Introduction

LINDA LOMBARDI

Optimality Theory (OT) (Prince and Smolensky 1993; McCarthy and Prince 1993) has rapidly become the dominant framework in formal phonological theory. In OT fundamental notions of generative grammar are revised by replacing rules and derivations with interacting well-formedness constraints. The phonological literature immediately prior to OT had shown that such constraints were crucial to the construction of explanatory analyses. (Prince and Smolensky 1993 give nearly half a page of such references; limiting ourselves just to works on segmental phonology by authors represented in this volume, we can point to such works as Ito 1986, 1989; Lombardi 1991, 1995; Mester 1986; Yip 1988.) But incorporating such constraints into derivational theory posed a number of problems.

One problem was that it was difficult to make a formal connection between constraints and the rules that accounted for the related alternations. (See Prince and Smolensky 1993, ch. 10, for discussion of some attempts.) It has long been observed that many languages exhibit phonological conspiracies (Kisseberth 1970): that is, there are different rules that have the effect of making forms obey the same surface well-formedness constraints. And many languages have been observed to have a kind of duplication problem: the constraints needed to restrict underlying forms and the rules that give the correct surface forms have the same effects. But in derivational theory the rules and the constraints were stated separately from each other, and also stated separately on the different levels, and were not formally linked.

Another difficulty was the issue of universality. The assumption that the constraints were constructed anew in each grammar, like rules, was unsatisfactory since the same constraints tended to recur in many different languages. But most constraints were not obeyed in all languages,

seemingly contradicting the idea that they could be part of Universal Grammar (UG). (For discussion of this in the literature see, for example, the debate on the Obligatory Contour Principle (OCP) (Odden 1986, 1988; McCarthy 1986).)

Given that rules alone were not sufficient to provide explanatory analyses, but that rules and constraints were difficult to combine formally, in hindsight it may seem to have been an obvious alternative to try using constraints alone. But how could this work? Constraints could describe static patterns, but what about phonological alternations?

Prince and Smolensky's proposals made it possible to construct the grammar entirely from constraints in a way that both allowed constraints to account for alternations and to be universal. Prince and Smolensky achieve this by claiming that constraints are violable and ranked: although all languages have a given constraint in their grammar, that constraint may be violated if necessary to obey a higher-ranked constraint. Constraint ranking and constraint violability thus allow language variation while using universal constraints: when two constraints make conflicting demands, languages differ in how the conflict is resolved. And one of the basic types of constraint conflict is that between markedness and faithfulness: when violation of faithfulness is forced, forms change, giving us phonological alternations.

The original work in OT (cited earlier) concentrated mainly on prosodic phonology and on the phonology-morphology interface. The clearest and perhaps most convincing success was Prince and Smolensky's analysis of crosslinguistic syllable structure requirements. Prince and Smolensky showed that these patterns were due to the interaction of a limited set of markedness constraints, which defined the well-formed structures, and faithfulness constraints, which regulated the extent of possible differences between underlying and surface forms. This analysis is a model for all work in the framework.

The immediate impact of Prince and Smolensky's syllable structure analysis was partly due to the fact that it was clear to all phonologists that there is indeed a restricted, well-attested pattern of syllable structure requirements across languages; their basic observations go back to Jakobson. And at the point when OT was introduced, prosodic phenomena were already being spoken of largely in constraint-based terms even in derivational phonology, using templates as constraints on prosodic shape (for example, Ito 1986 for syllables and McCarthy 1981 for prosodic morphology). Thus, these cases were ripe for translation into a framework where constraints are primary, and where a crucial limiting factor is that reranking of the proposed constraints should result in the appropriate crosslinguistic typology.

In contrast, segmental phonology was in a rather different position. Many analyses (like those cited earlier) did invoke constraints. But many phenomena were still generally analyzed exclusively in terms of rewrite rules. It was also much less clear what the restrictions were on the crosslinguistic typology of segmental processes. Segmental feature-changing alternations appear to show a good deal more variation than prosodic patterns do. There are simply more different features than there are prosodic constituents, and they do not all obviously have the same behavior. Vowel features and nasality assimilate long-distance; in contrast, Voice and Place assimilate only locally, but may dissimilate long-distance. Consonant features usually spread regressively, except that many languages have (progressive) postnasal voicing of obstruents. Neutralization of featural distinctions is common, but the context where it occurs differs depending on the particular feature involved. In sum, there are more kinds of interactions and patterns possible for features than there are for syllable structure constituents, and it was not as clear what the typological restrictions on their variation are, or how they can be accounted for.[1]

But although initially progress in OT phonology was concentrated in prosodic phenomena, there is now a body of work and a core research community that focuses on working out the details of featural phonology in the OT framework. This work shows that, with appropriate depth of analysis, there is in fact a restricted, universal set of possibilities for segmental phonology. The papers in this volume show us both what OT can tell us about segmental phonology, and what analysis of segmental phenomena has to contribute theoretically to OT.

The first part of this volume is "The Content of Representations." The tenets of OT, regarding constraint violability and ranking, make no particular claims about phonological representations. We could, for example, do OT with any kind of feature theory: SPE feature bundles or feature geometric representations, privative or binary features, and so on. Much early work in OT paid little attention to representational questions, simply taking over assumptions from previous work in derivational autosegmental phonology. But representational arguments are theory-internal and need to be reexamined in light of fundamental theoretical changes; the choice of correct representations to use in OT analyses must be based on arguments couched in OT terms. The chapters in the first part of the book make such arguments, and they show the additional insight into phonological patterns that can be gained by applying OT to these facts.

The first chapter, my "Why Place and Voice Are Different," gives an OT-internal argument that in order to explain certain asymmetries in

phonological patterns, Voice must be a privative feature, and the laryngeal consonants [h, ʔ] cannot be literally Placeless, but rather have a relatively unmarked Place feature. The former argument confirms earlier derivationally based work, but the latter contradicts it and shows that making this representational assumption within an OT grammar solves problems that were intractable in earlier frameworks.

Representational arguments are also central to Zoll's contribution, "Constraints and Representation in Subsegmental Phonology." This chapter is based on work in her dissertation, a crosslinguistic study of ghost segments and floating features that was one of the first in-depth OT studies of an area of featural phonology. Zoll shows that within OT, we can account for all of these phenomena with more restricted representational differences than those proposed in pre-OT analyses. Zoll's chapter shows that these subsegmental constituents submit to a more insightful and appropriately restrictive analysis in an OT framework.

Kirchner's chapter, "Phonological Contrast and Articulatory Effort," takes a novel approach to representations. He argues that phonological representations in an OT grammar can contain much richer phonetic detail than is normally assumed, without predicting unattested systems of contrasts. Constraints can then refer to more specifically phonetic properties than is otherwise possible. Using constraints regulating faithfulness to auditory features, and triggering constraints referring to articulatory effort and to a perceptually based fortition requirement, Kirchner gives a unified account of lenition, a segmental phenomenon that has been problematic for almost all previous phonological frameworks. This chapter thus represents an important alternative approach to segmental representations that has been made possible by the assumptions of OT.

In rule-based autosegmental theory, transparency and opacity phenomena were a crucial type of evidence for representational claims. Transparency effects were used to argue for lack of feature specifications (privative features or underspecification); opacity, in contrast, was often explained as due to the presence of a feature value blocking spread. These representational assumptions must be reexamined in light of our new theoretical assumptions. NíChiosáin and Padgett's chapter, "Markedness, Segment Realization, and Locality in Spreading," addresses the fundamental question of the locality of spreading, and what it means for a segment to be "transparent." Their paper is couched in terms of Dispersion Theory (Flemming 1995), an important line of research within OT segmental phonology. Similar to Kirchner's approach, the representations are assumed to contain rich phonetic

detail. In addition, an important difference between this framework and standard OT is that phonological contrasts arise directly from output constraints requiring the maintenance of contrasts; the distinctiveness of contrasts is also controlled by output constraints. Using these assumptions NíChiosáin and Padgett propose that at least some apparently long-distance spreading is actually local, and that the apparent fact of transparency can be an artifact of the phonological framework underlying the analysis.

Of course, all analyses must call on assumptions about both the form of phonological constraints and the form of representations, so along the way to their main points, these chapters also make theoretical contributions to topics in other parts of the book, particularly the second part on the form of phonological constraints. For example, along with arguments for the representations of voiceless sounds and laryngeal consonants, my chapter also argues for the existence of MaxFeature constraints and for the existence of CodaCond for Place but not for Voice. NíChiosáin and Padgett's chapter and Kirchner's as well, because of their different view of representations, must make novel claims about both faithfulness constraints and the markedness constraints triggering alternations.

The second part of the volume, "The Content of Constraints," addresses what can be seen as the central issue of what the content of UG is in OT. The basic tenets of OT are very general claims about constraint violability and interaction. They may imply the existence of the two basic conflicting categories of constraints, Markedness and Faithfulness, which give the conflict between well-formedness requirements and the constraints that penalize changes to input (although even this may not be required, as in Dispersion Theory, used in NíChiosáin and Padgett's chapter). But the specific formulation of the constraints can be worked out in many different ways. Many of the chapters in this volume assume the faithfulness theory of Correspondence (McCarthy and Prince 1995), for example, but within this theory there are still open questions about what kinds of correspondences are regulated by faithfulness relations. Similar questions arise perhaps even more strongly for the content of markedness constraints, which are the triggers of phonological alternations, since there is no general theory like Correspondence that constrains their form.

Pater's chapter, "Austronesian Nasal Substitution Revisited," is a continuation of research begun in his earlier paper, "Austronesian Nasal Substitution and other NÇ Effects" (Pater 1999). This paper was one of the earliest detailed treatments of the question of what constraints trigger a particular featural alternation, and what might be the explana-

tion for the existence of such constraints. In it, Pater proposes a constraint *NC̥ that penalizes a cluster of a nasal and a voiceless stop. He shows that such a constraint is phonetically motivated: it is grounded in the articulatory difficulty of such sequences compared to nasal–voiced stop clusters. The need for a constraint like *NC̥ to trigger simple postnasal voicing seems clear. But in the present volume, Pater reexamines the Austronesian nasal substitution data and suggests that the triggering factors for these more complex alternations must be seen in a different way.

Gerfen's chapter, "A Critical View of Licensing by Cue," addresses the question of the form of the constraints licensing contrast, which are crucial to achieving positional neutralization of phonological distinctions in OT analyses. As shown by the chapters by Kirchner and by NíChiosáin and Padgett in the first part, OT allows some novel approaches to the interaction of phonetics and phonology and to the issue of where the boundary between the two lies. This leads to the idea that instead of referring to traditional phonological elements, constraints refer directly to phonetic cues. As Gerfen points out, a problem with assessing these alternatives is that they often have the same empirical coverage. He examines in detail a particular case in Eastern Andalusian Spanish that he argues allows us to distinguish the two possibilities, and he shows that it can be best explained by licensing constraints that refer to syllable position rather than to phonetic cues. Gerfen's chapter is of particular interest in that its arguments are supported by both phonological patterning and phonetic data: measurements of duration support the proposed syllabification analysis, and thus, the argument that the context of neutralization is best described in terms of syllabic position.

Yip's chapter, "Segmental Markedness versus Input Preservation in Reduplication," looks at the segmental alternations involved in several types of reduplication in Chinese languages. Her chapter shows that these alternations can be seen as a direct result of contextual featural markedness; the segments that appear are the least marked for the various syllable positions involved. On the basis of this, she argues that the only faithfulness relationship that should be recognized is that between input and output (IO); there is no Base-Reduplicant faithfulness relation specific to reduplication. Thus, IO faithfulness is fully satisfied by a faithful realization in just one copy of the pair, leaving the other copy free to alternate under pressure of markedness.

The third and final part of this volume looks at broader questions of "The Structure of the Grammar," concentrating on how we should account for opaque phonological interactions.

Fukazawa's "Local Conjunction and Extending Sympathy Theory" shows that McCarthy's (1998) proposal of Sympathy, a type of candidate-to-candidate correspondence, can be shown to account for phenomena beyond the residual derivational opacity problems that it was designed to solve. This is an important point, showing that Sympathy has broader OT-internal application and is not just an appendage to explain away troublesome cases. The argument is based on a complicated OCP-triggered alternation in Yucatec Maya that is restricted to homorganic consonants. The analysis also contributes to the question of the constraints that trigger alternations, the topic of the second part, as it is shown that the combined requirements for violating this OCP constraint are accounted for by Smolensky's (1993) Local Conjunction of constraints.

Ito and Mester's chapter, "Structure Preservation and Stratal Opacity in German," addresses the details of an opaque interaction between two allophonic rules in German and shows that this type of interaction, where both rules are allophonic and affect each other's environment, cannot be accounted for using Sympathy Theory. From this interaction they draw the important and likely controversial conclusion that we must recognize some serialism within the phonology in OT. The two alternations, although both nondistinctive, otherwise differ in precisely the properties that are used to distinguish lexical and postlexical rules in Lexical Phonology. Ito and Mester show that if the two are the result of two different constraint rankings, one feeding into the other, the facts can be accounted for straightforwardly. Their model formalizes many of the key insights of Lexical Phonology, but within an OT grammar.

The chapters in this volume show both that OT allows new insights into segmental phonology, and that the details of segmental phonology provide insight into the broadest possible issues in the structure of UG. From basic questions of the representation of a segment to whether the fundamental architecture of the grammar is serial or fully parallel, these are important contributions to our formal understanding of the language faculty.

Notes

For their hard work on this volume, the editor would like to thank all of the contributors as well as our erstwhile representative at Cambridge University Press, Christine Bartels. Thanks also to Jill Beckman and Donca Steriade for their consideration early on in the project, although they were ultimately unable to contribute for reasons of scheduling. I would especially like to thank

John McCarthy for encouraging me to persist at some discouraging junctures. The usual caveats apply about none of the aforementioned being responsible for my or each other's errors and misconceptions. Finally, thanks to the technical support department at Three Cats Software for crucial assistance in coping with the many ways in which computers have made our lives so much easier.

1. Some earlier attempts at crosslinguistic typologies include Lombardi (1991, 1995) and Cho (1991). Of course, the intent of work in feature geometry and underspecification was also originally to achieve crosslinguistic restrictions on featural phonology by arriving at the correct structure for the segment. However, given the proliferation of different proposals for the representation of the same distinctions in this literature, it was unclear that this framework was approaching an agreed-upon and appropriately restrictive account.

References

Cho, Young-Mee Yu. 1991. Parameters of Consonantal Assimilation. PhD dissertation, Stanford University.

Ito, Junko. 1986. Syllable Theory in Prosodic Phonology. PhD dissertation, University of Massachusetts at Amherst.

Ito, Junko. 1989. A Prosodic Theory of Epenthesis. *Natural Language and Linguistic Theory* 7:217–259.

Flemming, Edward. 1995. Auditory Representations in Phonology. PhD dissertation, UCLA.

Kisseberth, Charles. 1970. On the Functional Unity of Phonological Rules. *Linguistic Inquiry* 1:291–306.

Lombardi, Linda. 1991. Laryngeal Features and Laryngeal Neutralization. PhD dissertation, University of Massachusetts, Amherst. Published by Garland, 1994.

Lombardi, Linda. 1995. Laryngeal Neutralization and Syllable Wellformedness. *Natural Language and Linguistic Theory* 13:39–74.

McCarthy, John J. 1981. A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry* 10:433–466.

McCarthy, John J. 1986. OCP Effects: Gemination and Antigemination. *Linguistic Inquiry* 17:207–263.

McCarthy, John J. 1999. Sympathy and Phonological Opacity. *Phonology* 16:331–399.

McCarthy, John J., and Alan Prince. 1993. Prosodic Morphology I: Constraint Interaction and Satisfaction. Technical Report No. 3, Rutgers University Center for Cognitive Science.

McCarthy, John, J., and Alan Prince. 1995. Faithfulness and Reduplicative Identity. University of Massachusetts Occasional Papers 18, pp. 249–384. Amherst, MA: GLSA. ROA 60-0000, http://ruccs.rutgers.edu/roa.html.

Mester, R. Armin. 1986. Studies in Tier Structure. PhD dissertation, University of Massachusetts, Amherst.

Odden, David. 1986. On the Obligatory Contour Principle. *Language* 62:353–383.

Odden, David. 1988. AntiAntigemination and the OCP. *Linguistic Inquiry* 19:451–475.

Pater, Joe. 1999. Austronesian Nasal Substitution and other *NC̥ Effects. In *The Prosody-Morphology Interface*, ed. R. Kager, H. van der Hulst, and W. Zonneveld, pp. 310–343. Cambridge: Cambridge University Press.

Prince, Alan, and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. To appear, MIT Press.

Smolensky, Paul. 1993. Harmony, Markedness, and Phonological Activity. Paper presented at Rutgers Optimality Workshop 1 and revised handout.

Yip, Moira. 1988. The Obligatory Contour Principle and Phonological Rules: A Loss of Identity. *Linguistic Inquiry* 19:65–100.

PART I

THE CONTENT OF REPRESENTATIONS