Harmonic Grammar, Gradual Learning, and Phonological Gradience

Joe Pater, University of Massachusetts, Amherst

June 29, 2007

Harmonic Grammar (HG; Smolensky and Legendre 2006) is a close relative of Optimality Theory (OT; Prince and Smolensky 1993/2004). In HG, the optimal candidate is the one with maximal Harmony, where Harmony is calculated using a simple linear equation. Given a representation's scores on a set of constraints, and a set of coefficients, or weights, Harmony is the sum of the weighted constraint scores. If, as in OT, constraints assign scores based on the number of violations, then the scores are the corresponding negative integers, and the weights are positive reals. An OT-style tableau using Harmony maximization as the criterion for optimality appears in (1). The weights are in the top row, and the rightmost column provides the Harmony values for the candidates. Output_1 has the highest Harmony, and is thus optimal.

(1) A weighted constraint tableau

	Weight	2	1	$\mathcal{H}$
	Input	Constraint 1	Constraint 2	
☞	Output_1		-1	-1
	Output_2	-1		-2

OT's use of ranking rather than weighting is sometimes justified in terms of its restrictiveness: weighted interaction is claimed to be too powerful for HG to function as a realistic model of human language (Prince and Smolensky 1993/2004: 236; 1997: 1608, Legendre et al. 2006b).

In this talk, I present results of several ongoing collaborative research projects on HG. I will discuss the following points:

(2) i. HG is (perhaps surprisingly) restrictive, due to inherent limitations on the types of languages that can be generated by an optimization system (Bhatt et al. 2007; Pater et al. 2007)

ii. HG is compatible with a simple correctly convergent gradual learning algorithm, the Perceptron algorithm of Rosenblatt (1958) (Boersma and Pater 2007; Pater 2007; see Jäger 2006, Soderstrom et al. 2006 for precedents).

iii. To deal with variation, HG can be implemented with noise, as in stochastic OT (Boersma 1998; Boersma and Hayes 2001). Testing shows that the noisy HG+Perceptron combination is robust, unlike the stochastic OT+GLA one (Boersma and Pater 2007).

iv. Gradual learning yields Harmony values that reflect frequency distributions. A problem for the HG account of gradient well-formedness (Keller 2006; Legendre et al. 2006a) raised by Boersma (2004) can be resolved with a revised HG acceptability metric (Coetzee and Pater 2007).

References

Bhatt, Rajesh, Joe Pater, and Christopher Potts. 2007. Harmonic Grammar with Linear Programming. Ms., UMass Amherst.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Boersma, Paul. 2004. A stochastic OT account of paralinguistic tasks such as grammaticality and prototypicality judgments Ms, University of Amsterdam.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86. Available on Rutgers Optimality Archive, http://ruccs.rutgers.edu/roa.html.

Boersma, Paul, and Joe Pater. 2007. Testing gradual learning agorithms. Ms, University of Amsterdam and UMass Amherst.

Coetzee, Andries, and Joe Pater. 2007. Weighted constraints and gradient phonotactics in Muna and Arabic. Ms, University of Michigan and UMass Amherst.

Jäger, Gerhard. 2006. Maximum entropy models and Stochastic Optimality Theory. In *Architectures, rules, and preferences: A festschrift for Joan Bresnan*, ed. Jane Grimshaw, Joan Maling, Chris Manning, Jane Simpson, and Annie Zaenen. Stanford, CA: CSLI. To appear.

Keller, Frank. 2006. Linear optimality theory as a model of gradience in grammar. In *Gradience in grammar: Generative perspectives*, ed. Gisbert Fanselow, Caroline Féry, Ralph Vogel, and Matthias Schlesewsky. Oxford: Oxford University Press.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. 2006a. The interaction of syntax and semantics: A Harmonic Grammar account of split intransitivity. In Smolensky and Legendre (2006), 417–452.

Legendre, Géraldine, Antonella Sorace, and Paul Smolensky. 2006b. The Optimality Theory–Harmonic Grammar connection. In Smolensky and Legendre (2006), 339–402.

Pater, Joe. 2007. Gradual learning and convergence. *Linguistic Inquiry* To appear.

Pater, Joe, Rajesh Bhatt, and Christopher Potts. 2007. Linguistic optimization. Ms, University of Massachusetts, Amherst.

Prince, Alan, and Paul Smolensky. 1993/2004. Optimality Theory: Constraint interaction in generative grammar. RuCCS Technical Report 2, Rutgers University, Piscateway, NJ: Rutgers University Center for Cognitive Science. Revised version published 2004 by Blackwell. Page references to the 2004 version.

Rosenblatt, Frank. 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65:386–408.

Smolensky, Paul, and Géraldine Legendre. 2006. *The harmonic mind: From neural computation to Optimality-Theoretic grammar*. Cambridge, MA: MIT Press.

Soderstrom, Melanie, Donald W. Mathis, and Paul Smolensky. 2006. Abstract genomic encoding of universal grammar in optimality theory. In *The harmonic mind: From neural computation to optimality-theoretic grammar. volume 2: Linguistic and philosophical implications.*, ed. Paul Smolensky and GÈraldine Legendre, 403–471. MIT Press.