

Parameter space of the DC language game

Bas Cornelissen

August 9, 2017

The Bayesian language game has a language strategy parameter η , a production strategy parameter ζ and a parameter γ for the life expectancy. How do those influence the resulting behaviour of the model? This appendix reports an experiment that systematically analysed the behaviour in a larger part of the space.

The experiment measures three new quantities besides coherence and reflectance. The first concerns the amount of *synonymy* in the language. If a language assigns all words the same probability, the synonymy is maximal, but if one word takes all probability mass, there is no synonymy. Synonymy is the inverse notion of efficiency in the naming games and formally defined as the relative Shannon entropy of the aggregate language,

$$S(t) := \frac{H(\bar{\pi}_t)}{\log_2(N)}, \quad (1)$$

where H is the entropy. $S(t) = 1$ indicates maximal synonymy, $S(t) = 0$ the complete absence of synonymy. Second, the *discrepancy* between the internal and external language is measured:

$$D(t) := \text{JSD}(\bar{\pi}_t, \psi_t), \quad (2)$$

where the aggregate language functioned as a proxy of all internal languages. Third, we measure the *variability* of the aggregate language as its standard deviation over time,

$$V(t) = \text{std}(\bar{\pi}_0, \dots, \bar{\pi}_t). \quad (3)$$

If the languages used in the populations were relatively stable throughout the game, the variability should be low.

The experiment simulated 20 runs of the Dirichlet-categorical language game for every combination of the parameters $\eta \in \{1, 2, 5, 50, \infty\}$, $\zeta \in \{1, 1.5, 2, 5, \infty\}$ and $\gamma \in \{1, 10, 50, 100, 1000, \infty\}$.¹ Every run had a duration of 100 000 iterations and all used the same, relatively flat, but nonuniform, prior. Trial experiments suggested

¹ To be completely clear: that amounts to 350 million rounds in 3500 independent simulations, using 175 different parameter settings.

that these parameter values sufficiently illustrate games in different parts of the parameter space. For example, $\eta > 100$ and $\zeta > 100$ already yield behaviour comparable to the infinite case and were left out. The coherence,² reflectance, synonymy, discrepancy, and variability were measured at the end of every game. Figure 1 shows all the results. Interpreting the results is tricky, and it helps to keep figure ?? in mind: The corners of every heat map in figure 1 mark the extreme strategies ($\eta, \zeta \in \{1, \infty\}$), which were also shown in figure ?. The life expectancies we considered earlier ($\gamma = 1, 10, 100$ and ∞) can also be found in figure ?. Now, the main conclusion is be that *our earlier findings are largely confirmed by the current experiment*. For intermediate parameter values, we observe a relatively smooth transition between the extreme cases $\eta, \zeta, \gamma \in \{1, \infty\}$.

We first discuss the last row, which corresponds to a naming game. The reflectance (column A) is much higher for samplers, and decreases quickly as soon as agents start to maximise their productions slightly (i.e. $\zeta > 1$). The synonymy (column B) suggests why: maximising production strategies result in a one-word language, that is, one with no synonymy. This explains why the reflectance is low for high ζ : the bias allows much more synonymy. The reflectance and synonymy also suggest that the language strategy (η) is far less influential than the production strategy (ζ). The fact that the change in vertical direction is smaller than the change in horizontal direction, and this seems to generalise to other life-spans as well. Looking at agents with shorter lifespans, we further see that the reflectance and synonymy increase as γ approaches 1 (iterated learning). As before, the reason is that the internal language of an agents with a short life span is nearly completely determined by the bias. Accordingly, reflectance is high and since the bias is fairly flat, so is synonymy. But, note that the discrepancy between the external language and the internal increases sharply for more maximising strategies. Finally the last column shows an increase in variability with intermediate lifespans. This could indicate continuous language change when agents have an intermediate lifespan of around $\gamma = 1000$ interactions, but more research is needed before such conclusions are warranted.

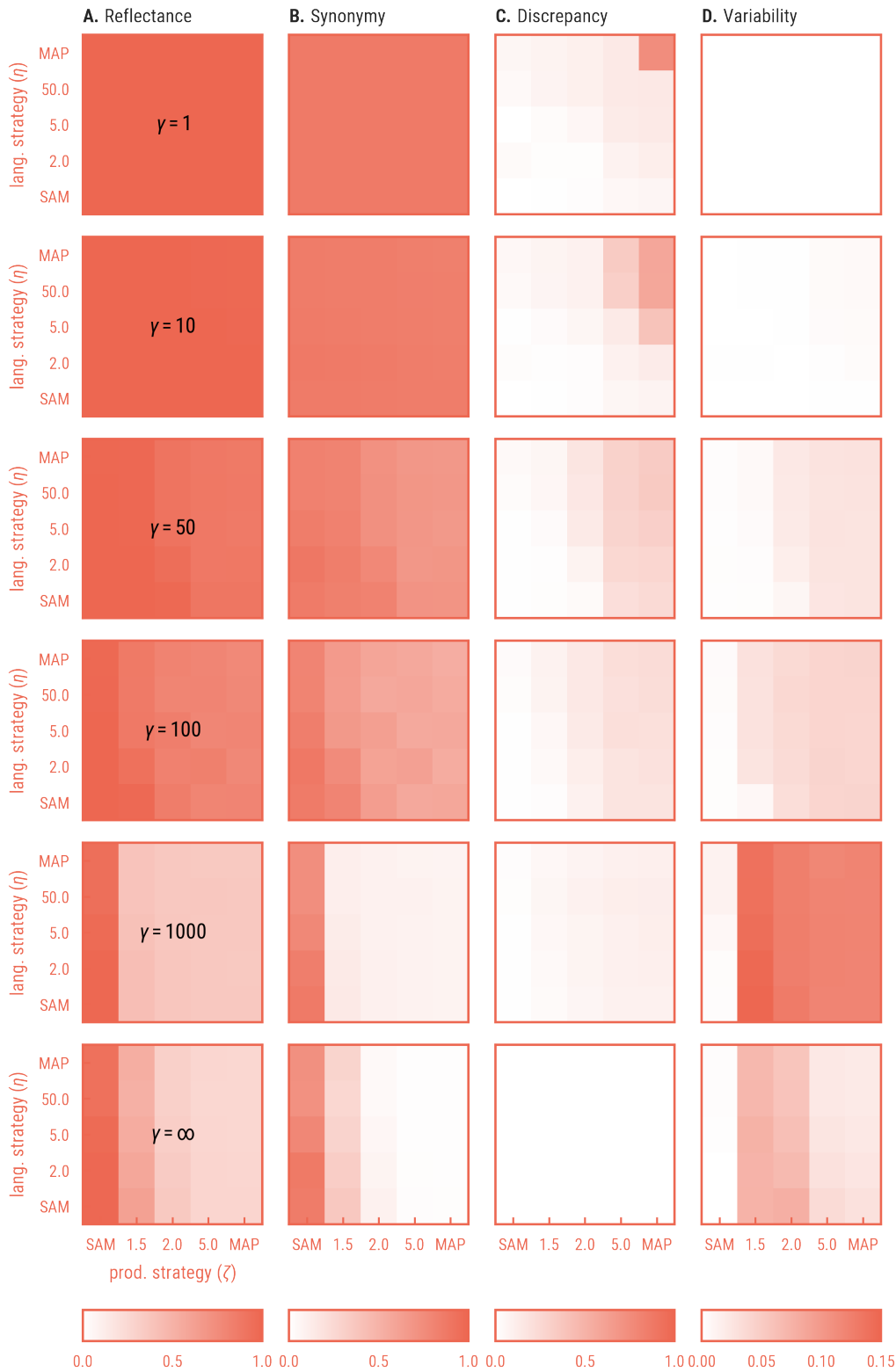


FIGURE 1 The behaviour of the Dirichlet-categorical language game across the parameter space (γ, η, ζ) . Rows corresponds to life expectancies (γ); columns show the coherence, reflectance, synonymy and variability for every strategy (η, ζ) . See figure ?? for the typical resulting languages in the extreme cases $\gamma, \eta, \zeta \in \{1, \infty\}$.

BNG06 Every cell is an average over 20 simulation runs. $K = 20$, $N = 10$, $b = 1$, $\gamma = \infty$, $\beta = 30$. Simulations used a deterministic hazard function.