Violable constraints in Classical Universal Phonology and beyond*

Joe Pater, UMass Amherst

Whither†OT mfm fringe workshop,
Manchester May 27th, 2015

Abstract

It appears to me that 'Classical Universal Phonology' (CUP) as a whole, rather than OT[1] in particular, is receiving less attention from phonologists in the 21st century. I highlight some of the contributions of violable constraints to CUP, and provide an overview of some developments in generative phonology outside of CUP *per se*, again emphasizing the role of violable constraints, especially in formalizing grammatical learning. I conclude with some speculations about why CUP seems to be less popular these days, and about what we might expect for the future.

1 Classical Universal Phonology

Goal: A theory of Universal Grammar that generates all and only possible phonological systems

Principles and parameters approaches of the 1980's made much progress towards that goal by developing theories of autosegmental and metrical representation that worked with parameterized rule + constraint systems.

Rule-only CUP is rare (creating issues for comparison of rule-based phonology *vs.* OT). Generative phonology prior to the 1980's did not pursue the 'all and only' goal to the same degree.

*Thanks to the Edinburgh P-Workshop for their feedback on a pilot version of this talk, in particular Jenny Culbertson, Patrick Honeybone, Joe Fruehwald, Pavel Iosad, and James Kirby, and to Alex Cristía and Presley Pizzo for a serendipitous lunch conversation in Paris. Its preparation was supported by NSF grant BCS-424077 to the University of Massachusetts Amherst, and by the city of Paris under a Research in Paris fellowship.

†Please note the absence of parentheses around the 'h' in my rendering of the name of the workshop, which is an intentional deviation from the advertised version.

[1]Optimality Theory, Prince and Smolensky (1993/2004). I haven't done proper citations for this handout because I didn't want to make the difficult decisions about what needed to be cited on what, and because the bibliography-free handout fit perfectly on 4 A4 pages. Please ask if you can't find something I mention or if you're curious about further literature on some topic. While the term Classical Universal Phonology may be my invention (UP or Youpie is pre-existing), I assume that the seminal contributions to the general approach to phonology that I'm describing with it are well-known.

1.1 Violability and CUP 1: Theories of typology

Pre-OT constraints were inviolable: an active constraint cannot be violated within its domain or level of application.

In OT, a constraint can be violated *iff* its satisfaction would require the violation of a higher ranked constraint.

– Violability often allows for more general constraints in the analysis of individual languages.
– With more general constraints, more explicit and testable theories of typology become achievable.[2]

1.2 Violability and CUP 2: Formalization of effects of constraints

See further McCarthy, Pater and Pruitt (to appear)

Constraints first appeared in the early 1970s in the work of Kisseberth, Sommerstein and others, who immediately recognized the difficulty of formalizing how constraints block and trigger rules. These difficulties are quite often glossed over in subsequent constraints + rules work.

Perhaps the hardest type of case: an ill-formed structure is created then repaired. How does the ill-formed structure get built in the first place?

e.g. Trochaic shortening or iambic lengthening – how do we build the ill-formed foot before fixing it?

Standard OT's answer: parallelism and violability

Harmonic Serialism's answer: just violability (more on HS in Kimper's talk)

1.3 Violability and CUP 3: Learning

We aren't solving the learning problem by giving the learner parameters - we also have to specify how they are set. In phonology, see Dresher (and Kaye) on setting of metrical parameters.

– The biggest problem for parameter setting is interdependence.
– This problem simply doesn't exist for violable constraints (see *esp.* Tesar and Smolensky 2000).

[2]Terminology here: OT is a framework, OT with a specified constraint set gives a theory of typology for some domain. Potential point of discussion given Heinz and Idsardi's work on formal language theory: are we seeking theories, or frameworks, with the right degree of generative power? Probably both, but the distinction seems important. A comment / question given Scheer's work on Government Phonology and Vaux's defense of rules: I suspect that part of the reason OT has been popular is that it helps us strike a good balance in between overly restrictive theories (e.g. those in GP?) and overly permissive ones (e.g. those with unconstrained rules?). A couple general issues: 1. Unfortunately theories are rarely, if ever, in a subset relation in terms of what they generate, so restrictiveness is often in the eye of the beholder. 2. Theorists also have different tastes or strategies that affect their choices on the restrictiveness – coverage continuum.

– This is not to say learning is 'easy' with violable constraints – hidden structure learning (e.g. Jarosz 2013: Phonology) and constraint induction (e.g. Hayes and Wilson 2008) are active areas of research – but learning research has clearly flowered with them.

When constraints are weighted (i.e. when we are using Harmonic Grammar; see Pater 2009, to appear on restrictiveness), there is a particularly wide range of learning results in cognitive psychology, natural language processing, and machine learning on which we can draw. My examples of the usefulness of violable constraints 'beyond CUP' are all instances of the usefulness of weighted constraints (*esp.* MaxEnt models) with their associated learning algorithms, though at least some of the results may well replicate with other grammatical theories and learning algorithms.

2 Beyond CUP

Broader goals of generative phonology:

– To provide a formal characterization of knowledge of phonology
– To explain how that knowledge is acquired
– To explain typological generalizations over phonological systems

Here are some examples of how these goals have been pursued in contemporary generative research, and how learning with violable constraints has played a role.

2.1 Laboratory learning experiments

In cognitive terms, Universal Grammar may best be thought of a theory of inductive bias: what biases are a language learner endowed with that explain the success of learning, the learning trajectory, and also the nature of the generalizations that the learner makes from the data?

A particularly direct way of studying inductive bias is to study learning in the lab, in which the input data are tightly controlled. This research has taken off in the 21st century; see Moreton and Pater's Language and Linguistic Compass review.

MaxEnt modeling has been prominent in this domain:

– Wilson (2006: Cognitive Science) shows how biases for phonetically natural processes can be formulated in terms of the prior of a MaxEnt model.
– Moreton, Pater and Pertsova (to appear: Cognitive Science) show how biases for featurally 'simple' patterns emerge from a MaxEnt model without a stipulated prior, and find support for fine-grained predictions of the model in a learning experiment.

2.2 Experiments on native speakers + learning simulations

Another popular stream of research in this century, starting with Ernestus and Baayen (2003: Language), Adam Albright and Kie Zuraw's dissertations, and Bruce Hayes' collaborations with them and others. (It's indebted to prior connectionist research, especially on the English past tense):

– Test native speakers to assess their knowledge of a phonological pattern, and test whether a learning algorithm arrives at a final state that matches that knowledge

This type of research is now widely pursued, and much of it involves MaxEnt modeling (more on MaxEnt in Becker's presentation):

– Hayes and Wilson (2008) train a MaxEnt phonotactic grammar on a dictionary corpus of English onsets, and show a good correlation between well-formedness judgments on nonce words and the MaxEnt scores. Daland *et al.* (Phonology: 2011) do further experimental work, and show that a MaxEnt model does better than several other models of phonotactics.
– Hayes, Zuraw, Siptár and Londe (2009) show how native speaker choices of suffixes for nonce roots match trends in the lexicon, and provide a MaxEnt learning account.

2.3 Typology with learning

What role does learning itself play in shaping typology?

Some recent work has addressed these questions using agent-based models in which the internal structure of the agents consists of violable constraint grammars and learning algorithms (MaxEnt + SGA):

– Staubs (2014: UMass diss.) shows that we can derive predictions about the relative frequency of stress systems using standard assumptions about grammar and learning (unattested vs. attested as the goal for typological modeling seems as arbitrary as ungrammatical vs. grammatical as the goal for human modeling).
– He also shows how phonetic pressures against final stress can shape the typological distribution of stress systems.
– Learning and agent interaction can also shape typology in ways relevant to theoretical debates: Staubs (2014), Hughto (2014: UMass ms.) and Hughto, Pater and Staubs (GLOW 2015) show that the output of these models can skew away from gang effects.

3 The future?

Why are phonologists doing less CUP?[3]

– There is almost certainly a mismatch between CUP's UG, and the UG that one might posit for a human learner.
– There are also CUP-internal foundational issues (*e.g.* What do we do about 'crazy rules'? How much phonetics belongs in phonology? How good are the data?)

I expect that there will be a continued integration of learning with phonological theory, with violable, especially weighted, constraints continuing to play a central role. I also expect that insights gained from the study and modeling of language change will help us better understand phonological typology.

In all likelihood, OT and rule-based phonology, combined with autosegmental and prosodic representations, will continue to be the lingua francas of phonologists for some time.

[3]I hasten to note that I believe that it's important to continue to pursue CUP, even if its UG is an *idealization* (see further Pater 2009: Cognitive Science: sec. 4.2).