

Review article

A review of two books introducing Optimality Theory

Archangeli, D.B. and Langendoen, T.D., editors, 1997: *Optimality Theory: an overview*. Oxford: Blackwell. 233 pp. ISBN: 0 631 20225 0 (£50 hardback).

Kager, R. 1999: *Optimality Theory*. Cambridge: Cambridge University Press. 452 pp. ISBN: 0 521 58019 6 (£42.50 hardback).

I Introduction

In the introductory chapter of the first of these two books, Diana Archangeli proclaims that Optimality Theory is 'THE Linguistics Theory of the 1990s'. Whether or not one readily agrees with this statement, there is no denying that Optimality Theory (OT) was one of the fastest-spreading linguistic theories in the 1990s. The rapidity with which the framework has developed and gained popularity since its inception in the early 90s can be easily attested in the ever growing amount of related work produced in the forms of working papers, conference papers, dissertations, journal articles and edited books. The publication of these two introductions indicates that the accumulation of such research effort now warrants textbook treatments. In this review, I will first give a brief outline of OT. The reviews of the individual volumes come next. The last section discusses the implications of OT for second language acquisition research.

OT is a version of generative grammar that defines well-formedness in terms of constraint interaction. For a given input to the grammar, the output (i.e., the surface realization of the input) is the form that best satisfies a set of constraints.¹ These constraints are held to be universal, and because they are of general nature, they are inherently conflicting. The conflicts are resolved through

¹ 'Input' here refers to any structure that enters the computation of a grammar, e.g. a lexical entry; not to be confused with the ambient data that a learner is exposed to.

the ranking of the constraints, which is determined on a language-specific basis. For any pair of constraints in conflict, the higher-ranked constraint takes absolute precedence, such that violations of a lower constraint are minimally permitted to the extent that they minimize violations of the higher constraint. The locus of language variation in OT is therefore the constraint ranking, while all other components are held constant across languages. These components, which make up Universal Grammar, are: the universal set of constraints, the functions GEN (for GENERator) and EVAL (EVALuator). GEN's role is to provide a set of all possible analyses for each input. The candidate analyses are then submitted to EVAL which selects the optimal output candidate, taking into consideration the language-specific ranking of constraints.

It is important to note that the inputs to the grammar and the output candidates generated by GEN are unrestricted (except by the requirement that they must be linguistic structures). In other words, the grammar has no direct control over the lexicon or the operations on the input structure, but only over the evaluation of wellformedness. It is equally important to keep in mind that wellformedness in OT is not only determined on the basis of the structure of output forms. This is because constraints come in two types. The first type, markedness, imposes structural restrictions on the output, while the second type, faithfulness, demands that the input and output be identical in various aspects. In effect, markedness constraints militate against marked structures in the output, but faithfulness constraints protect underlying contrasts between unmarked and marked values. Such tension between the demands of markedness and faithfulness is the essence of an OT grammar.

II Archangeli and Langendoen, 1997

The volume edited by Archangeli and Langendoen is the first in a series called 'Explaining Linguistics'. The book is intended to introduce OT to anyone 'who desires to understand the model, regardless of their background in formal linguistic theory itself'. The editors and contributors to the volume courageously tackle this Herculean task by starting off the book with an introduction to linguistics, ensuring that the basic concepts are defined on their first appearance, and glossing over much of the technicality in the formal analyses. Nevertheless, explaining OT on top of introducing phonology, morphology and syntax themselves in a book that is fewer than 250 pages long seems too ambitious a project, and the effort breaks down at several points in the book. The problems are

particularly apparent in the syntax chapters, which assume some familiarity with Principles and Parameters Theory (PPT). For instance, the first syntax chapter delves into a discussion of verb-second phenomenon couched in X-bar theoretic terms, but there is no preceding explanation of X-bar theory itself. In the second syntax chapter, the point is made that all PPT principles contain some hedge, which is suggestive of the effects of violable constraints. It would be surprising, though, if anyone without previous training in linguistics could grasp, for example, the Empty Category Principle from the simple statement 'A trace must be properly governed'. I am therefore sceptical that this textbook can serve the purpose it is intended for.

The book comprises an Introduction and 5 chapters on different linguistic domains, each written by an expert in the area. In addition, there is an Afterword with substantive discussion on the nature of input in an OT grammar. What is quite admirable about the make up of the chapters is the balance. There are two chapters on phonology (the first by Michael Hammond and the second by Douglas Pulleyblank), one on morphology (Kevin Russell) and two chapters on syntax (one each by David Pesetsky and Margaret Speas). Chapter 1 (Diana Archangeli), which introduces general issues, and the Afterword (no authors mentioned) also address all three of those areas. This may not sound particularly worthy of mention from the viewpoint of general introductions to linguistics, but in light of the fact that there is a high concentration of OT literature on phonology, the editors are to be commended on their effort to strike a balance, at least in quantity, among different areas. Of course, one cannot help notice the difference in approach to phonology and its related interfaces versus other areas. The phonology chapters and much of the morphology chapter focus on describing the more or less standard line of OT analysis in the respective domains, whereas the syntax chapters are highly speculative, guided by the question of whether OT is relevant to syntax at all. But this is quite understandable, considering the underdeveloped status of OT syntax at the time this book was written.

In many ways, OT provides an alternative perspective to the already established theories in linguistics, and this is likely to be a point of interest to all potential readers of introductions to OT. The contrast with previous approaches is most apparent in phonology. In fact it can be said that the direct impetus for OT was to solve perennial problems that plagued the rule-based approach adopted in most generative phonology analyses since Chomsky and Halle (1968). Rewrite rules were known to be too powerful, inadequate

in capturing the functional unity among seemingly unrelated processes (a problem known as 'conspiracy'), and redundant in that they mirror the restrictions imposed on the lexicon (the so-called 'duplication' problem). The use of constraints to remedy this situation predates OT, but it was not until the ideas of violability and constraint interaction were introduced by OT that a fully universal set of constraints could be shown to dispose of any language specific rules. Because constraints must be grounded in markedness theory, the combined effects are restricted and do not present the type of arbitrariness that is inherent of rules. The output orientation of OT overcomes the problems of conspiracy and duplication, since under OT's view, achievement of a particular surface configuration is what drives phonological processes, and restrictions on output forms need not be encoded in the lexicon.

These issues are discussed in the introductory and the phonology chapters albeit in non-technical terms. Hammond's chapter on prosody shows how an OT account of syllable typology is more parsimonious than its rule-based counterpart in that it correctly rules out impossible languages. Furthermore, a constraint-based analysis of English stress pattern is free from the incoherence exhibited by a rule-based account with its package of rules that seem to have nothing to do with each other. In his chapter on segmental phonology, Pulleyblank demonstrates that OT obtains segmental inventories from constraint interaction without having to specify them in the lexicon, thus overcoming the duplication problem.

It is unfortunate that the morphology chapter by Russell gives only a cursory treatment to the issue of 'cyclic effects'. This is one of the most interesting points of contention between OT and a derivational approach such as Lexical Phonology, which invokes multi-level analysis of the phenomenon. Most work in OT has abandoned the notion of cyclicity and captures the effects of morphological uniformity in terms of a phonological identity relation between morphological related forms. Russell misses an opportunity to showcase how OT differs radically from previous approaches to the phonology-morphology interface. Those who are interested in this issue might want to consult Chapter 6 of Kager's book, reviewed below.

The syntax chapters fully explore the comparison between OT and the approach of PPT and the Minimalist Program (MP), since that is precisely the task the writers have set up for themselves. Both Pesetsky and Speas see the crucial difference to be the (minimal) violability of OT constraint vs. the inviolability of

principles in PPA/MP, but the two take different approaches to examine this contrast. Pesestky's litmus test of (in)violability is ineffability: the state in which no grammatically acceptable output structure is obtainable for a given input. Ineffability calls for a model of syntax that completely blocks structures that violate certain principles, such as PPT/MP (or the 'clash and crash' approach as Pesestky calls it). On the other hand, OT will still select as a grammatical output form some structure that least violates the demands of the principles. Using this criterion, Pesestky argues that OT does not apply to core syntactic phenomena such as phrase structure and movement. Speas takes a different perspective, and focuses on the hedge contained in almost all syntactic principles. For example, the Case Filter states that an NP must have case *unless it is null*. The proviso that accompanies each principle seems to come from requirements of other syntactic principles, thus giving a flavor of interaction among violable constraints. Adopting the OT concept of minimally violable constraints can then free the principles from the hedge. Speas' conclusion, therefore, is much more optimistic than Pesestky's in that she finds a central role for interaction of violable principles in syntax.

The Afterword highlights another aspect of OT that separates it from other approaches. Due to the output orientation of the theory, there is more room for postulating different input structures. In phonology, this means that inputs can be underspecified for redundant features, fully specified, or lack distinct representations (because more than one input can lead to the same output). The third of these conceptions is unique to OT, and the same freedom of input representation can be implemented in morphological and syntactic structures. In syntax, there is an added question of whether or not a distinct input should be associated with a single semantic representation. Given the unresolved status of these issues, the discussions in the Afterwords are speculative but sure to whet the appetite of the more curious readers. It is regrettable that this last section of the book conspicuously lacks references to the relevant literature.

These discussions are only marginally extended to issues of language acquisition. In Chapter 1, Archangeli points out that characteristics of a developing grammar can be seen as effects of a nontargetlike ranking of constraints. OT constraints are universal, hence ensuring the continuity between child grammar and the target adult grammar, while in a rule-based account of child phonology, there is nothing that prevents the child from postulating a rule that lies outside the possibilities of natural language. In his chapter on morphology, Russell further introduces the well-

accepted hypothesis in the OT acquisition literature that the initial state is a hierarchy with the markedness constraints ranked above the faithfulness constraints. Why that has been proposed, however, is not explained (we will return to this point below). Second language acquisition is mentioned very briefly by Archangeli as one of the areas that can be explored using OT, although the point made – that different constraint rankings of the L1 and L2 may explain phonological transfer – is an obvious one. In general, language acquisition does not feature prominently in this book, leaving it up to reader's imaginations to determine exactly how OT can address issues in language learning.

This book shows the familiar strengths and weaknesses of multi-authored textbooks. On one hand, it benefits largely from the expertise each author brings into her or his area. Each chapter is a tour de force in and by itself. On the other hand, there are some organizational problems that would not have arisen in a single-authored book. To name one, although it is suggested that readers follow the sequence of the chapters, we find explanations of the goals of syntactic theory and the basic structure of an OT model of syntax in the second syntax chapter, not the first. Another problem, which could have been eliminated through editing, is the use of unconventional or highly confusing constraint names. For example, in most OT works, the constraint that requires all syllables to be parsed into feet is called PARSE-SYLLABLE. The same constraint is given the name *FOOTLESS in this book, and PARSE-SYLLABLE means something else. While the constraint NAS/VOI means 'A nasal must be voiced' (p. 74), NAS/CONT is defined as 'A nasal must *not* be continuant' (p. 76, emphasis mine). These abbreviations disguise the dissimilarity of their referents, and the reader should not be left to guess whether a constraint such as CONTRASTIVECODA means 'A coda must bear a contrastive feature' or 'A coda must not bear contrastive features' (it turns out to be the latter!). Not surprisingly, this haphazard naming of constraints has left some substantive errors unchecked, so we find LOWROUND defined as '[+low] vowels are [+round]' (p. 201) where it was meant to be: '[+low] vowels are [-round].

In sum, despite its intention, Archangeli and Langendoen (1997) may not be suitable for complete beginners in linguistics, but it does offer an well-rounded entry point into the theory for anyone with some basic training in linguistics, including researchers who wish to take a crash course in the theory. Interested readers, however, are strongly recommended to consult the primary sources since much of what is presented in the book has undergone substantial simplification.

III Kager, 1999

While Kager's book is also published as a textbook, its target audience is different from that of Archangeli and Langendoen. Kager assumes a readership with 'a basic knowledge of derivational Generative Phonology... and some knowledge of Minimalist Syntax'. This, coupled with the length of book (450 pages plus), has allowed him to go much more into depth without having to rush through the material as is the case in the volume by Archangeli and Langendoen. Both the content and the level of exposition seem to be pitched appropriately at the intended audience. The exercises that accompany the main chapters make the book suitable for use in a course (in phonology, for the reason given below) and for self-study.

Like Archangeli and Langendoen, Kager explicitly takes the position that OT is not only a theory of phonology, but also a general theory of linguistics. Yet, the coverage of the book is considerably skewed toward phonology. The first two chapters introduce the basic principles and architecture of the theory, but the perspective and empirical material are both phonological. Chapters 3 through 6 are assigned to various domains in phonology and its interface with morphology. Chapter 7, which is devoted to the learnability of OT grammars, focuses on the acquisition of phonology. Chapter 8 is the only chapter on syntax. Finally, Chapter 9 discusses residual issues, whose empirical implications are again mostly in phonology. Essentially, therefore, this book is an introduction to OT phonology, with bonus chapters on learnability and syntax. Within phonology, Kager emphasizes prosodic phenomena, so the four main chapters on phonology cover syllable structure, metrical structure, reduplication, and truncation/affixation, respectively. In contrast with Archangeli and Langendoen, therefore, the empirical data discussed in Kager's are narrower in scope.

The focus on phonology has several benefits. First, the motivations behind the basic tenets of OT are clearly contextualized within the development of generative phonology. In the introductory chapters, Kager provides a lucid account of the connection between OT's characteristics – abandonment of rules, output orientation and freedom of analysis – with classic problems in phonological theory. In particular, he devotes Chapter 2 to explaining how OT resolves the problem of conspiracy. The demonstration based on Pater (in press) shows that a range of attested phonological processes involving postnasal obstruents are unified under the interaction between a single markedness

constraint and faithfulness constraints.

Secondly, there is ample space to give a detailed OT analysis of complex data, and contrast it with an alternative rule-based account. For example, in Chapter 4, the intricate pattern of rhythmic lengthening in Hixkaryana is shown to arise from interactions of constraints on four dimensions. A comparison with Hayes' (1995) analysis is also presented to illustrate the advantages of the OT approach, which does away with a number of ad hoc intermediate representations the rule-based account relies on. This comparative formula of presentation is employed throughout the phonology chapters.

Thirdly, the book includes discussion of advanced and contentious issues in OT phonology, most notably the extension of Output-to-output Correspondence to the phonology-morphology interface (Chapter 6) and the issue of opacity (Chapter 9). The idea that transderivational transfer can be explained in terms of *Output-to-output Correspondence* – identity effects between surface forms of morphologically related words – is one of the boldest challenges OT has mounted against derivational phonology since 'cyclic' effects have long been viewed as a hallmark of serial derivation. On the other hand, *opacity* – the type of phonological phenomena that cannot be deduced from surface true generalizations – is the Achilles heel of OT. Opaque generalizations are easy to explain using rule ordering in a serial model, but they pose a serious problem for output-oriented OT. These two issues attract central attention in current OT research, and anyone interested in the latest developments of the theory would greatly benefit from reading Kager's comprehensive summaries in these chapters.

The main phonology chapters (Chapters 3 to 6) also profit from careful organization. The empirical domains treated in these chapters are ordered in a way so that each chapter only presupposes understanding of the phonological phenomena described in the preceding chapters. Chapter 3 deals with basic syllable structure. Chapter 4 introduces moraic theory and the foot inventory before embarking on a discussion of metrical structure. Chapter 5, which examines reduplication, builds on the representational assumptions of the prosodic hierarchy established in Chapters 3 and 4. Finally, Chapter 6 draws on various segmental and prosodic observations in investigating the phonology-morphology interface as attested in truncation and affixation. Kager successfully introduces formal analytical tools of OT in a piecemeal fashion at relevant points of empirical discussion rather than putting forth abstract models in a vacuum of data. Thus by the time the formal definitions of Generalized Alignment and Correspondence Theory are presented

in their complete form (in Chapter 3 and Chapter 5, respectively), readers will have gained familiarity with the workings of these concepts in actual data analysis.²

Unfortunately, the amount of attention Kager invests in phonology is not matched in the chapters on learnability and syntax. Not only are these chapters much shorter than others, but they also lack the kind of synthesis the phonological chapters offer. This is largely due to Kager's decision to anchor each of these chapters around a single piece of work. Thus, Chapter 7 discusses Tesar and Smolensky (1993) almost exclusively, and Chapter 8 is basically a summary of Grimshaw (1997). True, Kager uses this strategy in the phonology chapters as well. For instance, Chapter 2 is based on Pater (in press) and Chapter 5 on McCarthy and Prince (1995). But with practically seven chapters allocated to phonology, there is still plenty of space left to provide an overview of the area, whereas devoting the only chapters on learnability and syntax to single studies results in exclusion of issues not dealt within those studies. Among the important topics which are left out is the learnability argument for an initial-state ranking with markedness constraints overranking faithfulness constraints (Smolensky, 1996), which is now adopted by most OT-based research in phonological development (e.g., Demuth, 1995; Gnanadesikan, 1996; Pater, 1997). The crux of the argument is that if the initial state contains the ranking $F \gg M$, where F stands for a faithfulness constraint and M for a markedness constraint that potentially conflicts with F , a child exposed to a target ranking $M \gg F$ would never be able to learn it without negative evidence because the child's grammar simply generates a superset of the forms attested in the ambient data. To avoid this no-way-out situation, the learner must start out with the ranking $M \gg F$. This type of learnability issue has particularly important implications for second language acquisition in which the learner's initial state (presumably the L1 ranking) and the target ranking could stand in such a relation that cannot be 'relearned.'

The syntax chapter, because it focuses only on Grimshaw's work, provides a somewhat limited overview of the significant debate on the nature of input in OT syntax. Grimshaw's position is that candidates for a given input must be semantically nondistinct. This contrasts with the standard Minimalist assumption that the input is simply a numeration of lexical items which can generate semantically different outputs, and the similar position taken in the

² Generalized Alignment is a format of constraints that require matching of phonological and/or grammatical constituents. Correspondence Theory is a general theory of constraints that require identity between pairs of representations (including input and output, base and reduplicant, base and truncated form, and stem and affixed form).

OT syntax work of Legendre, Wilson, Smolensky, Homer and Raymond (1998) and Keer and Bokavic (1997). Although these references are mentioned in a footnote and in the further reading section, a more balanced treatment of this unsettled issue could have been achieved by giving a proportionate discussion of the alternative views. Another point of debate concerns the amount of structure that must be included in the input. In this respect, it would be informative to compare Grimshaw's assumptions with those of Speas, because Speas disagrees with Grimshaw on both accounts. The Afterword by Archangeli and Langendoen also presents a summary of the contrasting views on input in OT syntax.

In spite of these limitations, Kager's focus on Tesar and Smolensky (1993) and Grimshaw (1997) in these chapters seems appropriate as these are the studies that have set out the agenda for the OT research in learnability and syntax. Tesar and Smolensky (1993) is by far the most thoroughly worked out model of OT learning mechanisms.³ Because OT assumes that all constraints are universal and provided by Universal Grammar, the task of the learner is seen as one of arranging the initial state constraint ranking to match the target ranking. Tesar and Smolensky's model provides an algorithm that allows the learner to achieve this using only pairs of input and output. The learning procedure works on important concepts which underlie much subsequent learnability research in OT, such as mark cancellation and error-driven constraint demotion. Grimshaw's contribution, similarly, has been to construct a model of syntax within OT. While many aspects of OT syntax are still in the process of being developed, Grimshaw's work remains a key study.

To sum up, I would recommend this volume mostly to readers whose main interest is phonology. The material presented in the phonology chapters is authentic and accurately represents the state of the art in OT phonology. Researchers with a general interest in OT should also find the book valuable as a summary of the basic concepts and outstanding issues, and as a pointer to pertinent primary sources, as each chapter is followed by a thematically organized list of references.

IV OT and SLA

Perhaps reflecting the fact that OT's impact has been felt more strongly in phonological theory than in other areas of linguistics, or perhaps because there was a pressing need in phonology for a

³ A revised version of this work is published as Tesar and Smolensky (1998).

theory that enables the construction of an acquisition model with developmental continuity, research in child language acquisition using OT has so far been carried out mostly in phonology. This asymmetry is also reflected in second language acquisition research, although in this case, the total amount of output in OT-based research itself is still minuscule. Nevertheless the outcome of the few extant second language studies is quite promising. Hancin-Bhatt and Bhatt (1997) have demonstrated that a classic problem in second language acquisition (SLA) of the interaction between transfer and developmental effects receives an explicit account within OT. The advantage of OT lies in its capacity to formalize within a single model of grammatical development not only the effects of language specific properties of the native and target languages, but also the role of universal principles shared among all grammars. As shown by Broselow, Chen and Wang (1998), the effects of these universal properties are not necessarily apparent in all languages so that an unexpected phonological pattern can arise in the learner's grammar due to markedness constraints whose effects are masked in the native grammar. These preliminary investigations suggest that within SLA research there is great potential of OT. But that potential has been little exploited. It is hope that these introductory texts will encourage more researchers to explore the possibilities of OT as a tool to further our understanding of SLA.

In closing, I should mention the usefulness of Rutgers Optimality Archive (ROA) as a source of further reading. A large number of studies in OT, including many of the references cited in the two volumes reviewed here, can be downloaded from this electronic archive. ROA is accessible at: <http://rucss.rutgers.edu/roa.html>

Acknowledgements

Thanks are due to Margaret Thomas and Antonella Sorace for helpful suggestions and comments on earlier drafts of this review.

V References

- Broselow, E., Chen, S.-I. and Wang, C. 1998: The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* 20, 261-80.
- Demuth, K. 1995: Markedness and the development of phonological structure. In Beckman, J., editor, *Proceedings of the North East Linguistic Society* 25.
- Gnanadesikan, A. 1996: Child phonology in Optimality Theory: ranking

- markedness and faithfulness constraints. In Stringfellow, A., Cahana-Amitay, A. D., Hughes, E. and Zukowski, A., editors, *Proceedings of the 20th Annual Boston University Conference on Language Development*. Sommerville, MA: Cascadilla Press.
- Grimshaw, J.** 1997: Projection, heads, and optimality. *Linguistic Inquiry* 28, 373–422.
- Hancin-Bhatt, B. and Bhatt, R.M.** 1997: Optimal L2 syllables: interactions of transfer and developmental effects. *Studies in Second Language Acquisition* 19, 331–78.
- Hayes, B.** 1995: *Metrical stress theory: principles and case studies*. Chicago, IL: University of Chicago Press.
- Keer, E. and Bokavic, E.** 1997: Have FAITH in syntax. *Proceedings of the West Coast Conference on Formal Linguistics* 16.
- Legendre, G., Wilson, C., Smolensky, P., Homer, K. and Raymond, W.** 1998: Optimality and wh-extraction. In Barbosa, P., Danny, F., Hagstrom, P., McGinnis, M., and Pesetsky, D., editors, *Is the best good enough?* Cambridge, MA: MIT Press.
- McCarthy, J. and Prince, A.** 1995: Faithfulness and reduplicative identity. In Beckman, J., Walsh Dickey, L., and Urbanczyk, S., editors, *University of Massachusetts Occasional Papers in Linguistics* 18. Amherst: Graduate Linguistic Student Association.
- Pater, J.** 1997: Minimal violation and phonological development. *Language Acquisition* 6, 201–53.
- in press: Austronesian nasal substitution and other NC effects. In Kager, R., Van der Hulst, H. and Zonneveld, W., editors, *The prosody-morphology interface*. Cambridge: Cambridge University Press.
- Smolensky, P.** 1996: The initial state and ‘richness of the base’ in Optimality Theory. Johns Hopkins University, Baltimore, MD.
- Tesar, B. and Smolensky, P.** 1993: The learnability of Optimality Theory: an algorithm and some basic complexity results. Technical Report CU-CS-678-93, Department of Computer Science, University of Colorado, Boulder, CO.
- Tesar, B. and Smolensky, P.** 1998: Learnability in Optimality Theory. *Linguistic Inquiry* 29, 229–68.

Mitsuhiko Ota
Department of Theoretical and Applied Linguistics
University of Edinburgh