J. Linguistics **37** (2001), 127–143. Printed in the United Kingdom

REVIEW ARTICLE

The best of all possible words

T. MARK ELLISON & EWAN KLEIN

University of Edinburgh

(Received 16 December 1999; revised 17 June 2000)

Diana Archangeli & D. Terence Langendoen (eds.), *Optimality Theory: an overview***.** Malden & Oxford: Blackwell, 1997. Pp. xii + 233.

1. Introduction

According to the front page, Archangeli & Langendoen's *Optimality Theory: an overview* (henceforth A&L) is 'the first in a series of volumes of essays which are designed to introduce and explain major research areas in linguistic theory and practice'. On the back cover, we learn that it provides 'the first general introduction to optimality theory – arguably *the* linguistic theory of the 1990s'. And the Foreword states that the intended audience is 'anyone with a serious interest in language who desires to understand [Optimality Theory], regardless of their background in formal linguistic theory itself'.

In many respects, the book does indeed fulfill its announced goals. As far as we are aware, Archangeli & Langendoen were first to market with an introduction to OT, though since then, Kager (1999) has also appeared. The essays themselves are generally lucid and well-edited, and together they carry out a pretty comprehensive, though necessarily superficial, tour of the main landmarks in OT. Various aspects of phonology are covered in three chapters by Archangeli, Hammond and Pulleyblank; there is a chapter on morphology by Russell; and two chapters on syntax by Pesetsky and Speas. The six main essays are sandwiched between a useful Foreword and a more contentious and opaque Afterword.

Despite the deserved praise, we have some doubts about the coherence of the pedagogical strategy adopted by the volume. Clearly, OT is not yet sufficiently mature to provide the theoretical setting for a general introduction to phonology, on the model of Kenstowicz & Kisseberth (1979); this can be taken as a positive symptom of the vitality of OT as a research programme. Nevertheless, A&L does have pretensions to being an introduction not just to OT but to phonology and syntax more generally, thereby honouring its commitment to all readers 'regardless of their background in formal linguistic theory itself'. Thus, each of the six main

essays in A&L opens with an 'introduction to X', where X ranges from linguistics in toto through distinctive features in phonology to constituent structure and constraints on syntactic movement.

Intriguing though it is to imagine a generation of linguists reared on an initial diet of OT, this volume couldn't realistically be used as a first linguistics text for undergraduates, on at least two counts. Although most of the introductory material in A&L is pretty good as far as it goes, it just isn't comprehensive enough to provide a reasonable foundation for doing linguistics. Second, there is almost a head-on clash between the perspectives adopted in the two syntax chapters, a conflict that is guaranteed to dismay and confuse most undergraduate readers. In summary, then, the ambitions of A&L to serve as an introduction to phonology, morphology and syntax, as well as OT, seem misguided.

Perhaps more seriously, A&L also has significant shortcomings as an introduction to OT. There is no systematic guidance to further reading in the area. There are no exercises. And there is no discussion of methodology, of how to set about constructing an OT analysis of a particular phenomenon. Thus, even if you have managed to select an appropriate group of putatively universal constraints, how do you determine their ranking and what rules of thumb might you apply to check that you have tested them against a sufficient set of candidate forms?

2. The papers

2.1 *Archangeli*

The first chapter, entitled 'Optimality Theory: an introduction to linguistics in the 1990s' and written by Archangeli, sets the scene for the rest of the book, introducing in turn (generative) linguistics, OT, and the formative example of syllable structure. Mirroring the book as a whole, it has a closing section which introduces many more topics.

The most attractive attribute of this chapter is its clarity. For example, linguistics is defined as the study of LANGUAGE UNIVERSALS and LANGUAGE CHANGE. Language universals are attributed to rigidity in the innate language capacity, whose totality is named UNIVERSAL GRAMMAR. Linguists look for four kinds of language phenomena: (i) patterns within languages, (ii) variation between the patterns of different languages, (iii) universals resulting from our innate endowment, and (iv) markedness, a feature indicating the robustness of a particular property in a given language. These four points summarise with conciseness and clarity the ambitions of generative linguistics.

The characterisation is, appropriately, slanted toward the topic of the book, but deftly so. OT is defined by its account of the four language phenomena itemized above: universals are constraints. Language-specific

patterns, the corresponding cross-linguistic variation and markedness all flow from variation in the priority ranking of the universal constraints.

These general phenomena, and their realisation in OT, are exemplified with syllable structure. Differences of syllabification in Yawelmani (epenthesis), Spanish (elision), English (complex codas) and Berber (consonantal peaks) illustrate the effects of different rankings of four syllabification constraints. These constraints are FAITHV (faithful mapping of vowels from underlying to surface forms), FAITHC (faithful mapping of consonants from underlying to surface forms), PEAK (requiring syllables to have a vowel), and *COMPLEX (syllables only permitting single consonants at either edge).

As an introductory chapter, this presentation is excellent. There are boxed slogans to remember key points (a device which is also used to good effect in subsequent chapters). The examples are straightforward and pertinent. Clarity, however, is bought at a price. This chapter, like most of the book, is about polemic exemplification, rather than argumentation. A critical reader, not aware of the rest of the literature, might be left with many questions. In each of the examples, only one constraint is violated by the optimal candidates; for example, English allows violations of *COMPLEX. Furthermore, only one candidate is ever proffered which violates only this constraint. The imaginative reader can ask themselves whether it might not be just as well to drop one of the constraints in each of the languages. So why does OT say that English has the *COMPLEX constraint if this constraint does not select between candidates? This issue is not addressed.

After the examples, a conclusion section summarises the contents of the chapter. The chapter does not end with the summary, however. Instead this is followed, rather curiously, by three 'addenda'. The first gives a short history of OT, and what it has offered generative phonology and syntax. This is useful more to someone with a knowledge of phonology from an earlier period. The complete novice is less likely to be interested in the advantages of OT over SPE generative phonology.

The second addendum presents outstanding questions in OT itself. Questions on all aspects of the theory are considered: the input, GEN, CON, EVAL, the output and cognitive modularity. These discussions are certainly helpful. However, they seem restricted, given the amount of work going on in OT. There are a total of ten linguistic references cited. Excluding papers by authors in this book, and papers by authors responsible for the foundations of OT (Prince, Smolensky, McCarthy), this amounts to only three remaining papers to help acquaint the reader with the wider world of OT research.

The third addendum looks at extensions to the theory. These relate OT to: poetics, borrowings, second language acquisition, first language acquisition (as both an empirical and logical problem), language change, language perception, language production, and the computational modelling of language. Once again, this very interesting overview is marred by a lack of

references. There is a footnote (fn. 8, p.30) referring the reader to the Rutgers Optimality Archive (ROA) (also highlighted on page xii of the book's Foreword).[1] In the sections themselves, only two of the topics covered also mention who is pursuing the research. None have references to papers. This seems less than useful. If these papers are easily available on the ROA, why not find them and make reference directly to individual papers? Not to do so seems negligent.

In summary, this chapter presents with great clarity a view of the nature of generative linguistics. The presentation of OT is also remarkably clear. If you wish to be convinced about OT's claims, this chapter will not do it. But it will let you know what OT is and wants to do, if you don't already know. The addenda introduce the reader to some of the peripheral topics, but do not offer useful pointers to further reading.

2.2 *Hammond*

The second chapter, 'Optimality Theory and prosody', continues the axiomatic style of the previous chapter while narrowing down the exemplification to English prosody. This data is used to argue in favour of one OT approach over a rule-based account.

The chapter begins with the generative, mentalist claim that 'the unacceptability of *tkin* as a word of English is a consequence of unconscious knowledge about English' (34). The subject matter of linguistics is recharacterised as this unconscious knowledge and its acquisition. The claim is made then that syllables are not matters of physiology, but of linguistic knowledge, and are thus ripe for linguistic study.

Hammond proposes that OT can derive only the permissible syllable structures. According to this argument, some patterns of syllable structure – such as allowing codas but forbidding onsets – do not occur across the world's languages, and thus should not be included in the repertoire of possible syllable types. Hammond shows how OT constraints on onsets and codas can interact to preclude just these patterns. However, no reference is made to work arguing that some languages have these forbidden structures (e.g. Kunjun, Scottish Gaelic and Irish).

The chapter seeks to show the superiority of OT by comparing it with an account using rewrite rules. This is useful for the reader with a background in linguistics, less so for the complete novice who has no particular reason to think that phonology should be about rewrite rules. In order for rules to account for the typology of possible syllable structures. Hammond claims they must be transformed into constraints. As with earlier examples, this argument is indicative rather than conclusive. Hammond's proffered rule

[1] http://ruccs.rutgers.edu/roa.html

system (38) can easily be modified to produce an attractive typology of syllable structures according to the independent presence or absence of three rules. Here is such a set of rules, in order of application. Following Hammond, unparsed material (not found between square brackets) is presumed not to surface.

(R1) **CV→[CV]**, parse onset-nuclei, UNIVERSAL
(R2) **V→[V]**, parse other unparsed nuclei, OPTIONAL
(R3) **V→[CV]**, epenthesis for unparsed nuclei, OPTIONAL
(R4) **]C→C]**, parse codas, OPTIONAL

Rule (R1) expresses the putative universal requirement for onsets. If rules (R2) and (R3) are both absent, then underlyingly onsetless syllables cannot be parsed. If rule (R2) is present, then underlyingly onsetless syllables are parsed as-is. If rule (R3) is present, then onsetless syllables are parsed with epenthetic onsets. Both rules being present has the same effect as rule (R2) only being present, a redundancy paralleling the redundancy in the ordering of NoCoda and Ons when these outrank Faithfulness. Finally, codas can surface only if rule (R4) is present. These rules account for the typology in syllable structure with as few degrees of freedom as corresponding OT constraints. Consequently, Hammond's claim that 'under a rule-based account, the only way to predict the typology is to change rules into constraints' (box on page 39) holds only so long as one chooses the wrong rules.

The chapter moves on to discuss higher levels of prosodic structure, namely feet, and phenomena associated with them: *fuckin*-infixation, syncope and stress.

The integration of syllable structure with stress assignment is achieved, as it must be in all such OT analyses, through a constraint with implications in both systems. The constraint NoOnset requires that 'stressless medial syllables are onsetless' (52). As a result, word-medial consonants preceding stressless syllables are 'captured' into the coda position of the immediately preceding syllable. NoOnset is ranked above NoCoda and Ons, as it must be if it is to have any effect. At the same time, those two constraints must be lower ranked than the stress-assigning constraints which make the syllable stressless. Otherwise all medial syllables would become stressed in order to keep their onsets.

This discussion of stress and syllabification consonant capture concludes with a review of the psycholinguistic work of Cutler et al. (1983) on the syllabification of intervocalic consonants. Subjects were instructed to recognise CV or CVC strings (e.g. [pa] or [pal]) in spoken input; this input consists of words such as *palace* or *palmier* which contain both target strings. In French (which does not have consonant capture), subjects recognize the targets faster if the target string ends at the edge of a syllable in the input; for example, [pa] is recognized faster than [pal] if the input is [pa.las] (*palace*). However, in English there is no such variation.

Hammond argues that a rule-based account could not capture this effect. If the consonant capture rule does not apply, the experiment should show a correlation with underlying syllable structure, while if it does apply, surface syllable structure should be seen. However, in an OT account, EVAL does not complete the comparison of candidates in English, because it temporarily lacks a stress value for the following vowel. Because of the different constraint ranking in French, the same blockage does not leave an ambiguity in the surface syllable structure.

Once again, the argument is illustrative rather than conclusive. Hammond achieves ambiguity in syllabification by having the uncertainty in the stress level of the unheard following vowel vitiate the evaluation process of lower-ranked constraints. Rules are not allowed the same device, and neither syllable structure available before or after application of rewrite rules will cause the desired ambiguity. A fairer comparison would allow that if part of the context is as yet unavailable, rule application does not proceed willy-nilly, but rather the whole derivational process is delayed. The effect of the mismatch between target string and input syllable would be swamped by the delay in the derivation of that syllable structure, waiting for the stress value on the next vowel.

The chapter ends, like the previous one, with some pointers to other work. Unlike Archangeli, Hammond is not shy of giving references to the literature, providing approximately one per sentence.

In summary, this chapter provides a clear and readable introduction to syllable and feet, and the interaction between stress and syllable structure. While this chapter seeks to offer more in the way of argument than chapter 1, these arguments are better thought of as illustrative of those in the OT literature rather than conclusive.

2.3 *Pulleyblank*

Pulleyblank's chapter, 'Optimality Theory and features', begins with a discussion of the mapping from inputs to outputs. As is *de rigueur* in generative phonology, the input is a phonological representation of the same type as the output. These two levels of representation can be mediated by rewrite rules or constraints on relations. After a *reductio* of the rule based account, on the basis that the simplest, and therefore most desirable, rule-based grammar would change nothing between input and output, Pulleyblank describes how the OT assumptions sidestep this particular pitfall: IN OT THERE IS NO SIMPLEST GRAMMAR. Because all OT constraints are stipulated to be present in all languages, and languages differ only in the ranking of constraints, Occam's razor cannot distinguish between them. Languages which violate Faithfulness are as simple as those which do not.

The flaw in this argument is the incomplete application of Occam's razor. The same rubric should eliminate constraints from grammars where they

play no role, contradicting the universality of the constraint set. If the input contains a fully fledged phonological representation, no constraint ranked lower than the last Faithfulness constraint can have any effect on selecting candidates, and so should be eliminated. If we do this, then the fewest constraints are had by grammars with only Faithfulness constraints, and we are back to the same conclusion as for the rule-based account.

Berwick (1985) convincingly argues that the simplicity of a linguistic model should not be based on the final grammar, but on how much input must be provided to the learner to acquire that grammar. Under this view, grammars which perform the reductions made by infants are simpler than those which preserve all of the relevant distinctions of the language. Simplicity is independent of whether the linguistic system is constraint-based or derivational.

The meat of the chapter is the presentation in detail of two pillars of OT: Faithfulness and Alignment. In line with the chapter's title, the examples mostly involve sub- or supra-segmental phenomena such as neutralisation and tone. An extended example from Margi allows the illustration of Alignment and Faithfulness in interaction.

This chapter serves the main purpose of the book well. It introduces linguistic phenomena, such as features and tone, and illustrates how OT can analyse them. The main shortfall occurs when the author offers arguments against sketchy accounts of rule-based grammars. There is no systematic pointer to other literature on the topic beyond a footnote on page 101.

2.4 *Russell*

The fourth chapter, 'Optimality Theory and morphology', is written by Kevin Russell. Like most of the other chapters, this one attempts to be a 'popular science' work by presuming no knowledge of linguistics in general, or morphology in particular. It begins with an introduction to what morphemes are, and crisply presents a statement of the three questions which (according to the author) define morphology. These are, to quote (box, p. 104): '(i) What pattern of sounds is associated with each morpheme? (ii) How is the right allomorph chosen for each context? (iii) What determines the order of morphemes in a word?' Russell then offers a two-page account of the 'classical' (i.e. SPE) approach to morpheme combination and allomorph construction. In short, each morpheme has a unique underlying phonological representation; these are concatenated to build words; and variation in the form of a morpheme results from the action of phonological transformations. Some difficulties with this model are pointed out: for example, unnatural and natural rules have equal status, and words are hard to separate into morphemes. Russell dwells on a third difficulty: assigning a rule ordering can sometimes seem impossible. A rule-ordering paradox from Paamese reduplication exemplifies this point.

The chapter then introduces once again the pillars of OT discussed in the previous chapter, this time from the point of view of morphology. The first pillar is CORRESPONDENCE THEORY. According to this theory, the relationship between instances of a morpheme – between underlying and surface, or between two surface realisations – is a phonological object which can be governed by constraints. The second pillar, ALIGNMENT, appears to explain the Tagalog *-um*-infix. Alignment is, rightly, described as a constraint schema rather than an actual constraint. In describing these concepts, Russell neatly shows how OT answers the challenges of morphology. OT shares the classical notion of a unique underlying phonological representation for morphemes. Variation arises however by the interaction of correspondence constraints with structure constraints which limit possible surface forms. Alignment constraints determine the ordering in which phonological material from different morphemes is realised.

This presentation accords Russell his point that OT is a morphological theory. One can only presume – it is not discussed – that OT avoids the unnaturalness of the classical model by relying on the a priori claim to naturalness through a fixed set of constraints: all and only possible human languages are defined by the ordering of a fixed constraint set.

Like Hammond, Russell takes many of his examples from English. Among the phenomena discussed is the merger of the plural and possessive *-s* markers (e.g., *cats*' rather than **cats*'*s*), but the failure of either of these to merge with root-final sibilants. This example is neat and clearly handled.

For the linguistic novice, to whom at least the beginning of the chapter is addressed, some of the technical terms appearing in the text without explanation will be a little obscure: NOUN PHRASE, RELATIVE CLAUSE, HEAD NOUN, MODIFIERS (124); THE MORPHOLOGICAL COMPONENT (125); ALVEOLAR STOP (128); SYNTACTIC ZERO-BAR LEVEL (129). This contrasts with the helpful explanations afforded to INFLECTIONAL and DERIVATIONAL (129).

It is interesting to note that while Russell's chapter largely endorses a 'classical' OT story of morphology, an earlier paper of his (unpublished 1995, referred to in the chapter) offers an account of morphemes as constraints themselves. A little of this perspective does in fact find its way into the chapter, in the form of constraint schemas, such as ALIGN, which are instantiated differently in each language. For example, English does not have a constraint aligning the Tagalog *-um-* with the beginning of a word – let alone rank it in the constraint hierarchy. This is a loophole in the OT claim (box on page 17) that cross-linguistic variation in OT arises from differences in constraint ordering. Tagalog and English differ also in the existence of *-um-* alignment constraints.

The chapter concludes with a very useful section listing some of the interesting areas of research in OT morphology, with at least one reference per topic. These are: the place of morphology in production, more on morpheme ordering, cyclicity, one form/one function, paradigms, and acquisition.

2.5 *Aside on Syntax in OT*

The pair of papers by Pesetsky and Speas combine to offer an interesting perspective on the difficulty of transposing OT from phonology and morphology to syntax. Before talking in more detail about their individual contributions (which diverge considerably), it seems worthwhile to briefly comment on issues which they both discuss.

A field in flux Practically the only thing on which Pesetsky and Speas concur is that nobody really agrees how syntax should be treated in OT. For Speas, the area is 'in its infancy' (178), while Pesetsky comments on his own chapter, 'every aspect of this discussion is controversial. […] Too little has been accomplished and too much is unknown' (169). The level of controversy is well illustrated by these two papers: while Pesetsky claims that crucial syntactic constraints are inviolable, Speas explores 'the possibility that *all* syntactic constraints are violable' (171).

The nature of the input What should the input to GEN be when the output is a set of syntactic candidates? Pesetsky fails to explicitly address this question, though he implicitly seems to follow Speas in suggesting that the input is something like a 'string of words' (179). One might ask further what is meant by 'words' here: lexemes or inflected word forms? Bare phonological strings, or words associated with various kinds of lexical information? One might conclude from the fact that Speas counts *Pro* as a 'word' that she intends the latter to encompass more than just the phonological dimension, since there is no strictly phonological difference between the various inaudibilia (*Pro*, *PRO*, NP-trace, *wh-trace*, etc.) postulated by GB theory and its descendants.[2]

A further question concerns the semantic relation between the input and the candidate set. While Pesetsky implies (together with Grimshaw (1997) and Bresnan (2000)) that candidates are semantically (presumably, at least truth-conditionally) equivalent, Speas follows Legendre et al. (1998) in claiming there is no 'semantic mapping in the input' (179). Her argument seems to be one of economy: since the output of EVAL needs to be semantically interpreted, the input needn't be. This argument is less than overwhelming; surely there are a range of theoretical and empirical factors which might help determine whether Faithfulness constraints are sensitive to semantic content.

[2] This rich notion of word is also endorsed in the book's Afterword (231).

2.6 *Pesetsky*

Pesetsky's chapter, 'Optimality Theory and syntax: movement and pronunciation', opens with a brief introduction to derivational syntactic theory, covering some basic notions of constituent structure, followed by a section arguing for movement analyses of V2 in German, scrambling in Japanese and *wh*-movement in West Ulster English. He suggests these two types of phenomena are governed by distinct 'systems of principles'; later, he argues for a third 'pronunciation system' (167) which determines whether expressions are phonetically realised, particularly when they are associated with more than one structural position.

Although the arguments in this initial section are presented succinctly and persuasively, it is unlikely that any reader who comes to the chapter without a grounding in this kind of syntax will be able to weather the considerably more challenging material on OT and syntax that occupies the remainder of Pesetsky's contribution.[3]

The main thrust of Pesetsky's chapter is to address the interesting question whether an OT organisation of constraints might be inappropriate for certain aspects of grammar, that is, areas of syntax which do not interact in an 'OT manner'.

Pesetsky's discussion starts with the notion of ineffability, exemplified by the interaction of the following two constraints:

> *Obligatory* wh-*movement*: An embedded question must have *wh*-phrase in SPEC(CP).
> *Constraint on Movement*: *Wh*-movement cannot extract an adverbial from a relative clause.

He points out that his example (5.26) – given here in (1) – 'with *how* understood as modifying *treat their subordinates*, is completely impossible' (147).

(1) *I'm wondering how the company will fire [any employees who treat their subordinates __].

Moreover, leaving *how* in place fails to yield a satisfactory output, as shown in (2), his (5.27).

(2) *I'm wondering the company will fire [any employees who treat their subordinates how].

[3] It is also worth noting that there are some potentially very confusing typos in this chapter, especially in the use of through-scoring (e.g. '~~que~~') to indicate an unpronounced expression in a tableau. For example, *que* should be unpronounced in the second candidate in (5.58), while the indications of non-pronunciation of *qui* and *que* in (5.59) are completely absent. Possibly the task of finding the correct representations could be set as an exercise for the more able student.

So 'ineffability', according to Pesetsky, means that 'there is *no* acceptable alternative to [1] that involves the same words and the same interpretation' (148). Although the OT model claims that the conflict should be resolvable by ranking one constraint over the other and picking the candidate that satisfies the most highly ranked candidate, here we have an input that seems to yield no acceptable output. Instead, the interaction of the two constraints in question suggests what he calls a CLASH & CRASH model; conflict blocks the existence of any acceptable output. In fact, although Pesetsky takes ineffability as evidence for a difference between syntax and phonology, it might have been used to establish a parallel: ineffability is a problem in phonology as well, at least during the period of acquisition. Menn & Stoel-Gammon (1995: 346) comment on the fact that children, rather than getting as close to difficult adult words as their own phonological repertoire allows, will instead refuse to say them. It seems that under the childhood system of constraints, FAITHFULNESS sometimes breaks down so badly that the utterance is not worth producing.[4]

Despite his remark quoted above, Pesetsky concedes (150) that OT does not require a syntactic output to use the same words and structures as the input, but rather that it should be maximally faithful to the input. Indeed, he suggests, 'the output might not even receive the same semantic interpretation as that which might be associated with the input'. As an example, he suggests that (4) might belong to the same candidate set as (3), his (5.36c).

(3) *Mary wonders which book Bill to buy at the store.
(4) Mary wonders which book Bill should buy at the store.

In order to provide a persuasive alternative to the Clash & Crash model, Pesetsky claims, an OT account of ineffability would have to show that there is some relevant set *G* of constraints which can be uniformly ranked with respect to a unified FAITHFULNESS constraint. If instead it turns out to be necessary to postulate a family of FAITH(X) constraints, the resulting complexity would compare unfavourably with the Clash & Crash approach.

Despite these sceptical remarks, Pesetsky then develops a detailed comparison of the 'pronunciation' of *wh*-expressions and complementizers in English and French relative clauses, where he argues that an OT factorial typology (involving three constraints) is more explanatory than a Clash & Crash account. In the course of this discussion, he introduces the notion of a TIE. This widening of the original OT formulation allows constraints to be ranked equally and their violations are given equal weight. Strangely, he doesn't entertain the possibility that constraints which Clash & Crash might

[4] As pointed out to us by an anonymous *JL* referee, similar arguments might be constructed on the basis of defective inflectional paradigms. For example, there appears to be no optimal candidate corresponding to the input consisting of the English lexeme *forego* and PAST tense: **forewent*, **foregoed*, … .

be regarded as tied. Nor does he explore the grave consequences of ties for the formal expressivity of the theory. The number of possible grammars using constraint ranking with ties is significantly greater[5] than the already huge factorial possibilities of standard OT.

Overall, Pesetsky's comparison between the two syntactic paradigms is thoughtful and provocative; he raises the issues in an unpolemical fashion and clearly acknowledges that more research needs to be carried out on the issue of ineffability. Perhaps one avenue to explore would allow a failure to parse some semantic content in the input (such as a finiteness specification) to be ranked low in the constraint hierarchy so long as the violation has no truth-conditional consequences, as seems to be the case with (3) and (4). Alternatively one could envisage a more abstract account of lexical-conceptual structure which would admit a wide range of paraphrase-equivalent candidates, including a number of acceptable counterparts to Pesetsky's (5.26) such as the following:

(5) I'm wondering for what treatment of subordinates the company will fire employees __ .

Pesetsky points out that the 'repair' of an island violation by a resumptive pronoun appears to offer prima facie support for the OT model:

(6) (a) *There is one worker who the company actually fired the employee that treated __ badly.
(b) There is one worker who the company actually fired the employee that treated him badly.

However, he then goes on to observe that (6b), despite its superiority to (6a), is only marginally acceptable, and therefore falls outside the compass of OT: 'There are no silver or bronze medals in OT' (169). Now, since the gradience of grammaticality judgements is a well-attested phenomenon, and is robustly supported by psycholinguistic research, it certainly seems worth asking whether a grammatical framework can provide an account of gradient well-formedness, rather than simply relegating such data to the limbo of 'performance'. On the face of it, OT would appear to be a good starting point for such an effort, and in fact there is a small body of work which pursues this direction, e.g. Hayes (2000) and Keller (1998). Choi (1996), Keller (1998) and Müller (1998) have also proposed drawing a distinction between HARD or GRAMMATICAL syntactic constraints, whose violation leads

[5] If there are N constraints, and no ties, standard OT allows $N!$ grammars. With ties, there are $\sum_{K=1}^{k=N}$ Part $(N; k)k!$ grammars, where Part$(N; k)$ is the number of ways N items can be divided into k non-empty partitions. For N ranging from 1 to 7, there are 1, 2, 6, 24, 120, 720, 5,040 respectively grammars without ties, and 1, 3, 13, 75, 541, 4,683, 47,293 with ties. For $N = 10$, the difference is 3.5 million vs. 100 million.

to outright ungrammaticality, versus SOFT or MARKEDNESS constraints, whose violation lessens acceptability in a gradient fashion. In some sense, this approach might be regarded as striving for a Hegelian synthesis of Pesetsky's two opposed paradigms.

2.7 *Speas*

The final chapter, by Speas, is entitled 'Optimality Theory and Syntax: Null Pronouns and Control'. As indicated above, Speas takes a very different approach from Pesetsky – by comparison with his circumspection, she seems positively gung ho about the extent to which syntax yields gracefully to an OT analysis.

Taking Chomsky's (1995) Minimalist Program (MP) as a reference point, Speas's key observation is that syntactic constraints are accompanied, explicitly or implicitly, by hedges or exceptions. The substance of this claim can best be conveyed by considering some of her examples (184), illustrated in table 1. Let's look in a little more detail at Speas' discussion of the Extended Projection Principle (EPP). Following Grimshaw & Samek-Lodovici (1995) (see also Grimshaw & Samek-Lodovici 1998), Speas renames the EPP 'SUBJECT'. Viewed as an OT constraint, SUBJECT simply requires the presence of subjects. In English, SUBJECT outranks Full Interpretation (FULLINT), which dictates that *There can be no superfluous symbols in an output representation*. Consequently, sentences containing the semantically uninterpreted expletive subjects *it* and *there* are optimal, despite violating FULLINT. Speas goes on to suggest that the converse ranking of these two constraints can explain the absence of expletive subjects in languages like Yaqui; illustrated in (7), her (6.16).

(7) yooko yuk-ne
tomorrow rain-FUT
'It will rain tomorrow.'

Instead of providing the EPP with a hedge, then, the strategy is to look for an independently motivated constraint which can outrank it. Speas's important insight is that once hedges are brought to the foreground as interesting conditions in their own right, rather than embarrassing afterthoughts, they can be seen as an inevitable consequence of the interaction between violable constraints.

After elaborating and defending this position, Speas turns to an account of the distribution of null pronouns. Starting with a fairly detailed illustration and OT analysis of the distribution of *Pro* in English, she then focusses on the cross-linguistic pattern of *Pro* occurring in finite clauses; forbidden in English, allowed for both subject and object position in Thai and Korean, and allowed only in subject position in Spanish and Mandarin.

The discussion of null pronouns also affords an opportunity to reinforce the message about hedges. Speas points out that standard versions of control,

Principle	Essence	Hedge
Extended Projection Principle	All clauses must have a subject…	except for languages which lack overt expletives.
Binding Principle A	An anaphor must be bound in its governing category…	unless it is one of a special class of anaphor which need not be bound.
X-bar Principles	Every category has a head, a specifier and a complement…	unless a given head takes no complement or has no features to check with its specifier.

Table 1
Syntactic constraints and hedges

such as given below, run into trouble with examples like those in (8) (her (6.30)). As well as requiring the highlighted hedge, the definition of control domain has to be complicated so that there is no domain in such examples.

Control: A null pronoun must be controlled in its control domain, if it has one.

(8) (a) *Pro* to behave better in public would help Bill's reputation.
(b) Pat asked how *Pro* to make flowers out of kleenex.
(c) It's fun *Pro* to dance.

By contrast, in Speas's approach, no hedge is required for the CONTROL constraint. Its violability allows the sentences in (8) to emerge as optimal.

For the reader who has already mastered some version of Chomskyan syntax, this chapter provides a useful overview of how OT can provide an elegant account of interacting grammatical principles. In addition, the picture it presents of cross-linguistic patterns in the distribution of empty pronouns attests well to this OT story about variation.

2.8 *Afterword*

Before the references and index, there is a section of uncertain provenance entitled 'Afterword'.[6] A better title might have been 'Speculations on the nature of the input'. Whoever the author or authors, it seems as though they are placing here their own views on the input in phonology, morphology and syntax, so as to have the final word after the main chapters in the book.

The Afterword starts well, pithily distinguishing what the input must, may and cannot have. It considers how these three categories are pursued in different OT formulations of the input: lexical optimisation, minimal specification and an agnostic OT perspective.

After an ethnocentric claim about US English,[7] the paper considers the importance of paradigms for finding the necessary form of the input. A second topic is then introduced, one returned to again and again, namely the question of how much ordering is needed between elements in the input. Prosody, meanwhile, is shown to not be part of the input, at least in some words of English.

Attention then shifts to morphology. A short account of blocking is presented (206–207), with the example of the mono-morphemic *went* blocking a bi-morphemic form **go-ed*, courtesy of a constraint dubbed MONOMORPH which outranks FAITHFULNESS. In his chapter, Russell (128) briefly discussed

[6] The subject matter and style point to Archangeli as author, and in her chapter alone, there is a forward reference to the Afterword, with no author attribution.

[7] 'In most dialects of English, the words *petal* and *pedal* are pronounced identically' (204). This is a feature primarily of American English dialects. However, more dialects of English are identifiable in the British Isles than elsewhere, despite the U.S. having more native speakers (Wells 1982).

the case of subregular past tenses in English, such as *bent* < *bend*. In input, these past tense forms are presented as bi-morphemic. Furthermore, according to the *Oxford English dictionary* (Simpson & Weiner 1991: 2296) *went* is a subregular past tense of *wend*, albeit now used suppletively as the past tense of *go*. So according to Russell's approach, *went* is arguably not mono-morphemic after all, and hence resists the proposed account. While not damaging to the principle of blocking arguments in OT, this problem is indicative of the uneven reference the Afterword pays to the chapters earlier in the book.

The focus of this morphology section is to argue for three potential structures for words: inputs are mono-morphemic, or they consist of morphemes welded into compounds, or they combine roots with productive affixes. Rather than encoding order in the representation, a new device is introduced: tagging a morpheme with a syntactic tableau. This concept is not clearly explained, but purports to obviate the need for morpheme sequencing. So much so that the author reports that morpheme order must not be represented in the input (213).

3. Conclusion

A&L offers many clear and interesting introductions: to generative linguistics, to OT, to phonology, to morphology and to syntax. There are many well-presented examples, and few obvious gaffes or errors.

When it goes beyond this, it falls short. In discussions of extensions to standard OT, there is a shortage of references to this work, other than a general reference to the ROA. Similarly, the book does not provide more than illustrative arguments for OT against its seemingly only competitor in phonology, SPE-style rewrite rules systems.

In summary, then, if you want to know what OT is, and would appreciate a gateway to finding out what topics are being tackled in the field, this book may well be for you. However, if you want to learn how to do OT yourself, by analysing linguistic problems, this book won't teach you – Kager (1999) may be more helpful, at least in the case of phonology. And if you want to be convinced that OT is a better theory than the alternatives, you will need to look elsewhere.

REFERENCES

Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M. & Pesetsky, D. (eds.) (1998). *Is the best good enough?* Cambridge, MA: MIT Press.
Berwick, R. C. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.
Bresnan, J. (2000). Optimal syntax. In Dekkers et al. 334–385.
Choi, H.-W. (1996). *Optimizing structure in context: scrambling and information structure*. Ph.D. dissertation, Stanford University.
Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
Cutler, A., Mehler, J., Norris, D. & Segui, J. (1983). A language specific comprehension strategy. *Nature* **304**. 159–160.

Dekkers, J., van der Leeuw, F. & van de Weijer, J. (eds.) (2000). *Optimality Theory*: *phonology*, *syntax*, *and acquisition*. Oxford: Oxford University Press.
Grimshaw, J. (1997) Projection, heads and Optimality. *Linguistic Theory* **28**. 373–422.
Grimshaw, J. & Samek-Lodovici, V. (1995). Optimal subjects and subject universals. In Beckman, J., Dickey, T. W. & Urbanczyk, S. (eds.), Papers in Optimality Theory (University of Massachusetts Occasional Papers in Linguistics **18**). Department of Linguistics, University of Massachusetts. 589–605.
Grimshaw, J. & Samek-Lodovici, V. (1998). Optimal subjects and subject universals. In Barbosa et al., 193–219.
Hayes, B. P. (2000). Gradient well-formedness in Optimal Theory. In Dekkers et al. 88–120.
Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
Keller, F. (1998). Gradient grammaticality as an effect of selective constraint re-ranking. In Gruber, M. C., Higgins, D., Olson, K. S. & Wysocki, T. (eds.), *Papers from the* 34*th Meeting of the Chicago Linguistic Society*, vol 2. Chicago: University of Chicago. 95–109.
Kenstowicz, M. & Kisseberth, C. (1979). *Generative phonology*: *description and theory*. New York: Academic Press.
Legendre, G., Smolensky, P. & Wilson, C. (1998). When is less more? Faithfulness and minimal links in *wh*-chains. In Barbosa et al. 249–289.
Menn, L. & Stoel-Gammon, C. (1995). Phonological development. In Fletcher, P. & MacWhinney, B. (eds.), *A handbook of child language*. Oxford: Blackwell. 335–359.
Müller, G. (1998). German word order and Optimality Theory. Working Papers of the SFB 340, Report 126, University of Tübingen.
Russell, K. (1995). Morphemes and candidates in optimality theory. Rutgers Optimality Archive, ROA-44, http://ruccs.rutgers.edu./roa.html.
Simpson, J. A. & Weiner, E. S. C. (eds.) (1991). *The compact Oxford English dictionary*. Oxford: Oxford University Press.
Wells, J. C. (1982). *Accent of English*. Cambridge: Cambridge University Press.

Authors' addresses:

(Ellison)
Institute for Adaptive and Neural Computation,
Division of Informatics,
University of Edinburgh,
5 Forrest Hill,
Edinburgh EHI 2QL,
U.K.,

E-mail: T.Mark.Ellison@ed.ac.uk

(Klein)
Institute for Communicating and
Collaborative Systems,
Division of Informatics,
University of Edinburgh,
2 Buccleuch Place,
Edinburgh EH8 9LW,
U.K.
E-mail: Ewan.Klein@ed.ac.uk