in expressive power are to the advantage of Stratal OT in both directions. On the one hand, the OT has the power to express linguistically impossible generalizations which are correctly excluded in Stratal OT. On the other hand, the limited intrinsic seriality of Stratal OT allows OT to express genuine generalizations which go by the boards in strictly parallel OT. I will show that considerations of descriptive and explanatory adequacy uniformly favor Stratal OT over strictly parallel OT with Sympathy and O/O constraints.

The specific form of the excess power argument is based in the claim of Benua 1997 that actual differences between levels are limited to the ranking of faithfulness constraints, something which, if true, Stratal OT would have to stipulate. McCarthy also condemns Stratal OT for having to stipulate a special status for faithfulness constraints, but coming virtually in the same breath as his own stipulation that only faithfulness constraints can select Sympathy candidates this criticism lacks conviction. Within level-ordering the stipulation could at least appeal to some independent empirical motivation: in their study of the different native and borrowed strata of Japanese vocabulary, Itô and Mester 1995a,b found that non-derivational levels also differ just in the ranking of faithfulness constraints, and proposed that reranking in stratally organized lexicons is limited to faithfulness constraints. Within Sympathy, it is an altogether singular restriction related to nothing else.

As far as the empirical side is concerned, I concur with Itô and Mester 1997 in doubting the claim on which Benua's based her attack on Stratal OT. While her conceptual arguments for O/O constraints are interesting and important, her claim that levels in Stratal OT differ *only* in the ranking of faithfulness constraints is not based on serious study of Stratal OT. Moreover, the serial model she criticizes bears no relation to any version of Stratal OT that has ever been proposed, to my knowledge. She seems to have improvised it from various (pre-OT) Stratal OT analyses, extrapolating from the Strong Domain Condition. But it should be obvious that it is unsafe to transpose results from rule-based analyses into constraint-

based analyses without rethinking them from the ground up.

As Itô and Mester state, the generalization that interstratal ranking differences are restricted to faithfulness is likely to be an emergent property of grammars rather than a hard-wired fact of UG. My own analysis of Arabic in chapter ?? confirms that levels can differ in other ways as well, in this case, in what syllable structures are allowed and in what measures are taken to repair impermissible ones. The ranking of faithfulness constraints is no doubt the most frequent source of grammar-internal variation, and that includes even level-internal variation in ranking (such as that studied by Anttila 1997). One possible explanation for this is that at least one of the major sources of both inter-level and intra-level ranking variation is the process of sound change. On OT assumptions, sound change must be the reranking of a markedness constraint to a position where it dominates all faithfulness constraints (but not necessarily all other markedness constraints). For Stratal OT in particular, this reranking must be initiated in the postlexical system. Of course, even supposing that the privileged status of faithfulness constraints were an unexplained stipulation about lexical strata, and even supposing it did not have an analog in Sympathy theory, it would still be less troublesome than the irredeemable shortcomings of Sympathy theory discussed above. At worst, it would be stipulative. Stipulations are sometimes superseded as understanding grows, but a theory with fundamental structural inadequacies requires complete retooling.

As far as excess power is concerned, the shoe is truly on the other foot. Considerations of restrictiveness, learnability, and explanatory adequacy uniformly favor Stratal OT over strictly parallel OT with Sympathy and O/O constraints. On these grounds I reject McCarthy's first argument against Stratal OT.

3.4 The trivialization argument

Calabrese (2005: 460) objects that the Lexical Phonology notion of a stratum

“is trivialized by the kind of stratum that, say, Tiberian Hebrew or Levantine Arabic would require — a stratum of convenience rather than a meaningful correlation of phonological and morphological factors.”

This criticism¹⁰ is misdirected. The fragment of Levantine Arabic phonology in Kiparsky 2000 illustrates how the levels are not “trivial” or arbitrary, but clearly distinguishable on phonological as well as on morphological grounds. As for Tiberian Hebrew, work by Rappaport 1984, Malone 1993 (LPM) and Koontz-Garboden 2001 (Stratal OT) does not justify Calabrese’s claim. He could have found the gist of the Levantine Arabic analysis demonstrating the morphological correlates of the phonological levels in Kiparsky 2000. In fact, he cites that article, and in a curious reversal, dismisses it solely on the authority of the untraced quotation whose point it effectively refutes. Stratal organization is correlated with domain of constraint application, mutual interaction of phonological constraints, and phonology/morphology interaction. The small number of levels by itself imposes stringent constraints on the depth of ordering. If Stratal OT errs, it is more likely that it is *too* restrictive. This is by design: it is simply methodological common sense to start with the simplest and strongest theory that has a chance of being right. The charge of “trivialization” is in any case baseless.

3.5 Theoretical simplicity

“**Extra machinery**” Calabrese (2005: 460-1) states:

“OT requires extra machinery to deal with Opacity. It is unclear if this extra machinery brings greater insights into the theory, other than a pure account of opacity. This contrasts with the situation in the classical derivational model where opacity is accounted for by assuming extrinsic rule ordering, a notion that is independently required in that model.”

Calabrese’s objection holds only for classic OT, which requires (trans)derivational constraints for opacity. The very point of Stratal OT is that nothing needs to be added to account for opacity that is not also needed to account for cyclicity, prosodic morphology, and ordinary morphology/phonology interactions.

Non-uniformity A slightly different version of this objection is that Stratal OT does not provide a uniform theory of opacity. It countenances several sources of

¹⁰Calabrese attributes this quote to “McCarthy 2000: 9-10”, but I can’t find it my copy of this work (which has no pages 9-10 anyway) or in other works by McCarthy.

opacity. As in classic OT, constraint domination captures the blocking relation between special and general processes, as well as special cases such as “cross-derivational feeding” opacity (Bakovic 2006). Cyclicity captures the retention of phonological properties of bases into derived words. Level ordering accounts for morphology/phonology interactions, and unifies them with the sorts of opaque interactions that stipulated rule ordering is invoked for in SPE-type theories.

Rule ordering theories and transderivational OT also treat opacity in heterogeneous ways. In ordered rule theories, opacity arises from blocking, cyclicity, and ordering. Parallel OT provides analogous tools: Sympathy constraints or PREC constraints, Output/Output constraints and/or Paradigm Uniformity constraints (as well as Base/Reduplicant constraints and perhaps others, such as Targeted constraints). In fact, there is no real formal or conceptual unity between classic OT’s original idea of ranked violable markedness and faithfulness constraints, and the various constraint families that have subsequently been added to it.

While unification is certainly a hallmark of a good theory, it is futile to expect every pretheoretical notion to correspond precisely to a single theoretical concept or to have a single explanation. SPE’s treatment of stress as a feature like any other provided a formally unified theory of phonological representation, but it was metrical and autosegmental phonology’s formal differentiation between stress as a hierarchical prominence relation and features like voicing that led to the real progress in prosodic phonology. Generative syntax would not get very far by just reconstructing traditional grammatical categories. Historical linguists are content to live without a “uniform theory of linguistic change”, preferring the sharper understanding that comes with distinguishing between sound change, analogy, and borrowing, each having specific properties and causal mechanisms which can be further grounded in perception, production, acquisition, and sociolinguistic realities. They would rightly be unimpressed by the following “theory” of linguistic change: “all linguistic change is constraint reranking”. Wonderfully unified, but, in a world where constraint ranking is the only locus of difference between grammars, trivially true. A more articulated theory rich in empirical predictions is preferable to a conceptually unified untestable one.

It is useful to recall here the ill-fated “best theory” argument that once was put forward in support of generative semantics. Advocates of that approach argued for its superiority on the grounds that it provided a uniform treatment of the correspondence between deep structure and surface structure by a very general mechanism of derivational constraints (Postal 1972). In a rebuttal that is as relevant today as it was then, Chomsky pointed out that predictive content and

explanatory power trump conceptual minimality.

“Thus it is misleading to say that a better theory is one with a more limited conceptual structure, and that we prefer the minimal conceptual elaboration, the least theoretical apparatus. Insofar as this notion is comprehensible, it is not in general correct. If enrichment of theoretical apparatus and elaboration of conceptual structure will restrict the class of possible grammars and the sets of derivations generated by admissible grammars, then it will be a step forward (assuming it to be consistent with the requirement of descriptive adequacy). It is quite true that the burden of proof is on the person who makes some specific proposal to enrich and complicate the theoretical apparatus. One who takes the more ‘conservative’ stance, maintaining only that a grammar is a set of conditions on derivations, has no burden of proof to bear because he is saying virtually nothing.” (Chomsky 1972: 68)

In the opposite direction, even the unity of the type of opacity variously dealt with by stipulated rule ordering, Sympathy, and stratification has been questioned. Itô and Mester’s (2003) remark that the “quest for a catchall mechanism able to deal with all facets of opacity” might even be “in principle incorrect” is intended to apply to just this type. This lies behind the many partial accounts of opacity phenomena in the literature, using proposed devices such as Local conjunction (Kirchner 1996, Bakovic 2000) for counterfeeding, Łubowicz 2002 for counterbleeding), Targeted constraints (Wilson *Phonology* 18, 2001 Bakovic 2000), Turbidity (Goldrick 2000), Scalar faithfulness (Gnanadesikan 1997), Comparative Markedness (McCarthy *Theoretical Linguistics* 29, 2003), Contrast Preservation constraints (Łubowicz *WCCFL* 22, 2003), and conjunction of markedness and faithfulness constraints (Itô and Mester (2003). The project of dividing and conquering opacity by independently motivated principles is worth attempting, but fails empirically since Stratal OT is independently motivated and achieves better descriptive coverage than any combination of partial solutions as yet tried.

To summarize: while genuine conceptual unification and simplicity of theoretical apparatus are important desiderata, they should in the first place not be mistaken for each other, and secondly, neither of them can redeem a theory that has nothing interesting to say about its subject matter or one that makes massively false claims about it.

References

- ANTTILA, ARTO. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20: 1-42.
- ANTTILA, ARTO. 2006. Variation and opacity. *Natural Language and Linguistic Theory*. 24: 893-944.
- ANTTILA, ARTO, VIVIENNE FONG, STEFAN BENUS, AND JENNIFER NYCZ. 2008. Variation and opacity in Singapore English consonant clusters. *Phonology* 25.
- BAKOVIĆ, ERIC. 2000. *Harmony, Dominance, and Control*. Doctoral Dissertation, Rutgers University, New Brunswick, NJ.
- BAKOVIĆ, ERIC. 2011. Opacity and ordering. John A. Goldsmith, Jason Riggle, and Alan C. L. Yu (eds.) *The handbook of phonological theory*, 2nd ed.
- BENUA, LAURA. 1997. *Transderivational identity: Phonological relations between words*. Ph.D. Dissertation, U. Mass.
- BERMÚDEZ-OTERO, RICARDO 1999. *Constraint interaction in language change: quantity in English and Germanic*. Ph.D. Dissertation, University of Manchester.
- BERMÚDEZ-OTERO, RICARDO. 2001. *Underlyingly Nonmoraic Coda Consonants, Faithfulness, and Sympathy*. <http://www.bermudez-otero.com/DEP-mora.pdf>
- BERMÚDEZ-OTERO, RICARDO. 2006. Phonological change in Optimality Theory, in Keith Brown (ed.), *Encyclopedia of language and linguistics* (2 edn, vol. 9). Oxford: Elsevier, 497-505.
- BERMÚDEZ-OTERO, RICARDO. 2007. Diachronic phonology. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 497-517. Cambridge: CUP.
- BERMÚDEZ-OTERO, RICARDO. 2010. Currently available data on English t/d-deletion fail to refute the classical modular feedforward architecture of phonology. Paper given at the 18th Manchester Phonology Meeting, 20 May 2010. Handout at www.bermudez-otero.com/18mfm.pdf.
- BERMÚDEZ-OTERO, RICARDO. 2011. 'Cyclicity', in Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice (eds), *The Blackwell companion to phonology* (vol. 4: Phonological interfaces). Malden, MA: Wiley-Blackwell, 2019-48.

- BERMÚDEZ-OTERO, RICARDO. Forthcoming. The architecture of grammar and the division of labour in exponence, in Jochen Trommer (ed.), *The morphology and phonology of exponence: the state of the art*. Oxford: Oxford University Press.
- BERMÚDEZ-OTERO, RICARDO AND RICHARD M. HOGG. (2003). The actuation problem in Optimality Theory: phonologization, rule inversion, and rule loss, In D. Eric Holt (ed.), *Optimality Theory and language change*. Dordrecht: Kluwer, 91-119.
- BLUMENFELD, LEV. 2003. Russian palatalization in Stratal OT: Morphology and [back]. In W. Browne, J.-Y. Kim, B. H. Partee, and R. Rothstein (Eds.), *Proceedings of the Annual Workshop on Formal Approaches to Slavic Linguistics 11*, 141–158, University of Massachusetts, Amherst. Michigan Slavic Publications.
- BOERSMA, PAUL & DAVID WEENINK. 2007. Praat: doing phonetics by computer. Software version 4.6.09. Available at www.praat.org.
- BOOIJ, GEERT. 1996. Lexical Phonology and the derivational residue. In Durand, Jacques & Bernard Laks (eds.) *Current Trends in Phonology: Models and Methods*. European Studies Research Institute and University of Salford.
- BOOIJ, GEERT. 1997. Non-derivational phonology meets Lexical Phonology. In Iggy Roca (ed.) *Derivations and constraints in the lexicon*. Oxford: Clarendon Press.
- BOROWSKY, TONI. 1993. On the Word Level. In Sharon Hargus and Ellen Kaisse (eds.) *Studies in Lexical Phonology*. pp. 199-234. New York: Academic Press.
- BROME, MICHAEL. 1974. The cycle in phonology: stress in Palestinian, Maltese and Spanish. *Linguistic Inquiry*, 5.39-60.
- BURZIO, LUIGI. 1996. Surface constraints versus underlying representation. In Durand, Jacques & Bernard Laks (eds.) *Current Trends in Phonology: Models and Methods*. European Studies Research Institute and University of Salford.
- CHOMSKY, NOAM AND MORRIS HALLE. 1968. *The sound pattern of English*. New York: Harper & Row.
- CLARK, MARY. 1990. *The tonal system of Igbo*. Dordrecht: Foris.

- CLEMENTS, G. N. 1997. Berber syllabification: derivations or constraints? In I. Roca (ed.) *Constraints and derivations in phonology*, 289-330. Oxford: Oxford University Press.
- COETZEE, ANDRIES AND JOE PATER. 2008. The place of variation in phonological theory. To appear in John Goldsmith, Jason Riggle, and Alan Yu (eds.), *The handbook of phonological theory* (2nd ed.). Blackwell.
- COHN, ABIGAIL AND JOHN MCCARTHY. 1994[1998]. Alignment and parallelism in Indonesian phonology. Amherst ScholarWorks: University of Massachusetts.
- COLLIE, SARAH. 2007. English stress preservation and Stratal Optimality Theory. Ph.D. Thesis, University of Edinburgh.
- DONEGAN, PATRICIA AND DAVID STAMPE. 1979. The study of Natural Phonology. In Dinnsen, Daniel (ed.) *Current Approaches to Phonological Theory*. Bloomington: Indiana University Press.
- FILE-MURIEL, RICHARD J. 2004. An OT approach to vowel height harmony in Brazilian Portuguese. <https://www.indiana.edu/~iulcwp/pdfs/04-file.pdf>.
- GIEGERICH, HEINZ J. 1999. Lexical strata in English: morphological causes, phonological effects. *Cambridge Studies in Linguistics* 89. Cambridge: CUP.
- GNANADESIKAN, AMALIA. 1997. Phonology with ternary scales. <http://roa.rutgers.edu/files/195-0597/195-0597-GNANADESIKAN-7-0.PDF>
- GOLDRICK, MATTHEW. 2000. "Turbid Output Representations and the Unity of Opacity". In: Hirotani, Masako (ed.) *Proceedings of the North East Linguistics Society* 30. Amherst, MA: GLSA Publications, pp. 231-246.
- GOLDSMITH, JOHN. 1976. An overview of autosegmental phonology. *LA* 2: 23-68.
- GOLDSMITH, JOHN. 1990. *Autosegmental and Metrical Phonology*. Oxford and Cambridge, MA: Blackwell.
- GOLDSMITH, JOHN. 1993. Harmonic phonology. In J.A. Goldsmith (ed.) *The last phonological rule: Reflections on constraints and derivations*.
- GRIBANOVA, VERA . 2008. Russian Prefixes, Prepositions and Palatalization in Stratal OT. In Charles B. Chang and Hannah J. Haynie, *Proceedings of the 26th West Coast Conference on Formal Linguistics*. Cascadilla Proceedings Project.

- HALE, M. & KISSOCK, M. 1998. The Phonology-Syntax Interface in Rotuman. In Pearson, M., editor, *Recent Papers in Austronesian Linguistics*, volume 21. UCLA Department of Linguistics.
- HALLE, MORRIS, AND K.P. MOHANAN. 1985. Segmental phonology of Modern English. *Linguistic Inquiry* 16, 57–116.
- HARRIS, JAMES W., AND ELLEN M. KAISSE. 1999. Palatal vowels, glides and obstruents in Argentinean Spanish. *Phonology* 16: 117–190.
- HAYES, BRUCE, BRUCE TESAR, AND KIE ZURAW. 1993. OTSoft 2.1, software package, <http://www.linguistics.ucla.edu/people/hayes/otsoft/>.
- INKELAS, SHARON, ORHAN ORGUN, AND CHERYL ZOLL. 1994. Exceptions and static phonological patterns: cophonologies vs. prespecification. ROA-124. Rutgers Optimality Archive, <http://ruccs.rutgers.edu/roa.html>.
- INKELAS, SHARON, ORHAN ORGUN, AND CHERYL ZOLL. 1997. The implications of lexical exceptions for the nature of grammar. In I. Roca (ed.), *Derivations and Constraints in Phonology*, 393-418. Oxford University Press.
- ITÔ, JUNKO & ARMIN MESTER. 1995a. Japanese phonology. In John Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass.: Blackwell.
- ITÔ, JUNKO & ARMIN MESTER. 1995b. The core-periphery structure of the lexicon and constraints on reranking. In Jill N. Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*, pp. 181-209. University of Massachusetts Occasional Papers, 18; GLSA, UMass Amherst.
- ITÔ, JUNKO & ARMIN MESTER. 1999b. The Phonological Lexicon. In N. Tsujimura (ed.) *The Handbook of Japanese Linguistics*, 62-100. Oxford: Blackwell.
- ITÔ, JUNKO & ARMIN MESTER. 2002. Lexical and Postlexical Phonology in Optimality Theory. In Fanselow, Gilbert and Caroline Féry (eds.) *Resolving conflicts in grammars*, 183-207. Hamburg: Buske.
- ITÔ, JUNKO & ARMIN MESTER. 2003a. Lexical and postlexical phonology in Optimality Theory: Evidence from Japanese. In G. Fanselow and C. Féry (eds.), *Resolving conflicts in grammars*. *Linguistische Berichte, Sonderheft* 11.

- ITÔ, JUNKO & ARMIN MESTER. 2003b. Japanese morphophonemics: Markedness and word structure. Cambridge, Massachusetts: MIT Press Linguistic Inquiry Monograph Series 41.
- JAKER, ALEX. 2011. Prosodic Reversal in Dogrib (Weledeh Dialect.) Ph.D. Dissertation, Stanford University.
- JOSHI, S. D. & P. KIPARSKY. 1978. Siddha and asiddha in Paninian phonology. In D. Dinnsen (ed.), *The Differentiation of Phonological Theories*. Indiana University Press.
- KAGER, RENÉ. 1999a. *Optimality Theory*. Cambridge: CUP.
- KAISSE, ELLEN. 1985. *Connected speech: The interaction of syntax and phonology*. Orlando, Florida: Academic Press.
- KAISSE, ELLEN. 1990. Toward a typology of postlexical rules. In Sharon Inkelas and Draga Zec (ed.) *The phonology-syntax connection*, 127-143. Stanford: CSLI.
- KAPLAN, AARON. 2008. Pretonic non-prominence in Chamorro umlaut. http://people.ucsc.edu/~afkaplan/files/ocp_hndt.pdf
- KAVITSKAYA, DARYA, & PETER STAROVEROV. 2010. When an interaction is both opaque and transparent: the paradox of fed counterfeeding. *Phonology* 27: 255–288.
- KIM, HYO-YOUNG. 2003. An OT account of Auca stress. *Language Research* 39: 337-354.
- KIPARSKY, PAUL. 1971. *Historical Linguistics*. In W.O. Dingwall (ed.) *A Survey of Linguistic Science*. College Park, Maryland: University of Maryland Press. Reprinted by Greylock Publishers, 1978.
- KIPARSKY, PAUL. 1982. *Lexical morphology and phonology*. I.-S. Yang (ed.) *Linguistics in the morning calm*. Seoul, Hanshin.
- KIPARSKY, PAUL. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook* 2.82-138.
- KIPARSKY, PAUL. 1998. Covert Generalization. In Geert Booij, Angela Ralli, and Sergio Scalise (edd.) *Proceedings of the First Mediterranean Morphology Meeting*. University of Patras.
- KIPARSKY, PAUL. 1993. Blocking in Non-derived Environments. In Ellen Kaisse and Sharon Hargus (eds.) *Lexical Phonology*. New York: Academic Press.
- KIPARSKY, PAUL . 2000b. Opacity and cyclicity. *The Linguistic Review* 17:351-367.

- KIPARSKY, PAUL. 2003. Finnish noun inflection. In Diane Nelson and Satu Manninen (eds.) *Generative Approaches to Finnic Linguistics*. Stanford: CSLI.
- KIPARSKY, PAUL. 2011. Compensatory lengthening. In Cairns, Charles & Eric Raimy, *Handbook of the Syllable*. Leiden: Brill.
- KIPARSKY, PAUL. 1982. Lexical morphology and phonology. I.-S. Yang (ed.) *Linguistics in the morning calm*. Seoul, Hanshin.
- KIPARSKY, PAUL. To appear. *Paradigms and opacity*. CSLI Press.
- KOONTZ-GARBODEN, ANDREW. 2001. Tiberian Hebrew spirantization and related phenomena in stratal OT. <http://roa.rutgers.edu/files/607-0703/607-0703-KOONTZ-GARBODEN-0-3.PDF>.
- LAKOFF, GEORGE. 1993. Cognitive phonology. In J.A. Goldsmith (ed.) *The last phonological rule: Reflections on constraints and derivations*.
- LIBERMAN, MARK & ALAN PRINCE. 1977. On Stress and Linguistic Rhythm. *Linguistic Inquiry* 8: 249-336.
- ŁUBOWICZ, A. 2002. Derived Environment Effects in Optimality Theory. *Lingua* 112:243–280.
- ŁUBOWICZ, A. 2003. Contrast preservation in phonological mappings. Ph.D. dissertation, University of Massachusetts, Amherst.
- MCCARTHY, J., AND A. PRINCE. 1993a. *Prosodic Morphology 1986*. Technical Reports of the Rutgers Center for Cognitive Science.
- MCCARTHY, J., AND A. PRINCE. 1993b. *Prosodic Morphology I. Constraint Interaction and Satisfaction*. Ms. University of Massachusetts, Amherst and Rutgers University, New Brunswick. To appear, MIT Press.
- MCCARTHY, JOHN J., AND ALAN S. PRINCE. 1994. The Emergence of the Unmarked. In Mercè González (ed.) *NELS 24*, Vol. 2, 333-379.
- MCCARTHY, JOHN. 1999. Sympathy and phonological opacity. *Phonology* 16.331-399.
- MCCARTHY, JOHN. 1999. Sympathy and phonological opacity. *Phonology* 16.331-399.
- MCCARTHY, JOHN J. 2002. Comparative markedness. In Angela C. Carpenter, Andries W. Coetzee, and Paul de Lacy, eds., *Papers in Optimality Theory II* (University of Massachusetts Occasional Papers in Linguistics 26). Amherst, MA: GLSA. Pp. 171-246.

- MCCARTHY, JOHN J. 2003. "Sympathy, Cumulativity, and the Duke-of-York Gambit". In: Caroline Féry & Ruben van de Vijver (eds.). *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press, pp. 23-76.
- MCCARTHY, JOHN J. 2007. *Hidden Generalizations: Phonological Opacity in Optimality Theory*. *Advances in Optimality Theory*. London, UK: Equinox.
- MCCARTHY, JOHN AND ALAN PRINCE. 1995a. Faithfulness and reduplicative identity. In Jill N. Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*. University of Massachusetts Occasional Papers, 18, p. 249-384. GLSA, UMass Amherst.
- MCCARTHY, JOHN AND ALAN PRINCE. 1995b. Prosodic Morphology. In John Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass.: Blackwell.
- MCCARTHY, JOHN AND ALAN PRINCE. 1995. Prosodic Morphology. In John Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass.: Blackwell.
- MCHUGH, B.D. The phrasal cycle in Kivunjo Chaga tonology. In Sharon Inkelas and Draga Zec (ed.) *The phonology-syntax connection*. Stanford: CSLI.
- MOHANAN. K.P. 1982. *Lexical phonology*. Ph.D. Dissertation, MIT.
- MUTAKA, NGESSIMO. 1994. *The Lexical Tonology of Kinande*. Lincom Europa.
- ODDEN, DAVID. 1996. *The phonology and morphology of Kimatuumbi*. Oxford: Clarendon Press.
- ORGUN, ORHAN. 1996. *Sign-based phonology and morphology, with special attention to optimality*. Ph.D. dissertation. University of California, Berkeley.
- PADGETT, JAYE. 2003. Contrast and Post-velar Fronting in Russian. *Natural Language and Linguistic Theory* 21, 39-87.
- PAK, MARJORIE. 2005. Explaining branchingness in phrasal phonology. In John Alderete, Chung-hye Han, and Alexei Kochetov, *WCCFL* 24: 308-316.
- PAK, MARJORIE, AND MICHAEL FRIESNER. 2006. French phrasal phonology in a derivational model of PF. *NELS* 36.
- PATER, JOE. 2010. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker, (ed.) *Phonological Argumentation: Essays on Evidence and Motivation*. London: Equinox. 123-154.
- PESETSKY, DAVID. 1979. *Russian morphology and lexical theory*. MS, MIT. <http://web.mit.edu/linguistics/people/faculty/pesetsky/russmorph.pdf>.

- POTTS, CHRISTOPHER & GEOFFREY K. PULLUM. 2002. Model theory and the content of OT constraints. *Phonology* 19: 361–393.
- PRINCE, ALAN, & PAUL SMOLENSKY. 1993/2004. Optimality Theory: Constraint interaction in generative grammar. RuCCS Technical Report 2, Rutgers University, Piscataway, NJ: Rutgers University Center for Cognitive Science. Revised version published 2004 by Blackwell. Page references to the 2004 version.
- RIGGLE, JASON, MAX BANE, & SAMUEL BOWMAN. 2011. PyPhon. <http://clml.uchicago.edu>
- ROBERTS-KOHNO, RUTH. 1998. Empty root nodes in Kikamba: Conflicting evidence and theoretical implications. In Maddieson, Ian, and Thomas J. Hinnebusch (ed.) *Language description and linguistic history in Africa*. Trenton, NJ : Africa World Press.
- ROCA, IGGY. 2005. Strata, yes, structure-preservation, no. In Geerts, Twan, Ivo van Ginneken, and Haike Jacobs (eds.) *Romance languages and linguistic theory 2003*. Amsterdam: Benjamins.
- RUBACH, JERZY. 1984a. *Cyclic and Lexical Phonology: the structure of Polish*. Dordrecht: Foris.
- RUBACH, JERZY. 1984b. Segmental rules of English and Cyclic Phonology *Language* 60: 21-64.
- RUBACH, JERZY. 1997. Extrasyllabic consonants in Polish: derivational Optimality Theory. In I. Roca (ed.), *Derivations and Constraints in Phonology*, 551-581. Oxford University Press.
- RUBACH, JERZY. 2000. Glide and glottal stop insertion in Slavic Languages: a DOT analysis. *Linguistic Inquiry* 31: 271-317.
- RUBACH, JERZY. 2003. Duke-of-York Derivations in Polish. *Linguistic Inquiry* 34.
- SANDERS, NATHAN. 2003. *Opacity and Sound Change in the Polish Lexicon*. UCSC Ph.D. Dissertation. <http://roa.rutgers.edu/view.php3?id=809>
- SASAKI, KAN. 2008. Hardening alternations in the Mitsukaido dialect of Japanese. *Gengo Kenkyu* 134: 85: 118.
- SCHEER, TOBIAS. 2011. *A Guide to Morphosyntax-Phonology Interface Theories. How Extra-Phonological Information is Treated in Phonology since Trubetzkoy's Grenzsignale*. Berlin: Mouton de Gruyter.

- STAUBS, ROBERT, MICHAEL BECKER, CHRISTOPHER POTTS, PATRICK PRATT, JOHN J. MCCARTHY. 2010. OT-Help 2.0. Software package. Amherst, MA: University of Massachusetts Amherst.
- STERIADE, DONCA. 1999. Lexical conservatism in French adjectival liaison. In Marc Authier, B. Bullock, and L. Reed (eds.), *Formal perspectives on Romance linguistics*. Amsterdam: Benjamins.
- STONHAM, JOHN. 2007. Nuuchahnulth double reduplication and Stratal Optimality Theory. *Canadian Journal of Linguistics* 52: 105-130.
- STONHAM, JOHN. 2008. Level ordering in Nootka. In K. Hanson & S. Inkelas (eds.), *The Nature of the Word: Essays in honor of Paul Kiparsky*. Cambridge, MA: MIT Press.
- TOLSKAYA, INNA KONSTANTINOVNA. 2008. Oroch vowel harmony. M.A. Thesis, University of Tromsø. <http://www.ub.uit.no/munin/handle/10037/1404>.
- VIAPLANA, JOAQUIM. 2005. Velar verbs and verbal classes in Catalan. *Catalan Journal of Linguistics* 4: 225-247.
- WILSON, COLIN. 2001. Consonant cluster neutralisation and targeted constraints. *Phonology* 18: 147-197.
- YUN, JIWON. 2008. A Stratal OT approach to a noun-verb asymmetry with respect to opacity in Korean. *Penn Linguistics Colloquium* 32. <http://www.ling.upenn.edu/Events/PLC/plc32/revised/yun.pdf>.
- YUN, JIWON. To appear. Noun-Verb Asymmetries in Korean Phonology. *WC-CFL* 27.