

Background

Julian Grove and Aaron Steven White

Introduction

This course is about two major questions:

- How can semanticists use inference judgment datasets?
 - statistical properties of a dataset \leadsto semantic theory
- How should semantic theories be *modified* so that they may characterize statistical generalizations observed from datasets?

We:

- bring the compositional, algebraic view of meaning into contact with linguistic datasets by introducing and applying the framework of Probabilistic Dynamic Semantics (PDS; Grove and White 2024a; Grove and White 2025; Grove and White 2024b);
- PDS seamlessly integrates theories of semantic competence with accounts of linguistic behavior in experimental settings by taking a modular approach:

given a dataset involving some semantic phenomenon, and which exhibits certain statistical properties, this course offers an approach to developing both (a) *theories* of the meanings assigned to the expressions present in the dataset, and (b) *linking hypotheses* that directly relate these theories to linguistic behavior.

The ideas developed in this course build on and respond to existing probabilistic approaches to semantics and pragmatics, including those which use computational modeling to characterize inference (van Eijck and Lapin 2012; Zeevat and Schmitz 2015; Franke and Jäger 2016; Brasoveanu and Dotlačil 2020; Bernardy et al. 2022; Goodman, Tenenbaum, et al. 2024). Such models are motivated, in part, by the observation that linguistic inference tends to display substantial *gradiance*, giving rise to quantitative patterns that traditional semantic theory has difficulty capturing. Meanwhile, they often aim to explain Gricean linguistic behavior (Grice 1975) by regarding humans as Bayesian reasoners. Indeed, due to this emphasis on pragmatic principles, much modeling work blurs the semantics/pragmatics distinction, rendering the connection to traditional semantic theory somewhat opaque.

To take a paradigm case, models within the Rational Speech Act (RSA) framework consider human interpreters as inferring meanings for an utterance which maximize the utterance's utility relative to a set of possible alternative utterances (Frank and Goodman 2012; Lassiter 2011; Goodman and Stuhlmüller 2013; Goodman and Frank 2016; Lassiter and Goodman

2017; Degen 2023, i.a.). Probabilistic models of linguistic inference, including RSA, tend to encode Bayesian principles of probabilistic update in terms of *Bayes' theorem*, which states that the posterior probability of an event given an observation is proportional to the prior probability of the event, multiplied by the likelihood of the observation given the event. RSA models give an explicit operational interpretation to Bayes' theorem by assuming that prior distributions over inferences encode world knowledge, and that likelihoods represent the utility-maximizing behavior of a pragmatic speaker.

Despite their success in modeling a wide variety of semantic and pragmatic phenomena, probabilistic models of linguistic data remain largely divorced from semantic and pragmatic practice, both in theory and in implementation. RSA models, for example, regard the semantic interpretations which humans pragmatically reason about as being provided by a *literal listener* that determines a distribution over inferences, given an utterance (see Degen 2023). But aside from the constraint that the literal listener's posterior distribution is proportional to its prior distribution (i.e., it acts as a mere filter), the semantic components of RSA models are generally designed by researchers on an *ad hoc* basis: on the one hand, the space I of possible inferences must be decided by individual researchers in a way that depends on the task being modeled; on the other hand, the relation $\llbracket \cdot \rrbracket$ between utterances and inferences is typically assumed without a justified connection to any semantic theory using, e.g., an explicit grammar fragment in the style of Montague (1973).

Given the background, this course introduces students to a novel approach to probabilistic meaning which integrates traditional Montague semantics, as well as ideas in compositional dynamic semantics, with probabilistic computational models in a completely seamless fashion. The theoretical framework and methodology we introduce retain the crucial benefits of both kinds of approach to meaning: specifically, PDS may be used to construct probabilistic models of human inference data, and it is in principle compatible with existing probabilistic modeling paradigms such as RSA; meanwhile, it seamlessly connects probabilistic models to compositional dynamic semantics in the Montagovian tradition by providing a setting to write full-fledged grammar fragments.

PDS integrates probabilistic modeling with semantic theory by couching probabilistic approaches to meaning entirely within typed λ -calculus, following Grove and Bernardy 2023. Grove and Bernardy present a formalism for studying probabilistic effects—e.g., sampling from a probability distribution or performing Bayesian update—based on *monads*, a notion from category theory which has seen many fruitful applications within linguistic semantics (Shan 2002; Charlow 2014; Charlow 2020, i.a.). PDS builds on Grove and Bernardy's monadic framework by providing a full-fledged theory of dynamic discourse update, integrating aspects of discourse such as

the common ground, the question under discussion (QUD; Ginzburg 1996; Roberts 2012; Farkas and Bruce 2010), and uncertainty about lexical meaning. Crucially, given a semantic theory of some discourse phenomenon couched with PDS, one may obtain a probabilistic model of some linguistic dataset, given a particular *response function* (Grove and White 2024a; Grove and White 2025; Grove and White 2024b).

Semantic domains to be investigated in the course

This course introduces PDS in the context three empirical datasets which it uses it to model. These datasets study factivity, vagueness, and the QUD.

Factivity is a property associated with certain clause-embedding predicates, e.g., *love* and *hate*, which trigger the inference that their complement clause is true, even in entailment-cancelling contexts (e.g., *Jo {loves, doesn't love} that it's raining \leadsto it's raining*). Both Grove and White (2024a) and Grove and White (2025) use inference datasets to study the factivity of various clause-embedding predicates.

Vagueness is a property pertaining to certain natural language predicates; e.g., *Jo is tall* produces substantial uncertainty about the minimum bound on Jo's actual height, leading *tall* to be categorized as vague. The course instructors are currently collecting data on the dynamics of vagueness in discourse, which they will then model using PDS.

The QUD provides an important construct within PDS, since it is used to model the *question prompt* in the experimental tasks used to collect inference datasets. Grove and White (2025) manipulate the question prompt used to measure factive inferences and show that explicitly modeling it within PDS improves models of experimental data. The instructors also have plans to conduct further experiments which explicitly manipulate the QUD used in experimental tasks.

Tentative outline

Session 1

This session provides an introduction to existing approaches to probabilistic semantics and pragmatics, with an emphasis on RSA models, as well Bayesian statistical models of semantic inference data. We introduce RSA from a bird's eye view, and survey a few of its applications within linguistic semantics for illustration. We also use this session to discuss the limitations of RSA in practice—as well as those of statistical models of inference, more generally—mentioned above; i.e., that they have not been fully integrated with semantic analysis within the Montagovian tradition, presented in the form of grammar fragments.

Session 2

This session introduces the monadic approach to probabilistic semantics of Grove and Bernardy (2023), along with the foundational aspects of PDS, which build on it. The following aspects of PDS will be covered:

- Basic discourse constructs, such as the common ground and QUD, and the relationship between these constructs and the structure of inference tasks.
- The determination of semantic analyses within PDS and their associated prior probability distributions.
- The determination of response functions and their associated likelihood functions.

Session 3

This session applies PDS to the collection of linguistic datasets related to factivity, which are discussed in Grove and White 2024a.

Session 4

This session applies PDS to the collection of linguistic datasets related to the dynamics of vagueness.

Session 5

This session applies PDS to the collection of linguistic datasets related to the QUD, which are discussed in Grove and White 2025.

References

- Bernardy, J.-P. et al. (2022). “Bayesian Natural Language Semantics and Pragmatics”. In: *Probabilistic Approaches to Linguistic Theory*. Ed. by J.-P. Bernardy et al. CSLI Publications.
- Brasoveanu, A. and J. Dotlačil (2020). *Computational Cognitive Modeling and Linguistic Theory*. en. Vol. 6. Language, Cognition, and Mind. Cham: Springer International Publishing. DOI: [10.1007/978-3-030-31846-8](https://doi.org/10.1007/978-3-030-31846-8).
- Charlow, S. (2014). “On the semantics of exceptional scope”. PhD Thesis. New York: New York University.
- Charlow, S. (2020). “The scope of alternatives: indefiniteness and islands”. en. In: *Linguistics and Philosophy* 43.4, pp. 427–472. DOI: [10.1007/s10988-019-09278-3](https://doi.org/10.1007/s10988-019-09278-3).
- Degen, J. (2023). “The Rational Speech Act Framework”. In: *Annual Review of Linguistics* 9.1, pp. 519–540. DOI: [10.1037/rev0000186](https://doi.org/10.1037/rev0000186).

- Farkas, D. F. and K. B. Bruce (2010). “On Reacting to Assertions and Polar Questions”. In: *Journal of Semantics* 27.1, pp. 81–118. DOI: [10.1093/jos/ffp010](https://doi.org/10.1093/jos/ffp010).
- Frank, M. C. and N. D. Goodman (2012). “Predicting Pragmatic Reasoning in Language Games”. In: *Science* 336.6084. Publisher: American Association for the Advancement of Science, pp. 998–998. DOI: [10.1126/science.1218633](https://doi.org/10.1126/science.1218633).
- Franke, M. and G. Jäger (2016). “Probabilistic pragmatics, or why Bayes’ rule is probably important for pragmatics”. en. In: *Zeitschrift für Sprachwissenschaft* 35.1. Publisher: De Gruyter, pp. 3–44. DOI: [10.1515/zfs-2016-0002](https://doi.org/10.1515/zfs-2016-0002).
- Ginzburg, J. (1996). “Dynamics and the semantics of dialogue”. In: *Logic, Language, and Computation*. Ed. by J. Seligman and D. Westerståhl. Vol. 1. Stanford: CSLI Publications, pp. 221–237.
- Goodman, N. D. and M. C. Frank (2016). “Pragmatic Language Interpretation as Probabilistic Inference”. en. In: *Trends in Cognitive Sciences* 20.11, pp. 818–829. DOI: [10.1016/j.tics.2016.08.005](https://doi.org/10.1016/j.tics.2016.08.005).
- Goodman, N. D. and A. Stuhlmüller (2013). “Knowledge and implicature: modeling language understanding as social cognition”. eng. In: *Topics in Cognitive Science* 5.1, pp. 173–184. DOI: [10.1111/tops.12007](https://doi.org/10.1111/tops.12007).
- Goodman, N. D., J. B. Tenenbaum, et al. (2024). *Probabilistic Models of Cognition - 2nd Edition*.
- Grice, H. P. (1975). “Logic and Conversation”. In: *Syntax and Semantics*. Ed. by P. Cole and J. L. Morgan. Vol. 3, Speech Acts. New York: Academic Press, pp. 41–58.
- Grove, J. and J.-P. Bernardy (2023). “Probabilistic Compositional Semantics, Purely”. en. In: *New Frontiers in Artificial Intelligence*. Ed. by K. Yada et al. Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, pp. 242–256. DOI: [10.1007/978-3-031-36190-6_17](https://doi.org/10.1007/978-3-031-36190-6_17).
- Grove, J. and A. S. White (2024a). *Factivity, presupposition projection, and the role of discrete knowledge in gradient inference judgments*. LingBuzz Published In: submitted.
- Grove, J. and A. S. White (2024b). “Probabilistic dynamic semantics”. University of Rochester.
- Grove, J. and A. S. White (2025). “Modeling the prompt in inference judgment tasks”. en. In: *Experiments in Linguistic Meaning* 3, pp. 176–187. DOI: [10.3765/elm.3.5857](https://doi.org/10.3765/elm.3.5857).
- Lassiter, D. (2011). “Vagueness as Probabilistic Linguistic Knowledge”. en. In: *Vagueness in Communication*. Ed. by R. Nouwen et al. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 127–150. DOI: [10.1007/978-3-642-18446-8_8](https://doi.org/10.1007/978-3-642-18446-8_8).
- Lassiter, D. and N. D. Goodman (2017). “Adjectival vagueness in a Bayesian model of interpretation”. en. In: *Synthese* 194.10, pp. 3801–3836. DOI: [10.1007/s11229-015-0786-1](https://doi.org/10.1007/s11229-015-0786-1).

- Montague, R. (1973). “The Proper Treatment of Quantification in Ordinary English”. en. In: *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*. Ed. by K. J. J. Hintikka, J. M. E. Moravcsik, and P. Suppes. Synthese Library. Dordrecht: Springer Netherlands, pp. 221–242. DOI: [10.1007/978-94-010-2506-5_10](https://doi.org/10.1007/978-94-010-2506-5_10).
- Roberts, C. (2012). “Information Structure: Towards an integrated formal theory of pragmatics”. en. In: *Semantics and Pragmatics* 5, 6:1–69. DOI: [10.3765/sp.5.6](https://doi.org/10.3765/sp.5.6).
- Shan, C.-c. (2002). “Monads for natural language semantics”. In: *arXiv:cs/0205026*. arXiv: cs/0205026.
- Van Eijck, J. and S. Lappin (2012). “Probabilistic semantics for natural language”. In: *Logic and interactive rationality (LIRA)*. Ed. by Z. Christoff et al. Vol. 2. Citeseer, pp. 17–35.
- Zeevat, H. and H.-C. Schmitz, eds. (2015). *Bayesian Natural Language Semantics and Pragmatics*. en. Vol. 2. Language, Cognition, and Mind. Cham: Springer International Publishing. DOI: [10.1007/978-3-319-17064-0](https://doi.org/10.1007/978-3-319-17064-0).