Gradient Symbolic Representations in Grammar: The case of French Liaison

Paul Smolensky
(Microsoft Research AI &
Johns Hopkins University)

Matthew Goldrick
(Northwestern University)

At the foundation of the work reported here is the following hypothesis: theoretical debates about whether structure A or structure B is the correct analysis of phenomenon X can persist indefinitely because, in fact, the mental representation supporting X is a conjunctive blend (not a disjunctive probabilistic mixture) of structures A and B. The notion 'blend of structures' is formalized using Gradient Symbolic Representations, symbol structures in which each individual position is generally occupied by a sum of gradient symbols, each symbol having a partial degree of presence: its activity. The grammatical consequences of a Gradient Symbolic Representation are the sum of the consequences of all the symbols blended to form it; the consequences of a symbol - e.g., the costs of constraint violations - are proportional to its activity. Gradient Symbolic Representations enable formal expression of important gradient theoretical intuitions.

Our test case is a well-studied phenomenon of French phonology, liaison consonants, which alternate with zero. Theoretical intuitions expressed in the extensive existing literature concerning the structures underlying liaison - intuitions previously assumed to be in conflict - are blended to give a single analysis that covers a wide range of data not previously explicable within a single theory. The analysis formalizes the following intuitions about a liaison consonant $\mathscr{L}$ which may appear at the juncture of two consecutive words W_1 W_2: $\mathscr{L}$ is simply a consonant that is weak; different $\mathscr{L}$s are not all equally weak; $\mathscr{L}$ is underlyingly final in W_1; $\mathscr{L}$ is underlyingly initial in W_2; the greater the cohesion between W_1 and W_2 - i.e., the smaller the minimal prosodic unit containing both W_1 and W_2 - the greater the likelihood that $\mathscr{L}$ will appear. The proposed grammatical computation consists of optimization with respect to a numerical weighting of familiar constraints from Optimality-Theoretic phonology, straightforwardly extended to evaluate Gradient Symbolic Representations.

Gradient Symbolic Representations constitute the data of Gradient Symbolic Computation (GSC), a general computational architecture for cognition developed over the past several decades. The microstructure of a GSC system is a stochastic neural network deploying continuous, distributed representations; however this paper addresses only the macrostructure of the proposed GSC analysis, which is a Probabilistic Harmonic Grammar defined over Gradient Symbolic Representations. Although the network-level description is not utilized in this paper, the proposed analysis operates over gradiently active symbols which are ultimately emergent from patterns of gradiently active model neurons.