EMERGENCE OF PRAGMATIC REASONING FROM LEAST-EFFORT OPTIMIZATION

NOGA ZASLAVSKY[*1], JENNIFER HU[1], and ROGER LEVY[1]

[*]Corresponding Author: nogazs@mit.edu
[1]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, USA

It is widely acknowledged that pragmatics is an important driver of language evolution (e.g., Sperber & Origgi, 2010; Scott-Phillips, 2015; Woensdregt & Smith, 2018) and language change (Jucker, 1995; Traugott, 2012; inter alia). A sound theoretical understanding of the dynamics of pragmatic reasoning and computational principles that may give rise to this dynamics is thus crucial to the field of language evolution. Here, we study this dynamics within the Rational Speech Act framework (RSA: Frank & Goodman, 2012; Goodman & Frank, 2016).

RSA formulates pragmatic reasoning as probabilistic speakers and listeners recursively reasoning about each other. The speaker is defined by a production distribution $S(u|m)$ over possible utterances u given meaning m, and the listener is defined by an inference distribution $L(m|u)$. RSA recursively relates the speaker and listener by assuming a Bayesian listener—$L(m|u) \propto S(u|m)P(m)$, with $P(m)$ a prior distribution on speaker meanings that is assumed to be in common ground—and a speaker that is bounded-rational with respect to a utility function $V(u, m)$ (typically, $V(m, u) = \log L(m|u) - C(u)$ where $C(u)$ specifies the cost of u). That is, $S(u|m) \propto \exp(\alpha V(u, m))$, where α controls the degree to which the speaker maximizes utility. The framework enjoys broad popularity and empirical support (for review: Goodman & Frank, 2016), and while shallow recursion is often assumed, several studies have also explored and motivated deeper recursions (e.g., Camerer, Ho, & Chong, 2004; Franke & Degen, 2016; Bergen, Levy, & Goodman, 2016; Levy, 2018). These explorations have relied on numeric simulation (e.g., Yuan, Monroe, Bai, & Kushman, 2018; Peloquin, Goodman, & Frank, 2019), leaving much unknown regarding the dynamics of RSA recursion.

Here we present new analytic results, illustrated by implemented model instances, that answer key open questions about RSA dynamics. Because the RSA speaker is guided by (soft) optimization of utterance utility, the intuition is widely held that RSA recursion is guaranteed to (locally) optimize expected utility (e.g., Yuan et al., 2018). Our analysis disconfirms this intuition. We show that the RSA recursion is an instance of the alternating maximization algorithm (Csiszár & Shields, 2004), providing an optimization guarantee. However, the guarantee is

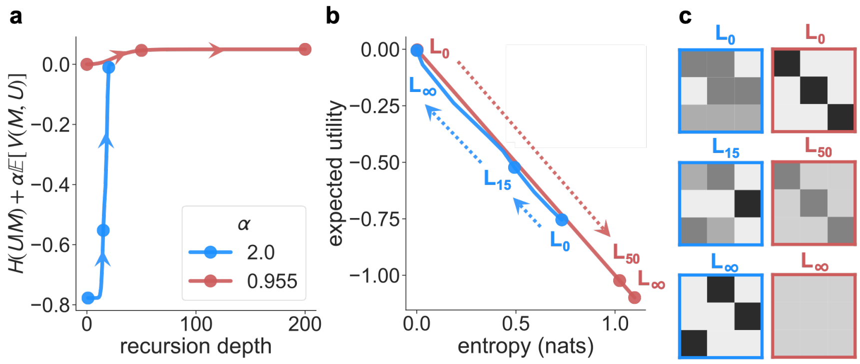

Figure 1. Model simulations with three uniformly distributed meanings and three possible utterances. **(a)** Trade-off between utility and effort improves with the depth of recursion. **(b)** Expected utility as a function of the speaker's entropy, $H(U|M)$. Expected utility may increase (blue; $\alpha = 2$) or decrease (red; $\alpha \approx 0.955$) as recursion depth increases. **(c)** Listener distributions at initial, intermediate, and converged conditions. Darker grays correspond to higher probabilities. (For L_0 in red, the off-diagonal elements are initialized to small but non-zero ϵ values.)

not to improve expected utility but rather a *tradeoff* between communicative effort and expected utility, namely $H(U|M) + \alpha\,\mathbb{E}[V(M,U)]$, where $H(U|M)$ is the conditional entropy of utterances given speaker meanings. This tradeoff can be thought of as an instance of Zipf's least-effort principle (Zipf, 1949), where here low communicative effort corresponds to high entropy of the speaker's production distribution. Our analysis also reveals that in general α does not simply trade off against recursion depth, as widely understood (e.g., Frank, Emilsson, Peloquin, Goodman, & Potts, 2018): the value of α determines the tradeoff between effort and communicative utility optimized by RSA recursion.

The model simulations of Figure 1 exemplify these results. RSA iteration always improves the utility–effort tradeoff (Figure 1a), but expected utility may increase (Figure 1b, blue trajectory), or decrease (red trajectory), depending on α and the initial listener (Figure 1c, L_0). We speculate that the possibility of RSA iteration decreasing expected utility has not previously been identified in numeric simulations because RSA initializations are typically (apart from structural zeroes arising when some messages do not satisfy the truth conditions of some utterances) already high in speaker conditional entropy $H(U|M)$.

This work shows that least–effort optimization, and not simply heuristic utility maximization, may give rise to human pragmatic reasoning. Therefore, the optimization principle we identified may shape the evolution of pragmatic skills and more generally, the evolution of language. In addition, we have directly linked the dynamics of RSA recursive reasoning to the dynamics of the known alternating minimization algorithm. This provides new theoretical grounds for further studying the dynamics of pragmatic reasoning and its role in language evolution.

References

Bergen, L., Levy, R., & Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, *9*(20).

Camerer, C. F., Ho, T., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, *119*(3), 861-898.

Csiszár, I., & Shields, P. (2004). Information theory and statistics: A tutorial. *Foundations and Trends in Communications and Information Theory*, *1*(4), 417-528.

Frank, M. C., Emilsson, A. G., Peloquin, B., Goodman, N. D., & Potts, C. (2018). *Rational speech act models of pragmatic reasoning in reference games.* (PsyArXiv preprint of July 2, 2018, https://doi.org/10.31234/osf.io/f9y6b)

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998.

Franke, M., & Degen, J. (2016). Reasoning in reference games: Individual- vs. population-level probabilistic modeling. *PLOS ONE*, *11*(5), 1-25.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818-829.

Jucker, A. (Ed.). (1995). *Historical pragmatics: Pragmatic developments in the history of English* (Vol. 35). John Benjamins Publishing.

Levy, R. P. (2018). Communicative efficiency, uniform information density, and the rational speech act theory. In T. Rogers, M. Rau, X. Zhu, & C. Kalish (Eds.), *Proceedings of the 40th annual meeting of the cognitive science society* (pp. 684–689). Austin, TX: Cognitive Science Society.

Peloquin, B. N., Goodman, N. D., & Frank, M. C. (2019). The interactions of rational, pragmatic agents lead to efficient language structure and use. In *Proceedings of the 41st annual meeting of the Cognitive Science Society* (pp. 912–917). Austin, TX: Cognitive Science Society.

Scott-Phillips, T. C. (2015). Nonhuman primate communication, pragmatics, and the origins of language. *Current Anthropology*, *56*(1), 56-80.

Sperber, D., & Origgi, G. (2010). A pragmatic perspective on the evolution of language. In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language: Biolinguistic perspectives* (p. 124–132). Cambridge University Press.

Traugott, E. C. (2012). Pragmatics and language change. In K. Allan & K. M. Jaszczolt (Eds.), *The Cambridge handbook of Pragmatics* (pp. 549–566). Cambridge University Press.

Woensdregt, M., & Smith, K. (2018). *Pragmatics and language evolution.* PsyArXiv.

Yuan, A., Monroe, W., Bai, Y., & Kushman, N. (2018). Understanding the Rational Speech Act model. In T. Rogers, M. Rau, X. Zhu, & C. Kalish (Eds.), *Proceedings of the 40th annual conference of the cognitive science society*

(pp. 2759–2764). Austin, TX: Cognitive Science Society.
Zipf, G. K. (1949). *Human behavior and the principle of least effort.* Addison-Wesley (Reading MA).