Capturing Phonological Shades Within and Across Languages

Capturing Phonological Shades Within and Across Languages

Edited by

Yuchau E. Hsiao and Lian-Hee Wee

Cambridge
Scholars
Publishing

Capturing Phonological Shades Within and Across Languages

Edited by Yuchau E. Hsiao and Lian-Hee Wee

This book first published 2015

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-4438-7179-6
ISBN (13): 978-1-4438-7179-2

Yuchau E. Hsiao is Professor of Linguistics at the Graduate Institute of Linguistics, National Chengchi University, Taiwan. Yuchau is a theoretical phonologist with special foci on tone and prosody. He has won numerous awards in Taiwan in recognition for his achievements as a scholar and humanist.

Lian-Hee Wee is Associate Professor at the Department of English Language and Literature, Hong Kong Baptist University. His research is mostly in the field of theoretical phonology with a special interest in Asian Englishes and Chinese languages.

TABLE OF CONTENTS

CONTRIBUTORS

ALDERETE John, Simon Frasier Univ.
ARCHANGELI Diana, Univ. of Hong Kong and Univ. of Arizona
BLENKIRON Amber, Simon Frasier Univ.
BROSELOW Ellen, Stony Brook Univ.
CHEUNG Winnie H.Y., Hong Kong Baptist Univ.
CHUNG Raungfu, Southern Taiwan Univ. of Science and Technology
DONG Yan, Univ. of Michigan
DUANMU San, Univ. of Michigan
HSIAO Yuchau E., National Chengchi Univ., Taiwan
HUANG Hui-chuan J., Academia Sinica, Taiwan
HUANG Tzu-chuan, National Chengchi Univ., Taiwan
HWANG Jiwon, Stony Brook Univ.
INKELAS Sharon, Univ. of California, Berkeley
KIPARSKY Paul, Stanford Univ.
LIN Hui-shan, National Taiwan Normal Univ.
LIN Yen-Hwei, Michigan State Univ.
LU Yu-an, National Chiao Tung Univ.
LO Chin-cheng, National Cheng Kung Univ., Taiwan
MYERS James, National Chung Cheng Univ., Taiwan
OTA Mitsuhiko, Univ. of Edinburgh
OU Shu-chen, National Sun Yat-Sen Univ., Taiwan
PULLEYBLANK Douglas, Univ. of British Columbia
WEE Lian-Hee, Hong Kong Baptist Univ.

PREFACE AND INTRODUCTION

A PHONOLOGICAL TRIO: ARCHITECTURE, PATTERNS AND LANGUAGE

YUCHAU E. HSIAO AND LIAN-HEE WEE

This volume is organized according to three different centers of attraction for phonological studies: (i) Architecture of phonological research; (ii) Phonological Patterning and (iii) Crossing language boundaries.

Part I is a selection of papers that deal with framework issues overarching approaches to phonological research. Kiparsky's paper explains that though optimality theoretic models are essentially correct, the insights from lexical phonology where morpho-phonological interweaving results in restricted depths of opacity must be preserved by stratification of H-EVAL. Inkelas re-interprets positional faithfulness and other effects of special faithfulness as due to differences in confidence scales, finding grounding in psychology. This offers intriguing new perspectives to issues like neutral vowels in vowel harmony, and non-derived environment blocking. Perhaps the most shocking idea is to be found in Archangeli & Pulleyblank invitation to analyze tonal alternations in Kinande as a simple matter of allomorphic selective. Though simple in itself, Archangeli & Pulleyblank show us how this eliminates the need for postulating underlyforms, which sets off a cascade of phonological rules and the innate language faculty falling like dominos. Myer's contribution here is in the incorporation of speeded acceptability judgments as an aspect for consideration in determination of models, in this case the model of the Mandarin syllable. Lo & Chung's contribution invites us to rethink the nature of constraint families in OT, so that a family of constraints may be defined by a distinctive feature (say, nasality) rather than by a principle (say, the OCP). These papers thus tell us less about what is the right analysis of a given phenomenon, but rather impress upon the reader on how to go about analyzing any given phenomenon.

Part II is a collection of papers that zoom in on phonological patterns that call for an explanation. Hsiao's analysis of Taiwanese tones argues that chain-shifts are the result of comparative markedness (which is an optimality theoretic interpretation of derived environment effects). T. Huang's treatment of blocking in Shanghai tone-spreading invokes the relevance of syntax, prosody and tone complexity. H-S Lin's study of reduplication in Tsou deals with patterns variation and argues for an EVAL module in OT that is dichotomized: constraints which violations must be in satisfaction of higher-ranked constraints and those which violations are generally tolerated. Through careful data study, Duanmu & Dong revives a Chinese puzzle by demonstrating that the orthodox understanding behind the preference for disyllabic words in Chinese cannot be attributed to homophone avoidance. H. Huang similarly questions accepted belief that /z/ is phonemic in Squliq Atayal, when her fieldwork data show that the alleged phoneme is complementary to [j] in some instances and in free variation in others. Blenkiron & Alderete explain that in Rotuman, reduplicants need to meet minimal word requirements and that interacts intricately with the trochaic stress system of the language.

Evidently, languages also come into contact, so that distinct phonological systems might collide to yield interesting insights. Part III is dedicated to this rather more complex issue of system interaction with Broselow's and Y-H Lin's study of loan phonology. Wee & Cheung provides a historical perspective on how modern Hong Kong English tonal patterns are coaxed out of Cantonese interpretations of English stress, a theme echoed in Ou & Ota's study of meter from the perspective of second language acquisition. The mysteries of acquisition are deepened in Lu & Hwang's rather curious experimental results where despite greater perceptual salience for onsets of target languages, for a given target segment such as [ð] second language learners also seem to make rather marked choices in codas.

Although we have arranged the papers along these three themes, it would be rather naïve to consider each paper as narrowly fitting into one category. In reality, each paper will touch on all three aspects albeit with varying weight. Our classification merely serves to highlight our interpretations on what is most valuable among all that is valuable in each paper. In so doing, the editors hope to put together a phonological trio that will unravel some phonological shades within and across languages.

Architecture

Kiparsky's contribution provides the first chapter of this volume. Since the advent of Optimality Theory in the early 1990s, analyses appear to favor morpho-phonological alignment in the treatment of effects that arise from stratification between different levels of morphology. Consequently, proponents of OT have been at pains in grappling with opacity effects otherwise easily captured in Lexical Phonology. Kiparsky provides detailed and intricate arguments in this paper on how and why Stratal OT is the right framework. Essentially, Stratal OT recognizes three strata: stem, word and phrase. With each of these strata having its own H-EVAL component, one predicts not only the existence of opacity in phonological patterning, but also a limit to the depth of opacity.

In Chapter 2, Inkelas demonstrates how the use of a confidence scale leads to a more finely shaded picture than the usual Special Faithfulness » General Faithfulness. Such a scalar re-interpretation of faithfulness is psychologically grounded, and correctly predicts a range of attested effects beyond usual canonical examples. The paper provides fresh perspectives on long-standing problems related to neutral vowels in vowel harmony (Finnish), abstract contrasts in segmental quality (Kashaya) and to non-derived environment blocking effects. By predicting a range of attested outputs beyond the canonical, confidence scales offer a handle on variation in a model that preserves the "strict dominance of constraints" properties of classical OT (unlike stochastic models, Boersma 1998 and Boersma & Hayes 2001).

A tantalizing surprise awaits us in Chapter 3 as Archangeli & Pulleyblank demonstrate how the tonal patterns of Kinande can be accounted for without postulating underlying. Different surfaces forms of the same morpheme are simply treated as allomorphy, thus the possibility of phonological analysis without phonology, and by extension linguistic behaviors (such as learning and articulation) without an innate language faculty. General intelligence for selection would suffice in their "solely allomorphic" treatment. As a by-product of their model, problems of opacity become totally irrelevant since no alternation takes place at all. Archangeli & Pulleyblank's work reminds us of the early 1990s, when Goldsmith's (1993) call for the "Last Phonological Rule" was the fanfare heralding the arrival of Optimality Theory (Prince & Smolensky 1994/2004). OT took away the need for language-specific order of language specific rule application in favor of ranked universal constraints. Archangeli & Pulleyblank now want to take away underlying forms and the language faculty altogether. The chapter promises to be seminal, and it

is not hard to see how many phonological analyses that involve an input-output relationship can be recast in allomorphic terms.

Myer's (Chapter 4) recognizes three forces that influence phonological modeling of the syllable: (i) articulation, (ii) parsing and (iii) association strength. The case in point is the Mandarin medial which has been variously argued to be part of the onset or part of the rime. Using acceptability judgments, Myers demonstrates that articulation encourages grouping of medials with onset, while parsing encourages grouping with the rime. Association strength which may be determined by phonological dependency appears to modulate between (i) and (ii). Myer's experimentation required a more subtle kind of investigation that probes at speakers' speeded judgments to non-lexical syllables to reveal if medials interact more with rimes or onsets. In methodology, this is groundbreaking for the kinds of experimental phonology that impact on phonological theorizing.

Specific to a serialist interpretation of Optimality Theory, Lo & Chung (Chapter 5) argues for a very different notion of a constraint family based on a featural paradigm (revolving around [nasal], thus OCP[nasal], *LINKONE[nasal], etc] rather than a phenomenal paradigm (say, revolving around OCP, thus OCP[nasal], OCP[high], etc). Though this is established on a number of added assumptions about Taiwanese nasals (i.e. the relevance of local constraint conjunction and the nature of the domain for conjunction), the idea to rethink the notion of constraint families is a timely one and well-grounded within the logical possibilities of any theoretical framework that invokes both principles (i.e. rules, constraints, laws), domains (e.g. segment, phrase, constituent) and features (e.g. distinctive features, tone).

Patterns

Though theoretical architectures help us grapple with phonological patterns, ultimately, these ideas can only be vindicated by the kinds of phonological patterns that require analysis. Hsiao (Chapter 6) addresses a difficult problem of opacity exhibited by the circular chain-shift patterns of Taiwanese tone sandhi. Though well studied and reported, the chain-shifts have hitherto not found satisfactory treatment with parallelist Optimality Theoretic analyses. Through comparative markedness, Hsiao's analysis demonstrates that the nature of chain-shifts is reducible to derived environment blocking effects, hence forcefully weakening the motivations for transderivational antifaithfulness in phonological theorization.

In Chapter 7, T. Huang reports on the blocking of tone spreading in Shanghai validated through a microsurvey. The patterns of tone spreading are conditioned by a combination of prosody and the complexity of the input tone contour. The prosody is in turned determined by the interface with syntax. In terms of methodology, T. Huang's approach is a nod both in Kiparsky's direction (Chapter 1) and Myer's (Chapter 4). Nonetheless, it remains to be seen if the patterns here require the power of Stratal OT. Similarly, the constituency of function words to prosodic domains and the ambivalence of blocking dependent on the complexity of tone contours provide opportunities of applying Myer's brand of experimentation.

H-S Lin study of reduplication in Tsou (Chapter 8) is founded on fieldwork both hers and others before her. More interesting, H-S Lin captures variation in the reduplication patterns by appeal to cut-offs in the hierarchy of constraints, an idea attributed to Coetzee's (2006) Ranked-ordering Model of Eval and supported by Liang & Wee (2006). Also invoking Coatzee, Y.H. Lin (Chapter 13) shows that variation could come from input underspecification. Reduplication is an issue taken up also by Blenkiron & Alderete (chapter 11) where in the case of Rotuman, the issue is with the need for the reduplicant to be a minimal word in a language that requires trochaic feet.

In Chapter 9, Duanmu & Dong explores the pattern of disyllabicity in Chinese and offer a critical review of the orthodox account that disyllabic words in Chinese are favored over monosyllabic ones as a result of homophone avoidance. Using an authoritative dictionary, Duanmu & Dong demonstrated that homophone avoidance, despite its appeal, is a failed theory. Homophone avoidance predicts there to be a positive correlation between degree of homophony and degree of disyllabification. This prediction is not substantiated, and earlier research in support of the orthodox view turned out to be rather inconclusive when their methodologies were more carefully scrutinized. Careful scrutiny also seems to be the moral of H. Huang's story on the Atayal /z/ in Chapter 10. Though widely believed to be phonemic, fieldwork research by H. Huang points to /z/ as a quasi-phoneme because of its complementary and free variation with /j/. She thus advocates a more stringent study of phonological relationship between two segments that include seven criteria set by Hall (2013): (i) predictability of distribution, (ii) lexical distinction, (iii) native speaker judgment, (iv) alternations, (v) phonetic similarity, (vi) orthography, and (vii) place in the system/ phonological behavior.

A noteworthy aspect of many papers in Part II is the value of data, many rarely documented and studied. Among them, Tsou (2130 speakers in 2002, Ethnologue), Squliq Atayal (84,300 in 2002, but figure includes

all varieties of Atayal, Ethonolgoue) and Rotuman (9000 speakers in 1991, Ethnologue) are languages clearly requiring attention, though one may say the same for Taiwanese and Shanghainese as Chinese youths today lean more and more heavily towards Standard Chinese (Mandarin). Kinande still boasts of 903,000 speakers in 1991 (Ethnologue) but that's hardly a large number considering that Cantonese would easily claim more than 8 million native speakers in Hong Kong alone. On that note, let us move to Part III where we shall glimpse in particular at the contact between English and Chinese languages (Mandarin and Cantonese in particular).

Language

Ethnologue reports that there are 55.6 million English speakers in the U.K., of which about 1.5 million are second-language users. Worldwide the number of English speakers swells to 335.15 million, with perhaps 505 million if one includes second-language users. Chinese, however, has at least 1500 million speakers, a rather overwhelming figure. If one considers only Mandarin Chinese (Putonghua, Standard Chinese), Ethnologue reports a more modest but still stunning figure of 840 million in China (2000 census). There are probably lots of people today learning Chinese given China's current status as the world's second largest economy, but Chinese speakers have been grappling with English for a much longer time, going back to when the west was busy colonizing and trading (Wee & Cheung, chapter 14). Also, the world's largest economy is still the English-speaking United States of America. Little wonder therefore that the selection of papers in Part III will be large devoted to the matter of loanwords from English or acquisition of English as a second language.

Broselow's contribution (Chapter 12) looks at the typology of vowel insertion as a strategy when languages that forbid consonant clusters take in loanwords that have complex onsets. Repair via vowel epenthesis involves two interacting resolutions: choice and position. Essentially, insertions in pre-cluster position are done with an invariant default vowel, while insertions between two consonants are contextually determined. Broselow explains that pre-cluster (usually having obstruent-obstruent sequences) insertions are more likely to be true insertion that is part of the production grammar while insertions that split up obstruent-resonant clusters are subject to listeners' misinterpretation of the acoustics of the foreign language.

In Chapter 13, Y.H. Lin deals with a rather intricate problem where loanwords from English manifest themselves in varied forms in Standard Mandarin. When one thinks about it, given a particular input, say a word

from English, and a particular grammar, say Standard Mandarin, one would expect the output loan to be consistent with the grammar of Standard Mandarin. The variations of loanwords are thus unexpected and raise non-trivial questions on our understanding of how grammars actually work. Y.H. Lin's solution to the problem rightly queries the unsubstantiated assumption that loan inputs are fully borrowed into the host language. If loan inputs are underspecified, then it is likely that a range of outputs would be equally parsimonious given a particular ranked constraint hierarchy.

Offering a historical perspective, Wee & Cheung (Chapter 14) study of the Cantonese transliterations of English words as documented in the *Chinese-English Instructor* (1862) revealed a potential source of the tonal properties of modern Hong Kong English. Apparently, differences in prominence of English syllables were interpreted through different tonal heights of Cantonese syllables. The syllables that were "higher" in tone in transliteration eventually settled to become high-toned syllables in modern Hong Kong English, offering evidence of tonogenesis triggered by prosody rather than by deletion of coda consonants or by onset (de)voicing (Hyslop 2009).

However, does history square with language acquisition patterns? Ou & Ota (Chapter 15) look at how stress is acquired by Mandarin speakers of English and found that Mandarin speakers' acquisition of stress is not guided by mechanisms of metrical phonology. So, perhaps Wee & Cheung are right that stresses are interpreted as tones which might then impact on the how English is acquired to become varieties different from the source (British or US) in terms of tone and stress.

Looking deeper into the syllable, Lu & Hwang (Chapter 16) unearths the asymmetry between onsets and codas that impact on second language learning. Taiwanese Mandarin speakers articulate [l] in place of the English voiced interdental [ð] when used as an onset, but articulate [θ] when used as coda. Lu & Hwang takes a rather fresh perspective in the observation and framed the issues as why in the onset [l] is favored over less marked options like [t] when the target is [ð]; similar why coda [n] would have been a less marked option than [θ]. When contextualized in Mandarin Chinese where voiced obstruents are unattested, the choice of [l] in the onset suggests that learners perceive voicing despite [ð] being obstruent whereas the choice of [θ] for the coda suggest that the voicing was not perceived. One continues to wonder at how the interdental fricative could have surfaced in the coda when Mandarin does not have that segment in its inventory. However, like all research, efforts at uncovering explanations lead on to other questions. The hope of this

volume lies as much in presenting the shades of phonological patterning as in inviting others to add to the tapestry of phonological theorizing.

A short story

There is so much to write about phonology, that if this volume is successful in capturing the shades of phonological patterning, it must only be a very modest one, offering no more than a mise-en-scène of flowers sampled from the bloom of spring. The cast of authors in this volume is connected, among other ways, through the Theoretical Phonology Conferences (TPC) held first in 2005, then 2009, 2011 and 2013 at the National Chengchi University, Taipei (with Acadmia Sinica as co-host, excepting 2005, in all cases supported by the Ministry of Science and Technology and the Ministry of Education in Taiwan). The TPC series accepts international submissions and aims to provide a platform for formal phonology researchers and graduate students to exchange their ideas and new findings in theoretical studies in the generative framework. These conferences had no parallel sessions and were attended by more than 200 participants each time. In effect, TPC is a festival, bustling with animated phonological discussions. Keynote speakers who have graced the TPC include Paul Kiparsky, John McCarthy, Robert Ladd, Matthew Chen, Sharon Inkelas, Ellen Broselow, Duanmu San, Yen-hwei Lin, Bao Zhiming and Chilin Shih, all of whom also spent after conference hours with graduation students. Through this volume, we hope to spread some of the sensation that permeated the TPC.

The volume would not have happened if not for Carol Koulikourdi and Sam Baker of the Cambridge Scholars Publishing who approached us with the opportunity of getting these works in print. In putting this volume together, we are much indebted to Angela Carpenter, Anna Lubowicz, Frantisek Kratochvil, Jie Zhang, Karen Chung, Larry Hyman, Long Peng, Marc Oosterndorp, Maria Gouskova, Nancy Hall, Rina Kreitman, Seung-hun Lee, Srinivas Sampath Kumar, Tian-hsin Hsing, and Xu Zheng who provided helpful and incisive reviews to the papers. Several authors also pitched in on this very important, and often under-appreciated task: Duanmu San, Hui-chuan J. Huang, Hui-shan Lin, John Alderete, Shu-chen Ou and Yu-an Lu. On behalf of all the authors and also ourselves, thank you!

The editors would probably have gone blind while trying to set all the papers to the necessary consistency for style and format if not for the meticulous efforts of Liu Yang, Queenie K.Y. Chan and Srinivas Sampath Kumar. We couldn't afford to pay them and they graciously shrugged their shoulders and grimaced in retaliation.

References

Boersma, P. (1998) *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Ph.D. dissertation, University of Amsterdam. LOT International Series 11. The Hague: Holland Academic Graphics. Retrieved from http://www.fon.hum.uva.nl/paul/diss/.

Boersma, P. and B. Hayes (2001) Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-86.

Coetzee, A. W. (2006) Variation as accessing "non-optimal" candidates. *Phonology* 23: 337-385.

Ethnologue. http://www.ethnologue.com, accessed 31 May 2014.

Goldsmith, J. (1993) *The Last Phonological Rule*. Chicago: University of Chicago Press.

Hall, K. C. (2013) A typology of intermediate phonological relationships. *The Linguistic Review* 30(2): 215-275.

Hyslop, G. (2009) Kurtöp Tone: A tonogenetic case study. *Lingua* 119: 827-845.

Liang, Y. and L. H. Wee (2006) Beijinghua de lianshang biandiao. *Zhongguo Yuwen* 311(2): 151-163.

Prince, A. and P. Smolensky (1993/2004) *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell.

PART I:

THE ARCHITECTURE OF PHONOLOGICAL RESEARCH

CHAPTER ONE

STRATAL OT: A SYNOPSIS AND FAQS

PAUL KIPARSKY

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 & 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 sections 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 & Prince 1995a). 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 the form 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 as DERIVATIONAL OT and TRANSDERIVATIONAL OT respectively. Derivational and transderivational constraints are too powerful in some ways and too weak in others, and 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 2014; Kiparsky 2014), 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 accounts. 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 means true of all of them. Some provide quite compelling evidence in favor of Stratal OT over any ordered rule theory including

[1] If 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.

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-categorial domains stem, word, and 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 & Mester 1995b). A part of (2b) is sometimes implicitly or explicitly adopted in mainstream OT work in that lexical and post-lexical phonology are treated as separate constraint systems, with the output of the former providing inputs to the latter. (2c) is

a generalization of generative phonology's phonological cycle (Chomsky & 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).

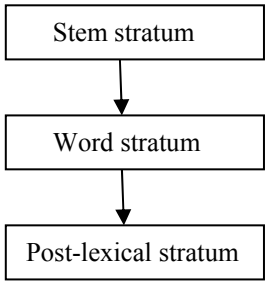
Stratal OT in turn can be implemented in several possible ways. Applying the theory requires commitment to a specific implementation, sometimes even beyond a point that can be empirically justified at present.

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 constraints, and provide the input to the word system, which in turn provides the input to the syntax and post-lexical 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

[2] The 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.

correspondence between the stem phonology and the post-lexical phonology.

(4)



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 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-categorial morphosyntactic units. A theory that conflates stem and word phonology, leaving only a lexical and post-lexical 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 post-lexical 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 post-lexical stratification comes from studies by Kaisse (1985, 1990), Kiparsky (1985), Clark (1990), McHugh (1990), Mutaka (1994), Koontz-Garboden (2001), Pak (2005), Pak and Friesner (2006). Cyclicity and opacity *within* post-lexical phonology would not only be consistent with such a post-lexical 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 Kimatuumbi (Odden 1996) and Dogrib (Jaker 2011). 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 the 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 post-lexical level than at the word level, it is undominated at the post-lexical 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 by derivatives from bases, 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 extensive and inventive OT sub-literature 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 refuting the idea that constraints evaluate output representations in parallel, 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 looks 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 & Zuraw 2003), the Praat OT workbench (Boersma & 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 and Pullum's (2002: 361) point that these constraint types can't be handled in a formal reconstruction of OT