mfm FRINGE meeting

A workshop, timed to coincide with the 23rd Manchester Phonology Meeting, organised by: Ricardo Bermúdez-Otero, Patrick Honeybone & Marc van Oostendorp

W(h)ither OT?

Wednesday 27th May 2015

Arts Theatre, Samuel Alexander Building
University of Manchester

2.00-2.10	**Opening**	
2.10-2.55	**Joe Pater**	*Violable constraints in Classical Universalist Phonology and beyond.*
2.55-3.40	**Wendell Kimper**	*What changes and what stays the same: is Harmonic Serialism with positive constraints still Optimality Theory?*
3.40-4.25	**Michael Becker**	*MaxEnt as a baseline theory of grammar.*
4.25-4.40	**Break**	
4.40-5.25	**Bert Vaux**	*Evaluating OT: Abkhaz stress requires process-specific directional iterative intra-morphemic computation*
5.25-6.10	**Tobias Scheer**	*We need a theory of computation.*
6.10-6.25	**Break**	
6.25-7.10	**Jeffrey Heinz & Bill Idsardi**	*Beyond Optimality: the computational nature of phonological maps and constraints.*
7.10-7.30ish	**Discussion**	

Marc van Oostendorp and **Ricardo Bermúdez-Otero**, as discussants, will get the first comment/question after each speaker.

W(h)ither OT?

Abstracts

Joe Pater (University of Massachusetts Amherst)

"Violable constraints in Classical Universalist Phonology and beyond."

Since the advent of OT in the mid nineties, the empirical scope and theoretical goals of the field of phonology have changed quite dramatically. In what we might call "Classical Universal Phonology" (CUP), which started with the principles and parameters work of the early eighties, phonologists construct analyses of individual languages, and aim for theories of Universal Grammar that generate all and only the possible systems that are inferred from this analytic work. OT as it was originally proposed is firmly in this tradition. CUP continues to be fruitfully practised in much of the current work in OT and its violable constraint relatives, Harmonic Grammar and Harmonic Serialism, as well as in other phonological frameworks. Increasingly, however, phonologists are studying knowledge of phonology and how it is acquired, Universal Grammar, and typology in new ways. I'll point out some ways in which violable constraints are useful in research that goes beyond CUP, which suggest to me that the shelf-life of OT, broadly construed, will be quite long for phonology, also broadly construed. I'll also include some reminders of ways that violable constraints are useful in CUP.

Wendell Kimper (University of Manchester)

"What changes and what stays the same: is Harmonic Serialism with positive constraints still Optimality Theory?"

In its original formulation, Optimality Theory presented both an architectural framework – the components GEN, CON, and EVAL – and specific theoretical proposals about the contents and functioning of those components. In the years since, many proposed modifications to these original specifics have been put forward: the theory of CON has been enriched to include a number of new constraint types (including targeted constraints, anti-faithfulness constraints, morphological indexation, constraint conjunction, and others) and modifications to GEN and EVAL have produced Harmonic Serialism and Harmonic Grammar respectively. In this talk, I'll discuss some of the ways in which modifications of separate components are interconnected, focusing in particular on the opportunity that a restricted GEN allows for positively-formulated constraints in CON. I'll also argue that it's the enduring framework of OT itself that has made this kind of exploration both fruitful and productive – at its core, OT is a powerful architecture for generating predictions, and as such will continue to remain relevant beyond the lifespan of any specific proposal within it.

Michael Becker (Stony Brook University)

"MaxEnt as a baseline theory of grammar."

1. I will suggest that we adopt MaxEnt grammars as the baseline theory. Reasons:
   (a) the constraint weights are learnable using general-purpose, mathematically-sound, well-understood algorithms, unlike OT with ranked constraints
   (b) since MaxEnt is a kind of regression, it allows tighter integration between theory and experimental results, as both are analyzed with regression models
   (c) variation and gradience are built-in and don't require any additional mechanisms
   (d) the integration of biases (phonetic and otherwise) is also built-in; a beautiful recent example is Jamie White's work on saltation

2. Serial vs. parallel grammars: an orthogonal question that can be pursued separately. The theory can use MaxEnt grammars either way.

3. I will suggest that the traditional desideratum of phonology, i.e. producing a single-grammar, unified analysis of a language's phonotactics and morphophonology, is losing traction because it is misguided and unlearnable. I will propose that the culprit is our desire for a single grammar. I will show how a principled approach to grammar proliferation (i.e. "sublexical phonology") greatly expands the range of learnable phenomena, which may in turn make the traditional goal of broad coverage viable again.

Bert Vaux (University of Cambridge)

"Evaluating OT: Abkhaz stress requires process-specific directional iterative intra-morphemic computation."

This paper explores one of the cells in the space of predictive differences delineated by the extensions of OT (McCarthy and Prince 1993) and RBP (Kenstowicz 1994). Globalist theories such as OT differ from localist theories such as RBP in their predictions concerning intra-morphemic operations; stress assignment in Abkhaz appears to fall into the relevant zone of predictive difference. Examination of this system, particularly with regard to the interaction of lexical accents and clash deletion, leads to the conclusion that the phonological component of the human language faculty crucially requires the ability to scan directionally and iteratively within morphemes. This functionality sits comfortably in highly localizable derivational theories such as RBP, but is fundamentally incompatible with globally oriented theories such as OT. Even recent versions of OT that acknowledge the need for extrinsically staged computation fail to generate the desired facts, and it is not clear how they might be repaired to address this problem without being reduced to notational variants of RBP.

Dybo 1977 first saw that the core facts of Abkhaz stress can be captured as in (1), on the assumption that vowels can bear lexical accents in underlying representations:

(1) Primary surface word stress falls on the leftmost lexically accented vowel that is followed by an unaccented vowel or word edge. If there is no lexical accent in the word, surface stress falls on the last vowel of the word.

Spruit, Kathman, and Trigo have shown that Dybo's insights can be derived straightforwardly in Rule-Based Phonology (RBP). In the Halle & Idsardi 1995 formalism, for example, we can derive the basic Abkhaz system via the operations in (2).

(2) i. Project a right bracket) for all lexical accents.
ii. Line 0 Edge Marking: LLL
iii. Clash Deletion:) → Ø / _ *) [iterative, L→R]
iv. Project rightmost element of Line 0 feet to Line 1
v. Project leftmost element of Line 1 feet to Line 2

The conflicting directionality identified by Dybo results from Left vs Right headedness on Lines 0 and 1 respectively (iv, v), and the iterativity and directionality via iii (cf. Howard 1972). On the other hand, no version of Optimality Theory currently on the market is able to account straightforwardly for a system of the Abkhaz type, due to the ability of the Clash Deletion rule to apply iteratively and directionally within individual morphemes.

In this talk I present the Abkhaz facts and their analysis in RBP and the leading variants of OT. I show that the Classic OT tenets of globalism/parallelism and minimal violation favor outputs which do the global minimum necessary to avoid stress clash, which harmonically bound the desired winners with their greater number of clash deletions. Versions of OT that allow staged computation sensitive to morphological structure (e.g. Stratal OT or Orgun's Cyclic OT) can deal with at least a subset of the cases where no more than one deletion happens per morpheme, but Stratal OT fails with forms involving more than one deletion per stratum and Cyclic OT fails with forms involving more than one deletion per morpheme. Harmonic Serialism (and perhaps OT-CC) have the power to generate the desired outputs, but cannot rule out equally harmonic outputs produced by derivations that do not apply clash deletion in L→R order.

The problem in brief is that the directionality of clash deletion has to be built into the process itself, which isn't allowed in existing forms of OT, as they do not recognize the existence of processes.

REFERENCES

Alderete, John. 1999. *Morphologically governed accent in Optimality Theory*. Ph.D. dissertation, University of Massachusetts, Amherst.

Dybo, V. 1977. 'Zapadnokavkazskaja akcentnaja sistema i problema ee proisxozhdenija.' In Konferencia Nostraticheskie jazyki i nostraticheskoe jazykoznanie, Tezisy doladov, 41-45.

Halle, Morris, and William Idsardi. 1995. 'General properties of stress and metrical structure.' In J. Goldsmith (ed.) The Handbook of Phonological Theory, 403-443. Cambridge, MA: Basil Blackwell.

Kathman, David. 1992. 'Stress and accent in Abkhaz.' In Germán Westphal et al. (eds.), ESCOL '91: Proceedings of the Eighth Eastern States Conference on Linguistics, 210-221.

Kiparsky, Paul. 2000. 'Opacity and cyclicity.' The Linguistic Review 17.

McCarthy, John. 2007. 'Restraint of analysis.' In Freedom of analysis, ed. M. Krämer S. Blaho, P. Bye, 203–31. Berlin and New York: Mouton de Gruyter.

Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology with special attention to Optimality Theory*. Ph.D. dissertation, University of California, Berkeley.

Revithiadou, Anthoula. 1999. *Headmost accent wins*. Ph.D. dissertation, Leiden University.

Riggle, Jason, and Colin Wilson. 2005. 'Local optionality.' In Proceedings of NELS 35. Amherst, MA: GLSA.

Spruit, Arie. 1985.. 'Stress in Abkhaz.' Studio Caucasica 6, 31-81.

Trigo, Loren. 1992. 'Abkhaz Stress Shift.' in George Hewitt (ed.), Caucasian Perspectives. Unterschleissheim/München: Lincom Europa.

Zoll, Cheryl. 1997. 'Conflicting directionality.' Phonology 14.2, 263-286. [ROA 151-10962]

Tobias Scheer (CNRS, Université de Nice)

"We need a theory of computation."

It will be recalled that OT is a theory of computation, not a theory of phonology. And that in order to make a theory of phonology you need a theory of computation and a theory of representations, which must not reducible one to another. OT has no, and has never had the ambition to have or develop genuine representations: it works with the stock inherited from the 80s, in the segmental area typically with SPE-type features alone. That is, given the natural inclination of OT towards computation, on the one hand no effort was put into the development of representations (we are still driving a car from the 80s), and on the other hand representations have been demoted to decoration. This follows from the fact that in an OT grammar (with unmarshalled GEN and Richness of the Base), the only locus of decision about (a)grammaticality is the constraint chamber, i.e. computation. There is no independent and unoutrankable arbitral award coming from representations. The OT mainstream was happy with that and actually advertised the final elimination of representations as a goal (introduction to de Lacy's 2007 *Handbook*). There is no principled reason, though, to go this way, and a situation where OT computation coexists with a theory of representations, each autonomous and not the slave of the other, is perfectly feasible (and there is a minority movement in OT going into this direction).

On the basis of this I'd like to suggest that one reason of the loss of velocity of OT is this: the misconception that phonology reduces, or may reduce, to computation. Coming from Government Phonology, I am aware that this theory may sometimes have been prone to the opposite temptation, i.e. to believe that a theory of representations will suffice to make a good theory of phonology. GP is kind of the symmetrical enterprise with respect to OT: it does not (really) have a theory of computation, and computation is often only decorative, existing at best in prose statements of the kind "and then X spreads to Y" or "and then the floating suffix-initial vowel associates to the final empty nucleus of the stem".

The computation-is-king direction that OT was misguided enough to engage in betrays Prince & Smolensky's original conception, which was the idea that constraint-based computation relate representations (i.e. allow us to go from one derivational stage to another). I suggest that the decline of OT is also due to the fact that there are a few more promises that OT did not bring home and/or betrayed: universality of the constraint set, anti-derivationalism, absence of evaluation of intermediate forms (surface-orientation), modularity, Freedom of Analysis, free ranking, constraint violability, dominance.

Looking ahead, I believe that there are two dangers: 1. OT will go the way it came: without any confrontation or argument exchange with its historical competitor, extrinsic rule ordering; 2. discussion will not be between theory X and theory Y as it has always been, but between people who believe that there ought to be a theory, and those who either believe that we don't need any, or who don't care because you can go to conferences, publish etc. just by measuring this and that and "see whether this concords with previous measurements or not".

What is needed, I submit, is a collective consciousness that we are part of a field, phonology, which like any other scientific endeavour needs a theory (or theories), and more specifically a theory of representations and a theory of computation. In the past the front lines were too much theory-oriented: somebody who works on representations (GPers for example) should be aware that they will also need a theory of computation, and get interested in what is out there with sympathy. And the symmetric attitude should be manifested by people who work on computation. Since Stephen Anderson's book on the history of phonology in the 20th century, phonologists should

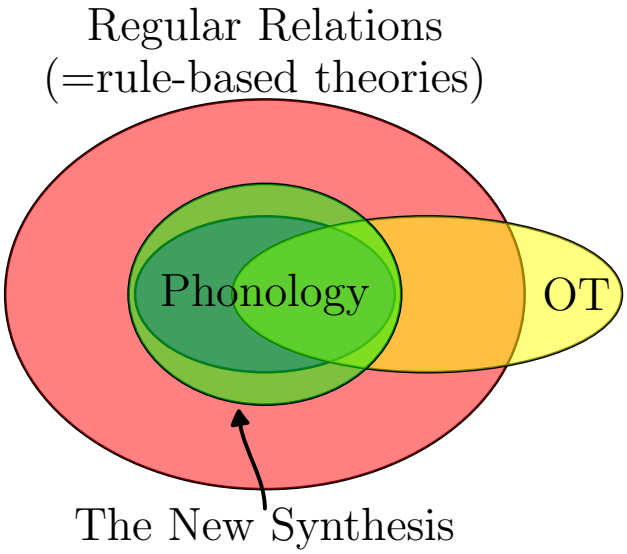
have understood that representational (80s) as much as computational (90s, 00s) imperialism is a dead end.

Jeffrey Heinz & Bill Idsardi (University of Delaware & University of Maryland)

"Beyond Optimality: the computational nature of phonological maps and constraints."

Optimization is a common metaphor in many fields (Alexander 1996, Kersten et al 2004, Pyle 1984, Schoemaker 1991). For the past twenty plus years, OT (Prince and Smolensky 2004) has been the dominant framework for understanding phonological patterns.

We have argued elsewhere (Heinz and Idsardi 2011, 2013) that phonological generalizations are restricted to natural, small, subregular classes of patterns and that the defining properties of those classes ought to be the cornerstone for a new synthesis of phonological theory. We argue this synthesis combines the best insights of rule-based and constraint-based theories (cf. Idsardi 2000). This provides a solution for the over-generation problem in both theories, and the under-generation problem in OT (see figure).



As in OT, phonology needs a theory of surface constraints and phonotactics. But we also argue that, as in rule-based theories, phonology needs a theory of underlying representations, which is incompatible with the principle of a rich base. The new synthesis includes both a theory of surface constraints and a theory of underlying representations. Furthermore, as with OT, the synthesis comes with learnability results (Heinz 2010, et seq., Chandlee et al. 2014), and, as with rule-based theories (and certain variants of classic OT), opaque patterns can be represented (Chandlee et al. 2015).

Of course much work still needs to be done. Phonology also needs a theory which will correctly correlate violations of constraints with repairs. Rule-based theories were too rigid because repairs and constraints were inseperable; on the other hand, in OT they are completely separate with the consequence that there are "too many solutions" in OT (Blumenfeld 2005).

Finally, we believe elements of the grammars are in a relationship which has a more complicated data structure than a list (that is, not just a list of rules or a ranking of constraints). The new synthesis, grounded in mathematical, logical, and computational

theories of representations and the transformations of those representations, allows for this possibility to be explored with unprecedented rigor.

REFERENCES

Alexander, R. McNeill. 1996. Optima for animals. Revised edition. Princeton University Press.

Blumenfeld, Lev. 2005. Too many solutions: prosody and segmental effect. Proceedings from the Annual Meeting of the Chicago Linguistics Society. 41:17-31.

Chandlee, Jane, Jeffrey Heinz, and and Adam Jardine. 2015 Representing and Learning Opaque Maps. Workshop on Computational Phonology at GLOW 38.

Chandlee, Jane, Remi Eyraud and Jeffrey Heinz. 2014. Learning Strictly Local Subsequential Functions. Transactions of the Association for Computational Linguistics, v. 2, pp.491-503.

Heinz, Jeffrey. 2010. String Extension Learning. Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. pp. 897--906.

Heinz, Jeffrey and William Idsardi. 2011. Sentence and word complexity. Science. 333:295-7. Erratum in: Science. 333:1825.

Heinz, Jeffrey and William Idsardi. 2013. What complexity differences reveal about domains in language. Topics in Cognitive Science. 5:111-31.

Idsardi, William. 2000. Clarifying Opacity. The Linguistic Review. 17:337-50.

Kersten, Daniel, Pascal Mamassian and Alan Yuille. 2004. Object Perception as Bayesian Inference. Annual Review of Psychology. 55:271-304.

Prince, Alan and Paul Smolensky. 2004. Optimality Theory: Constraint Interaction in Generative Grammar. Wiley-Blackwell.

Pyle, Graham H. 1984. Optimal Foraging Theory: A Critical Review. Annual Review of Ecology and Systematics. 15:523-575.

Schoemaker, Paul J. H. 1991. The quest for optimality: a positive heuristic for science? Behavioral and Brain Sciences 14:205-215.

We aim for there to be time and opportunity at the end of the session for general discussion, so that all members of the audience have the opportunity to contribute.

[The following abstract was submitted for the workshop and accepted for presentation, but the authors had to withdraw – we retain it here for reference.]

Paul Smolensky, Matt Goldrick, Geraldine Legendre, Colin Wilson, Jennifer Culbertson, Akira Omaki, Kyle Rawlins, & Benjamin Van Durme

"Gradient Symbolic Computation: situating grammar in the mind/brain."

Since the advent of generative phonology (e.g., Chomsky and Halle, 1968), the phonological grammar has been viewed as part of the mind/brain of individual speaker-hearers. As such, developing a complete grammatical theory requires a clear specification of how grammar relates to the overall cognitive architecture – the functional, algorithmic, and ultimately neural specifications of the cognitive system. Optimality Theory (OT, Prince & Smolensky 1993/2004) and Harmonic Grammar (HG, Legendre, Miyata & Smolensky 1990) can be imbedded as part of the Integrated Connectionist/Symbolic (ICS) computational architecture of cognition (Smolensky and Legendre, 2006). In this architecture, continuous-activation processing defines the micro-structure of cognition – in the form of neural computation – and discrete processing defines the macro-structure of the very same system – in the form of symbolic computation. In this work, activation and structure collaborate to characterize cognition, but each is primarily relegated to its own level of description. The macro-structure (characterized by OT and HG) consists of symbolic functions defined over discrete combinatorial representations; the microstructure consists of continuous activation-spreading through networks to compute those functions.

More recently, in linguistic domains ranging from phonetics to semantics, we have been examining a perspective on ICS that takes a radical step beyond OT and HG: Gradient Symbolic Computation (GSC, Smolensky, Goldrick & Mathis, 2014). At the level of cognitive macro-structure, GSC incorporates computational and representational principles from the micro-structure of neural-network processing. This yields the novel claim that grammatical computations operate over Gradient Symbol Structures comprising blends of symbols – blends of structural fillers (e.g., both /d/ and /t/ occupy the coda position, but to varying degrees) as well as blends of structural roles (e.g., to varying degrees, /t/ occupies the coda of one syllable and the onset of another). GSC, in emphasizing the role of grammar within the general cognitive system, has the potential to provide new insights into behavioral and neural data relating to performance with linguistic structure.

In the realm of competence, we are exploring how GSC may yield new insights into grammatical phenomena via the hypothesis that grammatical patterns can often not be comprehensively explained within a single structural analysis. Linguistics is rife with cases in which putatively inconsistent structural analyses of a single phenomenon persist indefinitely as each analysis can explain aspects of the phonemenon that other analyses cannot. The GSC hypothesis is that in such cases a complete explanation demands a gradient blend of mutually inconsistent structural accounts. We illustrate with such an analysis of a classic, challenging set of phonological alternations: latent consonants and vowels in French (e.g., Tranel, 1981).