Harmonic Grammar growing into Optimality Theory:

Maturation as the strict domination limit (or vice-versa)

Tamás Biró

Eötvös Loránd University, Budapest; tamas.biro@btk.elte.hu

In the last decade, Harmonic Grammar (HG), the predecessor of Optimality Theory (OT), re-emerged as an alternative to OT. Their typological predictions and learnability have been thoroughly compared. Some argue for the advantages of Harmonic Grammar, while other scholars are more than happy with the capacities of Optimality Theory. How do they relate to each other?

In 1990, before the advent of OT, Harmonic Grammar was presented by Legendre, Miyata and Smolensky in a series of papers as a connectionist model that maximizes "harmony" (or minimizes "energy"). *The Harmonic Mind* by Smolensky and Legendre contains the full-fledged theory of how a "higher level" symbolic rule system and a "lower level" neural network can approximate each other (depending on one's perspective). A candidate is an activation pattern $\{a_i\}_{i=1}^{N}$ of the N nodes in the network, whereas constraint C_k is a set $\left\{W_{i,j}^{k}\right\}_{i,j=1}^{N}$ of partial connection weights. The connection $W_{i,j}$ between the network nodes i and j is obtained as a linear combination over the n constraints $C_1, C_2, \dots C_n$: namely, $W_{i,j} = \sum_{k=1}^{n} w_k \cdot W_{i,j}^{k}$, where w_k is the weight of constraint C_k. Hence, the harmony of a candidate A – that is, of an activation pattern $\{a_i\}_{i=1}^{N}$ – can be expressed as:

$$H(A) = \sum_{i,j=1}^{N} a_i \cdot W_{i,j} \cdot a_j = \sum_{i,j=1}^{N} a_i \cdot \sum_{k=1}^{n} w_k \cdot W_{i,j}^{k} \cdot a_j = \sum_{k=1}^{n} w_k \cdot \sum_{i,j=1}^{N} a_i \cdot W_{i,j}^{k} \cdot a_j = \sum_{k=1}^{n} w_k \cdot C_k[A]\,.$$

The result is that a constraint can be viewed as a function of the candidate, and the total harmony assigned to that candidate is obtained by a linear combination of the constraints. Thus emerges the symbolic version of HG, increasingly popular nowadays, even among non-connectionist linguists.

The key difference between OT and HG is the *strict domination* nature of the constraint interaction. A strict domination hierarchy $C_k \gg C_{k-1} \gg \cdots \gg C_2 \gg C_1$ can be realized with weights $w_k = q^k$, provided that no constraint ever assigns more violations than $q-1$ (that is, $C_k[A] \leq q-1$ for all constraints C_k and all candidates A). Yet, what guarantees an exponentially growing weight system? The Harmonic Mind (e.g., vol. 1, p. 87) lists "the emergence of OT's strict domination constraint interaction (…) from network-level principles" as one of the major open problems in their *Integrated Connectionist/Symbolic Cognitive Architecture*.

In order to get closer to the solution of this problem, the current paper presents some observations regarding the role of the base q of an exponential weight system. Growing q infinite shall be called the *strict domination limit.* It has been noted that in a symbolic simulated annealing implementation of OT or HG, increasing q speeds up calculating the optimal form (Biró 2009). In the strict domination limit, OT is more time-efficient than HG. At the same time, even if computation is performed very slowly, OT is prone to make errors, while simulated annealing with the real-valued target function in HG is proved to be able to find the optimal candidate. Consequently, speed and error rate compete. It will be argued that phonology and morphology prefer OT because speed is more important on these lower levels, where numerous computations need to be performed in each second. At the same time, phrase and sentence level computations in syntax and semantics may be offered more time but less opportunity to errors; hence, their inclination for HG.

Finally, we shall also propose that the ideal base q of the exponential weight system might develop during ontogeny. Vanishing consonant cluster reduction in child language shall illustrate how HG grows into an OT phonological grammar, probably in order to accelerate mental computation. Pronoun resolution by children, however, will be explained as their gradually decreasing the value of q.