Phonology **26** (2009) 217–226.
doi:10.1017/S0952675709001766 Printed in the United Kingdom

Review

Paul Smolensky and Géraldine Legendre (2006). *The harmonic mind: from neural computation to optimality-theoretic grammar*. Cambridge, Mass.: MIT Press. Vol. 1: *Cognitive architecture*. Pp. xxiv + 563. Vol. 2: *Linguistic and philosophical implications*. Pp. xxiv + 611.

Joe Pater
University of Massachusetts, Amherst

The harmonic mind is a two-volume collection of 23 chapters authored by Smolensky, Legendre and their collaborators.* Some of these are reprints or slight updates of previously published material, but approximately three-quarters of it has never appeared before. Every phonologist who takes at all seriously the notion that phonology is a branch of cognitive science should own a copy of this book.

Smolensky & Legendre (vol. 1, p. 1) present a view of cognition that incorporates two distinct, but related levels of formal description: 'the continuous, numerical lower-level description of brain', characterised in terms of a connectionist network, and 'the discrete, structural, higher-level description of mind', characterised in terms of Optimality Theory (OT). The depth and breadth of the presentation are awe-inspiring, and quite likely intimidating in some places for any single reader. In their preface, Smolensky & Legendre characterise the intended audience as interdisciplinary groups of researchers in cognitive science, and a joint reading of this book by such a group would undoubtedly be extremely educational for all. Bridges between the disciplines are built by careful exposition of the fundamental ideas, and the experts in each area will be satisfied by the wealth of insights, and the formal precision of their presentation. Since phonology is a central focus, this book provides phonologists with a unique opportunity to engage with both the research and the researchers in other realms of cognitive science.

There is furthermore much that phonologists (and other linguists) working on their own can gain from this book. The presentation of the overall cognitive framework in the first few chapters of Volume 1, *Cognitive architecture*, should be accessible to readers with little or no background in mathematics or computational modelling of cognition. This section should particularly appeal to phonologists curious about connectionism, but who are put off by reductionist applications that seek to eliminate the types of representations used in generative

* I would like to especially thank Paul Smolensky for sharing an advance copy of portions of the manuscript, and for comments on an earlier version of this review. Thanks also to the participants in Ling 730, Fall 2005 for discussion, and to Michael Becker, Shigeto Kawahara and John McCarthy for comments on the draft review.

linguistics. Another key point of entry to the material in the first volume is Smolensky & Legendre's 'Formalizing the principles II: optimization and grammar', which discusses the relationship between connectionist optimisation, grammatical well-formedness and formal language theory. A further point of access to the model of cognition for linguists is found in Smolensky's 'Principle-centered connectionist and generative-linguistic theories of language' in Volume 2. The final paper in Volume 1, 'Optimality theory: the structure, use, and acquisition of grammatical knowledge', by Paul Smolensky, Géraldine Legendre & Bruce Tesar, consists of an extremely lucid exposition of OT in the context of the overall generative enterprise, along with a set of cogent responses to common worries about the theory. Another chapter in Volume 1 worth highlighting for its interest to linguists is Smolensky's 'Optimization in neural networks: harmony maximization'. For anyone who would like to seriously explore the connectionist formalisation of linguistic grammar, this chapter will serve as an essential guide to the relevant properties of different types of neural network.

Volume 2, *Linguistic and philosophical implications*, is more directly targeted at a linguistic audience. For phonologists, Smolensky's 'Optimality in phonology II: harmonic completeness, local constraint conjunction, and feature domain markedness' is a watershed publication. In it, Smolensky presents for the first time in print his already widely influential theory of local constraint conjunction, and also a theory of headed domains for phonological representations. Smolensky shows that headed domains allow a constraint-based theory to deal with phenomena in vowel harmony that can otherwise pose a considerable challenge, including directionality effects and other restrictions on the 'trigger' of spreading (see e.g. Cole & Kisseberth 1995, McCarthy 2004 for related proposals in OT). This theory of vowel harmony has received relatively little attention in comparison with other OT accounts based on Agree constraints (e.g. Baković 2000) or on Alignment constraints (e.g. Kaun 2004). Smolensky's successful headed domains reanalysis of Archangeli & Pulleyblank's 1994) account of Lango ATR harmony, which involves a particularly complex interacting set of conditions on the direction, trigger and target of spreading, provides ample motivation for further exploration of this approach.

The local conjunction (henceforth LC) theory of constraint interaction is better known than the headed domains theory of representation, and is also much more controversial (see e.g. McCarthy 1999, 2003, Padgett 2002 for critiques; I draw on these below). Much of the controversy surrounds the worry that LC introduces into OT the power to generate unattested linguistic patterns that are ruled out by the strict domination property of ranked constraints. The relationship between LC and a version of OT that uses weighted constraints, as in OT's predecessor, Harmonic Grammar (henceforth HG), is addressed in Géraldine Legendre, Antonella Sorace & Paul Smolensky's chapter 'The Optimality Theory–Harmonic Grammar connection'. The remainder of this review focuses on this relationship, which as Legendre *et al.* point out, can easily be misconstrued. Throughout, references to Legendre *et al.* are to the just-cited chapter, and those to Smolensky are to 'Optimality in phonology II'.

LC is an elaboration of OT originally proposed by Smolensky in a series of unpublished, but much-cited conference presentations. The literature on

OT-LC is now quite large: along with the work cited in Smolensky's chapter, an especially important further source is Ito & Mester's (2003) monograph. Smolensky (vol. 2, p. 43) defines local conjunction as follows:

(1) *Local conjunction within a domain* $\mathcal{D}$

$*A \&_{\mathcal{D}} *B$ is violated if and only if a violation of *A and a (distinct) violation of *B both occur within a single domain of type $\mathcal{D}$.

This operation of conjunction yields a new constraint that is separately rankable in the constraint hierarchy.

The original motivation for OT-LC, and the one discussed at length in Smolensky's chapter, is to reduce complex markedness constraints to more basic primitives. A simple example that appears in his earlier unpublished work, but not in this chapter, is that of a constraint against voiced obstruent codas (see also Ito & Mester 2003). This can be expressed as the conjunction of two independently needed constraints, NoCoda ('assign a violation mark to a consonant in syllable final position') and *VoiceObs ('assign a violation mark to a voiced obstruent'). When these are conjoined in the domain of a segment, the resulting constraint penalises voiced obstruent codas, but neither voiceless codas nor voiced obstruents. If the conjoined constraint is ranked above the faithfulness constraint Ident-Voice ('assign a violation mark to a segment whose Input and Output voice specification differs'; McCarthy & Prince 1999), and the component constraints are ranked beneath Ident-Voice, the result is coda devoicing:

(2) *Local conjunction analysis of final devoicing*

/bad/	NoCoda & *VoiceObs	Ident-Voice	NoCoda	*VoiceObs
☞ a. [bat]		*	*	*
b. bad	*!		*	**
c. pat		**!	*	
d. pad	*!	*	*	*

Obviously, the devoicing pattern could not be produced by just the unconjoined constraints in OT. Perhaps less obviously, it also cannot be produced by weighted constraint interaction in HG.

Before showing why HG cannot duplicate this effect of LC, it's worth first seeing how weighted versions of OT constraints can produce the devoicing pattern if we include *CodaVoice, which, like the conjoined constraint, assigns violations to voiced codas. In HG, optimality is defined in terms of maximal numeric harmony: the optimal form has the highest summed weighted violation score. The HG tableau in (3) adds a row to show the weights of each constraint, and a column to show the harmony scores of each candidate. Following a convention introduced by Legendre *et al.*, the violation marks are changed into the corresponding negative integers. Because *CodaVoice has greater weight than Ident-Voice, a candidate that incurs one violation of that constraint, like [bad], receives a greater weighted penalty ($2 \times -1 = -2$) than a candidate that has a single violation of Ident-Voice, like [bat] ($1 \times -1 = -1$). Since [bat] has the highest harmony in the candidate set, it is optimal.

(3) *Final devoicing in HG*

	2	1	
/bad/	*CodaVoice	Ident-Voice	
a. bad	−1		−2
☞ b. [bat]		−1	−1
c. pat		−2	−2
d. pad	−1	−1	−3

To help show why HG cannot replicate the effect of local conjunction illustrated in (2), the table in (4) provides the violation profiles of the candidates using negative integers. I leave out the final, harmonically bounded, candidate from (3), whose violation marks on the basic constraints would be identical to [bat].

(4) *Violation scores of unconjoined constraints*

/bad/	Ident-Voice	NoCoda	*VoiceObs
a. bat	−1	−1	−1
b. bad		−1	−2
c. pat	−2	−1	

The question is whether there exists some set of weights that will choose [bat] as optimal.

The answer is no. For [bat] to beat [bad], the weight of *VoiceObs must be greater than that of Ident-Voice. And for [bat] to beat [pat], the reverse relationship must hold. These inequalities are inconsistent, just as the rankings would be. Both [bat] and [bad] violate NoCoda; that shared violation cannot affect the choice between them.

Legendre *et al.* (vol. 2: 348 ff) provide a similar example of the inability of HG to handle a pattern analysed in terms of OT-LC, and show that this inability generalises to a set of cases discussed in Smolensky's chapter. The example is that a ban on [x] cannot be reduced to independent constraints against velars and fricatives, when the inventory contains both velars and non-velar fricatives. The other cases are other OT-LC analyses of 'basic inventories'. As they point out, in comparison with HG's linear model (the harmony function is a linear equation), LC is a superlinear theory of constraint interaction.

In linguistic theory, greater power is of course a double-edged sword. The superlinearity of OT-LC also allows it to generate unattested linguistic patterns not generated by HG. One example, discussed first by Itô & Mester (1998), can be produced by conjoining NoCoda with Ident-Voice in the domain of the segment (thanks to Matt Wolf for bringing this to my attention).

(5) *Local conjunction analysis of initial devoicing*

/bad/	NoCoda & $_{Seg}$Ident-Voice	*VoiceObs	NoCoda	Ident-Voice
☞ a. [pad]		*	*	*
b. bat	*!	*	*	*
c. pad		**!	*	
d. pat	*!		*	*

The *VOICEOBS ≫ IDENT-VOICE ranking leads to devoicing, which is blocked by the dominant conjoined constraint when it occurs in coda position, where it violates both NOCODA and IDENT-VOICE. This pattern is referred to as a markedness reversal, since the feature is now protected in only the marked environment (see Łubowicz 2005 for further cases).

With just the unconjoined constraints in (5), this pattern cannot be produced by HG: [bat] and [pad] have identical violation marks, and no weighting will make [pad] solely optimal. In a full OT/HG constraint set that includes *CODAVOICE, [bat] harmonically bounds [pad]. The fact that [pad] can be made optimal in OT-LC but not HG points to an important difference between the theories: pairwise harmonic bounding relationships from OT are preserved in a version of HG with positive weights (Prince 2002), but not in OT-LC.

Unattested systems can also be produced by conjoining markedness constraints. One such system devoices an obstruent if there is a coda anywhere in the word; this system emerges if we conjoin NOCODA and *VOICEOBS over the domain of the word.

(6) *Local conjunction analysis of long-distance 'coda devoicing'*

	/balatak/	NOCODA & $_{Wd}$*VOICEOBS	IDENT-VOICE	NOCODA	*VOICEOBS
a.	i. balatak	*!		*	*
	☞ ii. [palatak]		*	*	
b.	/balata/				
	☞ i. [balata]				*
	ii. palata		*!		

Again, this pattern cannot be replicated by HG, for the same reason that it cannot do local coda devoicing with the unconjoined version of these constraints: the shared violation of NOCODA is irrelevant to the choice between [balatak] and [palatak].

This last case is an instance of the locality problem for LC. To deal with it, Łubowicz (2002) proposes to limit LC to conjunctions over the smallest shared domain of two constraints. However, LC also produces unattested patterns when constraints are conjoined over their smallest shared domain. To construct an example, we can add a constraint that forces place assimilation; for familiarity's sake, I will use AGREE-PLACE ('assign a violation mark to every pair of adjacent of non-homorganic segments'). The first two tableaux show the effect of *CODAVOICE and AGREE-PLACE being dominated by the relevant faithfulness constraints (IDENT-PLACE penalises change in place of articulation): neither coda devoicing nor place assimilation usually applies. In the last tableau, we see that when a consonant with underlying voicing is in the context of a non-homorganic consonant, it devoices. Although place assimilation and coda devoicing are well attested, this cumulative interaction is to my knowledge undocumented.

(7) *Cumulative interaction of coda devoicing and place agreement in OT-LC*

a.

/bad/	IDENT-VOICE	*CODA-VOICE
i. bat	*!	
☞ ii. [bad]		*

b.

/sat+ma/	IDENT-PLACE	AGREE-PLACE
☞ i. [satma]		*
ii. sapma	*!	

c.

/bad+ma/	*CODA-VCE & $_{Seg}$AGR-PL	ID-PL	ID-VCE	*CODA-VCE	AGR-PL
☞ i. [batma]			*		*
ii. babma		*!		*	
iii. badma	*!			*	*

In HG, this unattested cumulative interaction is impossible with these constraints. If IDENT-VOICE has a greater weight than *CODAVOICE, as needed to block devoicing in the first tableau, then [badma] will be optimal in the last tableau. Here, the shared violation of AGREE-PLACE is irrelevant to the choice between [batma] and [badma].

This is an instance of a co-relevance problem for LC. Empirically, place agreement is not relevant to coda voicing, but LC allows these independent factors to interact. Baković (2000) and Łubowicz (2005) propose restrictions on LC that are aimed at addressing this problem. Interestingly, the restriction that Łubowicz (2005) imposes on LC cumulative interactions emerges from the nature of HG cumulativity: for two markedness constraints to have a cumulative effect, they must be satisfied by violating a single other (faithfulness) constraint. For example, in Japanese loanword devoicing (Kawahara 2006), a voiced geminate is devoiced only in the presence of another voiced obstruent. In Pater (2009), this is analysed as IDENT-VOICE being violated only if both OCP-VOICE and *VOICE-GEMINATE are at issue. Because this sort of asymmetric trade-off in violations is a prerequisite for cumulative effects in HG, it imposes inherent restrictions on such interactions that must be stipulated in the LC approach.

These instances of restrictiveness of HG relative to LC may come as a surprise to some readers, as they did to me and the other participants in a graduate seminar when we read Legendre *et al.* for the first time.[1] Two questions immediately arise. The first is whether HG might actually be sufficiently restrictive to serve as a viable alternative to OT as a theory of grammar. Along with Prince & Smolensky (1993), Legendre *et al.* argue that this is not the case; see Pater (2009) for a detailed consideration of these arguments that yields a more positive conclusion for the prospects of HG. One point that is relevant to the

[1] It would be a mistake to claim that HG and LC are in a subset relation in terms of the languages they generate. Because LC requires that constraints be conjoined over some domain, there are instances of HG interaction that cannot be captured by LC (McCarthy 2002). For instance, Moreton & Smolensky (2002) claim that MAX and markedness constraints do not share a common domain, so they cannot be conjoined, but there is no reason that they could not participate in a gang effect in HG.

current HG-LC comparison is that the unattested pattern that Legendre *et al.* show that HG can problematically generate can equally be produced by self-conjoining their 'MAINSTRESSR' constraint. This constraint assigns a violation to each syllable separating a main stress from the right edge the word. The self-conjoined version would produce a 'Power Hierarchy' of constraints, with one penalising stress one syllable from the edge, another penalising stress two syllables away, and so on. These constraints could be used to produce a stress window of any size, just as Legendre *et al.* show that weightings of MAINSTRESSR can.

The other question is to what extent HG can, and should, deal with patterns that escape the original version of OT, and which have been attributed to LC. This is a very complex question, partly because some of the phenomena attributed to OT-LC may be best explained not as cumulative interaction at all (see recently McCarthy (2007) on chain shifts as well as Wolf (2008) on derived environment effects; cf. Łubowicz (2002) and Moreton & Smolensky (2002) in OT-LC). Here I will discuss only the examples already mentioned in this review.

First, whether complex markedness constraints like NOCODAVOICE above, and *[x] in Legendre *et al.*, should be reduced to more basic constraints is a matter of some controversy. The idea that NOCODAVOICE is the sum of the effects of a constraint against codas and one against voiced obstruents hearkens back in some ways to theories of prosodic licensing (e.g. Itô 1986, Goldsmith 1990, Lombardi 1991, Steriade 1995), which see contextual markedness as the inability of marked prosodic contexts to license marked segments. Much work in OT has questioned this approach, analysing contextual markedness as the effect of rather specific, substantively motivated constraints (e.g. Pater 1999, Steriade 1999). The empirical motivation for these alternatives is that contextual markedness displays asymmetries that are not captured by prosodic licensing: the set of marked contexts is not the same for every marked segment, and markedness relationships between segments can be reversed across contexts (see Barnes 2006 for an extensive recent critique of prosodic licensing theory). Like prosodic licensing, OT-LC also fails to express these asymmetries.

Smolensky's comparison of the OT-LC analysis of Lango ATR harmony with that of Archangeli & Pulleyblank (1994) brings up a concrete example of this issue (see Potts *et al.* 2009 for a reanalysis of the Lango data in HG). In Lango, only [+high] vowels trigger regressive ATR harmony. Smolensky (vol. 2, p. 91) analyses this as the conjunction of the constraint against [−high, +ATR] vowels with the constraint demanding that heads be leftmost in the domain. Leftward spreading from a non-high [+ATR] vowel would violate this constraint. Smolensky (vol. 2, p. 138) points out that the rule-based analysis predicts that a rule that spreads ATR only from the marked vowel could equally exist, and cites Archangeli & Pulleyblank (1994) for the observation that the resulting pattern does not exist. It is clear, however, that other features do preferentially spread from a marked vowel: Kaun (2004) discusses a typology of rounding harmony in OT that instantiates just this observation.

Whether building a phonological theory with very specific phonetically grounded universal constraints is the right response to such asymmetries is of course also a matter of controversy (see Blevins 2004, Hayes *et al.* 2004). However, the greater reductionism possible in an OT-LC account of contextual markedness does not seem to be a knockdown argument for it over OT and HG, especially since that reductionism is only obtained in a somewhat abstract sense, since Smolensky proposes no mechanism for the learning of conjoined constraints.

If the conjecture that HG can serve as a theory of grammar proves correct, the consequences for Smolensky & Legendre's integrated connectionist-symbolic theory of mind, and for the general relationship of grammar to theories of learning and processing, are quite deep. Smolensky & Legendre aim to provide an explicit mathematical characterisation of the relationship between neural processing and higher-level symbolic computation, including linguistic grammar. They see HG as a bridge between these domains, with OT as the theory of grammar. However, this bridge is not yet fully built. Building on discussion in Prince & Smolensky (1993: 236), Legendre *et al.* (vol. 2, p. 347) explain the formal difficulties in constructing a connectionist harmony maximising network that is guaranteed to behave as a strict domination hierarchy. HG grammars, on the other hand, have been realised in connectionist networks (see also Goldrick & Daland, this issue). In phonology, Legendre *et al.* present connectionist implementations of HG versions of Prince & Smolensky's (1993) analyses of Berber syllabification, while Melanie Soderstrom, Donald W. Mathis & Paul Smolensky's chapter 'Abstract genomic encoding of Universal Grammar in Optimality Theory' provides a connectionist learning algorithm for HG, along with a proposal for how universal constraints might be encoded in the genome (the title of the chapter is somewhat misleading, in that the model uses weighted, rather than ranked constraints).

Smolensky, Legendre and their colleagues revolutionised linguistic theory by introducing optimisation into generative grammar. As readers of this journal certainly know, the impact has been especially strong in phonology. Along with some contemporaries (e.g. Goldsmith 1990 and papers in Goldsmith 1993), they can also be credited with introducing numerically weighted constraints into generative linguistics in their 1990 presentation of HG (an update of their 1990 work with Yoshiro Miyata: 'The interaction of syntax and semantics: a Harmonic Grammar account of split intransitivity' appears in Volume 1). Current statistical approaches to constraint interaction like the Maximum Entropy model of Goldwater & Johnson (2003) are formally stochastic versions of HG. Not only does *The harmonic mind* clearly elaborate the connectionist foundations for OT, but it also opens up vast new territory for phonological theory, for formal analysis of phonological systems as well for the study of learning and other types of processing.

REFERENCES

Archangeli, Diana & Douglas Pulleyblank (1994). *Grounded phonology*. Cambridge, Mass.: MIT Press.

Baković, Eric (2000). *Harmony, dominance, and control*. PhD dissertation, Rutgers University.

Barnes, Jonathan (2006). *Strength and weakness at the interface: positional neutralization in phonetics and phonology*. Berlin & New York: Mouton de Gruyter.

Blevins, Juliette (2004). *Evolutionary Phonology: the emergence of sound patterns*. Cambridge: Cambridge University Press.

Cole, Jennifer & Charles Kisseberth (1995). An optimal domains theory of harmony. Ms, University of Illinois.

Goldsmith, John (1990). *Autosegmental and metrical phonology*. Oxford & Cambridge, Mass.: Blackwell.

Goldsmith, John (ed.) (1993). *The last phonological rule: reflections on constraints and derivations*. Chicago: University of Chicago Press.

Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenador, Anders Eriksson & Östen Dahl (eds.) *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm: Stockholm University. 111–120.

Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.) (2004). *Phonetically based phonology*. Cambridge: Cambridge University Press.

Itô, Junko (1986). *Syllable theory in prosodic phonology*. PhD dissertation, University of Massachusetts, Amherst. Published 1988, New York: Garland.

Itô, Junko & Armin Mester (1998). Markedness and word structure: OCP effects in Japanese. Ms, University of California, Santa Cruz.

Ito, Junko & Armin Mester (2003). *Japanese morphophonemics: markedness and word structure*. Cambridge, Mass.: MIT Press.

Kager, René, Harry van der Hulst & Wim Zonneveld (eds.) (1999). *The prosody–morphology interface*. Cambridge: Cambridge University Press.

Kaun, Abigail R. (2004). The typology of rounding harmony. In Hayes *et al.* (2004). 87–116.

Kawahara, Shigeto (2006). A faithfulness ranking projected from a perceptibility scale: the case of [+voice] in Japanese. *Lg* **82**. 536–574.

Lombardi, Linda (1991). *Laryngeal features and laryngeal neutralization*. PhD dissertation, University of Massachusetts, Amherst.

Łubowicz, Anna (2002). Derived environment effects in Optimality Theory. *Lingua* **112**. 243–280.

Łubowicz, Anna (2005). Locality of conjunction. *WCCFL* **24**. 254–262.

McCarthy, John J. (1999). Sympathy and phonological opacity. *Phonology* **16**. 331–399.

McCarthy, John J. (2002). *A thematic guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCarthy, John J. (2003). Comparative markedness. *Theoretical Linguistics* **29**. 1–51.

McCarthy, John J. (2004). Headed spans and autosegmental spreading. University of Massachusetts, Amherst. Available as ROA-685 from the Rutgers Optimality Archive.

McCarthy, John J. (2007). *Hidden generalizations: phonological opacity in Optimality Theory*. London: Equinox.

McCarthy, John J. & Alan Prince (1999). Faithfulness and identity in Prosodic Morphology. In Kager *et al.* (1999). 218–309.

Moreton, Elliott & Paul Smolensky (2002). Typological consequences of local constraint conjunction. *WCCFL* **21**. 306–319.

Padgett, Jaye (2002). Constraint conjunction versus grounded constraint subhierarchies in Optimality Theory. Ms, University of California, Santa Cruz. Available as ROA-530 from the Rutgers Optimality Archive.

Pater, Joe (1999). Austronesian nasal substitution and other NC̥ effects. In Kager *et al.* (1999). 310–343.

Pater, Joe (to appear). Weighted constraints in generative linguistics. *Cognitive Science*.

Potts, Chris, Joe Pater, Karen Jesney, Rajesh Bhatt & Michael Becker (2009). *Harmonic Grammar with Linear Programming: from linguistic systems to linguistic typology*. Ms, University of Massachusetts, Amherst.

Prince, Alan (2002). Anything goes. In Takeru Honma, Masao Okazaki, Toshiyuki Tabata & Shin-ichi Tanaka (eds.) *A new century of phonology and phonological theory: a Festschrift for Professor Shosuke Haraguchi on the occasion of his sixtieth birthday*. Tokyo: Kaitakusha. 66–90.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.
Steriade, Donca (1995). Underspecification and markedness. In John Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell. 114–174.
Steriade, Donca (1999). Phonetics in phonology: the case of laryngeal neutralization. *UCLA Working Papers in Linguistics: Papers in Phonology* **3**. 25–145.
Wolf, Matthew (2008). *Optimal interleaving: serial phonology-morphology interaction in a constraint-based model*. PhD dissertation, University of Massachusetts, Amherst.