Modeling individual differences in a pragmatic reference game as a consequence of variable disengagement from unsuccessful strategies

John Duff[1], Alexandra Mayn[1], Vera Demberg[1,2]
[1] Department of Language Science and Technology, Saarland University
[2] Department of Computer Science, Saarland University

Individuals are known to vary in their likelihood of drawing Gricean pragmatic inferences. We present the first algorithmic-level model of a pragmatic reasoning task, in ACT-R, to formalize the way in which this variability could come from individual differences in domain-general cognitive variables.

In non-linguistic reference games (RefGames), participants select a referent for a possibly-ambiguous message. RefGame participants vary in their tendency to derive "simple" and "complex" Gricean inferences (Franke & Degen, 2016).

Mayn & Demberg (2022) found that more complex RefGame inferencing was strongly associated with successful problem-solving in Raven's Progressive Matrices (RPM). One shared mechanism which may underlie successful RPM and RefGame performance is the ability to use internal negative feedback to efficiently disengage from unsuccessful strategies: Stocco and colleagues' (2021) ACT-R model captures an observed relationship between one's negative feedback strength (F_{neg}) and RPM success.

Based on the Stocco model, we present an ACT-R model of RefGame where success requires disengaging from strategies which fail to provide a high-quality (unique) answer. Our model stochastically applies interpretive strategies of varying complexity until arriving at a unique answer or else guesses after a timeout corresponding to individual persistence. Deriving complex inferences ultimately requires faster disengagement via stronger F_{neg} (or higher persistence).

Our model generates several novel predictions: (a) RefGame performance should correlate with specific F_{neg} and persistence metrics, (b) more complex trials will take longer to solve, (c) incorrect responses will have shorter response times and (d) correct responses will more often demonstrate eye movements predicted for complex strategy execution.