Learning to Reason Pragmatically with Cognitive Limitations

Adam Vogel
acvogel@stanford.edu
Stanford Computer Science

Andrés Goméz Emilsson
nc07agom@stanford.edu
Stanford Psychology

Michael C. Frank
mcfrank@stanford.edu
Stanford Psychology

Dan Jurafsky
jurafsky@stanford.edu
Stanford Linguistics

Christopher Potts
cgpotts@stanford.edu
Stanford Linguistics

Abstract

Recursive Bayesian models of linguistic communication capture a variety of intricate kinds of pragmatic enrichment, but they tend to depend on the unrealistic assumption that agents are invariably optimal reasoners. We present a discriminative model that seeks to capitalize on the insights of such approaches while addressing these concerns about inferential power. The model relies on only approximate representations of language and context, and its recursive properties are limited to the training phase. The resulting behavior is often not optimal, but we present experimental evidence that this suboptimal behavior is closely aligned with human performance on both simple and complex reference games.

Introduction

Recursive Bayesian models of language production and comprehension capture a variety of intricate phenomena concerning context dependence and pragmatic enrichment (Jäger, 2007; Franke, 2009; Frank & Goodman, 2012; Bergen et al., 2012; Vogel et al., 2013; Smith et al., 2013). In these models, speakers and listeners reason about each other recursively in order to achieve ever more optimal communication systems. These approaches offer precise, algorithmic perspectives on philosophical and linguistic theories of communication (Lewis, 1969; Grice, 1975; Horn, 1984), and they make robust predictions about experimental data (Stiller et al., 2011; Rohde et al., 2012; Degen et al., 2013).

Despite the success of these models, they raise concerns about inferential power, in that they assume that the agents are invariably optimal reasoners with unbounded computational resources. These concerns can be mitigated by stipulations about depth of iteration (Camerer et al., 2004; Franke, 2009; Jäger, 2007, 2012), but the models remain computationally demanding and powerful.

We seek to capitalize on the insights of these approaches while addressing these concerns. We define a discriminative model of pragmatic reasoning that requires no explicit representation of the context. Rather, it relies only on features of the environment and language. In addition, the recursive aspects of the model are limited to training: we employ a *self-training* regime in which, starting with basic models of the speaker and the hearer, we use the speaker to generate supervised training data for the listener, and vice versa. Once this phase is complete, the model makes decisions without any recursion. The models are both more efficient and more fallible than the above generative ones.

Our model also offers a way to reconcile previous explanations of the interpretation or production of pragmatically complex utterances: slowly via complex recursive inferences made as each sentence is processed, (Geurts, 2009; Huang & Snedeker, 2009) or quickly via inferences that are precompiled and cached based on previous interactions (Levinson, 2000; Grodner et al., 2010; Smith et al., 2013). Instead, our discriminatively-trained model instantiates a third possibility, extending Jurafsky 2004: learning to directly map surface linguistic cues to speaker intent. Like the interpretive models, a learned model explains how context-sensitive inferences could be drawn at communication time; like the cached models, it explains why processing could be fast and direct.

Our central question is whether our model's behavior matches human performance across a wide range of situations. To address this, we use collaborative reference games (Rosenberg & Cohen, 1964; Clark & Wilkes-Gibbs, 1986; DeVault et al., 2005) in which a speaker refers to an object in a shared visual scene and the listener uses the speaker's message to try to guess the intended referent. By manipulating the properties of the scene and the speaker's available messages, we can ensure that pragmatic reasoning is required for reliable success. We report several experiments that identify the bounds on human performance in these reference games, and we compare human performance to our model, showing that its inferences closely align with human performance.

Reference Games

Figure 1 depicts a reference game scenario. There are three potential referents (A, B, and C), each with a pre-specified set of properties (wearing glasses, wearing a hat, having a mustache). The *speaker* is privately assigned one of the referents. Using a message from a pre-defined vocabulary, the speaker tries to convey the identity of this referent to the *listener*. The listener uses the message to choose one of the referents. Intuitively, the two participants' goals are aligned: they win just in case the speaker's referent is the hearer's choice.

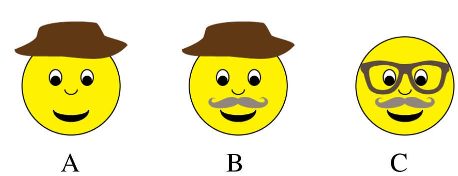


Figure 1: Scenario for a simple reference game.

Code and data: github.com/acvogel/discriminative-ibr

Definition

Formally, a reference game is a tuple $G = (M, T, [\![\cdot]\!])$, where T is the set of targets, M is the set of messages, and $[\![\cdot]\!] : M \rightarrow 2^T$ is the semantics of the messages, which dictates which targets each message is true of. A speaker S is a (possibly stochastic) function $G \times T \rightarrow M$ from a game and target to a message. Similarly, a listener L is a function $G \times M \rightarrow T$ from a game and a message to a target. For a given run of a game G and target t, the speaker produces the message $m = S(G,t)$, the listener selects a target $t' = L(G,m)$, and they win iff $t = t'$.

Literal Speaker and Listener

To initialize the learning procedure described in the next section, we define literal speakers and listeners, who rely entirely on the semantics of the messages, with no consideration of the other participant's behavior. The literal speaker S_0, when given a target to refer to, picks uniformly at random from messages which are true of the target:

$$S_0(G,t) = \text{Uniform}\left(\{m | t \in [\![m]\!]\}\right) \tag{1}$$

Similarly, the literal listener L_0, when given a message m, picks randomly from targets in the semantics of the message:

$$L_0(G,m) = \text{Uniform}\left(\{t | t \in [\![m]\!]\}\right) \tag{2}$$

Discriminative Best Response

Our model relies on iterated *self-training*, alternating between discriminatively training the listener and the speaker. This method requires no generative model of the context or messages, but rather only a shallow feature vector representation. Furthermore, there is no explicit recursive reasoning procedure of the sort common in the generative approaches discussed above. At evaluation time, our trained listener simply generates features for the problem and utterance, and computes the model activation for each possible target.

Algorithm 1 describes the training algorithm. It takes as input a set of reference games $\mathbb{G}$, and a number of training iterations to perform, N. Starting with the literal speaker, we train a listener using the input reference games and the speaker (Algorithm 2). Then, using this newly trained listener, we retrain the speaker (Algorithm 3), using the current listener to create training data. We use an artificial neural network (ANN) as the discriminative classifier, but other learning algorithms would work as well. This procedure repeats for N iterations, yielding a self-trained ANN listener and speaker.

To train a listener (Algorithm 2), we use a speaker S to generate training data as follows. For each reference game $G \in \mathbb{G}$, and for each target $t \in T$, we query the speaker to get the speaker's chosen message $m = S(G,t)$. We then form a training set in which the input features are games combined with messages and the gold label is the target t that the speaker intended to refer to. The listener ANN uses a simple feature representation: a binary feature for the presence or absence of each feature for each target, and also a binary feature for each possible message. For three targets, three properties, and three utterances, this yields twelve binary features. The listener ANN has three outputs, one per target. Given a reference game and a message, the listener selects the target corresponding to the highest output activation.

The training procedure for speakers (Algorithm 3) proceeds similarly. For each reference game $G \in \mathbb{G}$ and for each target $t \in T$, we query the listener to determine whether there is a message m that the listener interprets as target t. If so, we add a training example with (G,t) as the features and m as the label. For some listeners, there is no message it interprets as a particular target, so we do not add those to the training set.

The listener and speaker ANNs have the same structure: a fully-connected network with 12 input features, one a hidden layer whose size we vary in the experiments, and 3 output nodes. The models are trained with back propagation (Rumelhart et al., 1986), using the PyBrain library (Schaul et al., 2010).

Algorithm: SelfTrain
Input: Reference games $\mathbb{G}$, Number of iterations N
Output: Listener L
Initialize $S = S_0$
for *i from* 1 *to* N **do**
    $L = TrainListener(\mathbb{G}, S)$
    $S = TrainSpeaker(\mathbb{G}, L)$
end
return L

Algorithm 1: Train listener and speaker ANNs for a given number of iterations, starting from the literal speaker S_0.

Algorithm: TrainListener
Input: Reference games $\mathbb{G}$, Speaker S
Output: Listener L
Initialize training data $X = Y = \emptyset$
foreach *Reference game* $G = (T,M) \in \mathbb{G}$ **do**
    foreach *Target* $t \in T$ **do**
        $m = S(G,t)$
        Append (G,m) to X
        Append t to Y
    end
end
return *ANN-Train*(X,Y)

Algorithm 2: Train a listener ANN from a given speaker.

Simulations with Synthetic Data

We first evaluate our ANN listener model on a variety of automatically generated reference games.

Experimental Design

We first generated the full set $\mathbb{G}_{\text{sep}}$ of three-target, three-feature reference games that have fully separating equilibria, that is, games that can be resolved to totally unambiguous speaker and listener strategies in the iterated best response (IBR) model of Franke (2009) and Jäger (2007, 2012). From

Algorithm: TrainSpeaker
Input: Reference games $\mathbb{G}$, Listener L
Output: Speaker S
Initialize training data $X = Y = \emptyset$
foreach *Reference game* $G = (T, M) \in \mathbb{G}$ **do**
  foreach *Target* $t \in T$ **do**
    if $\exists m$ *such that* $L(G, m) = t$ **then**
      Append (G, t) to X
      Append m to Y
    end
  end
end
return *ANN-Train*(X, Y)

Algorithm 3: Train a speaker ANN from a given listener.

$\mathbb{G}_{\text{sep}}$, we generated 1,206 specific reference instances, where an instance is a game G, a message m, and an intended target t. We separate these instances by the depth of reasoning that an IBR listener requires to identify a unique target. Level 0 problems require only the literal semantics of messages, i.e., the message uniquely identifies the target; level 1 problems require reasoning about a speaker that reasons about a literal listener; and level 2 problems require reasoning about a speaker that reasons about a level 1 listener. Using these instances (G, m, t), we self-train our listener model using only the game information G and the message m, holding out the target t for evaluation, as described in Algorithm 1.

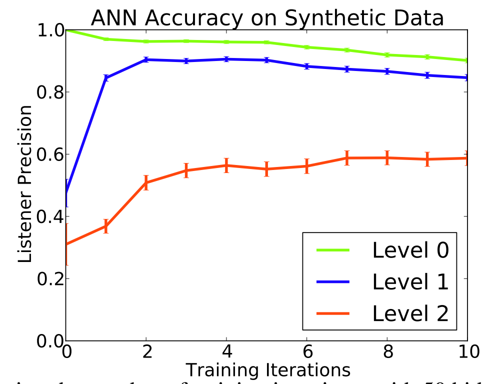
Results

We first evaluated how the number of training iterations affects model performance. Figure 2(a) shows how the accuracy of the training model changes over training iterations. The x-axis is how many training iterations the listener underwent. (The listener with 0 training iterations is the literal listener.) The y-axis is the precision of the model on the specific type of problem. We next explored how the size of the hidden layer affects listener performance, by varying the size of the hidden layer and evaluating each model's performance after 10 training iterations (Figure 2(b)).

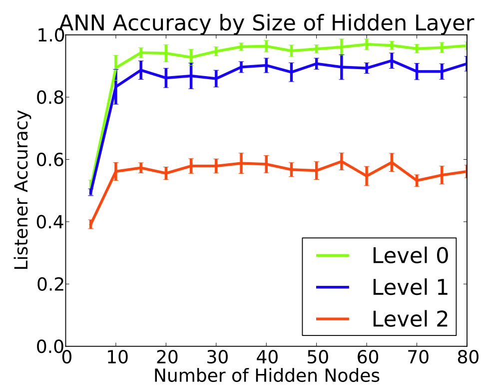
Discussion

Model accuracy decreases as the problem level (complexity) increases. The literal speaker performs perfectly on level 0 problems, as expected. It is surprising that the trained models perform less than perfectly on these easy problems. A listener trained *only* on level 0 problems achieves perfect accuracy on level 0 problems, so it seems that training on the more difficult problems leads to some degradation in performance on easy problems. The trained listeners perform well, with 91% accuracy on level 0, 85% on level 1 problems, and 59% on level 2 problems. Figure 2(b) shows that the size of the hidden layer has some effect on model performance, but that after a certain threshold the performance stabilizes.

We also evaluated IBR listener models for a variety of recursion depths (Figure 3). All have perfect accuracy on the level 0 problems (barely visible along the top of the plot), with rapidly increasing accuracy on the level 1 and 2 problems. The accuracies shown here are using the probabilities of each target given a message (Frank & Goodman, 2012; Bergen et al., 2012). If we instead always choose the target with the highest probability given a message (Franke, 2009; Jäger, 2007, 2012), the IBR model solves the problems exactly after one and two levels, respectively.



(a) Varying the number of training iterations, with 50 hidden nodes.



(b) Varying the number of hidden nodes, for 10 training iterations.

Figure 2: ANN Listener accuracy on synthetic data.

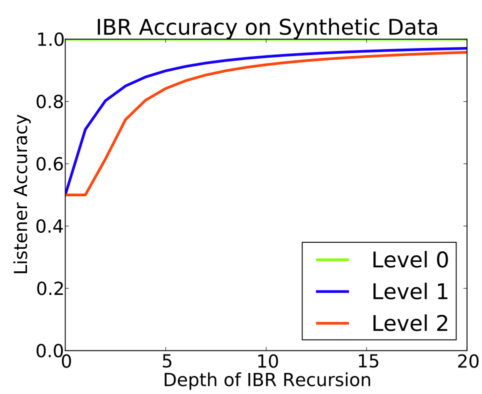


Figure 3: IBR listener accuracy on synthetic problems.

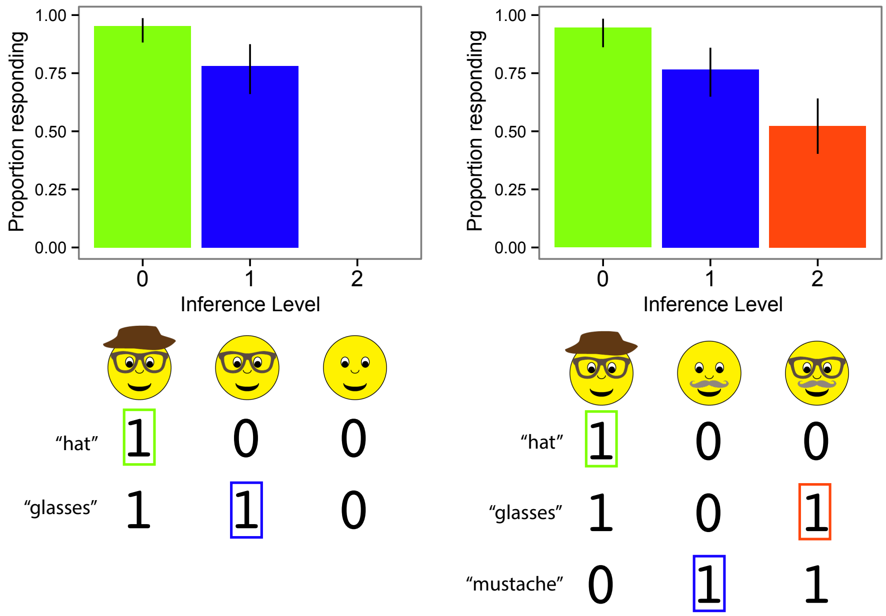


Figure 4: Human data and experimental matrices from Experiments 1 and 2.

Experiment 1: Simple Scalar Implicature

To quantitatively compare our model with human data, we conducted two experiments with human participants in which we asked them to play reference games that varied in their inferential complexity. We then compare performance of human participants to that of our discriminative model. Our first experiment used the simple referential context shown in the left of Figure 4, following Stiller et al. (2011).

Methods

Participants We recruited 120 participants on Amazon Mechanical Turk, of whom 65 received the level 0 stimulus, and 55 received the level 1 stimulus.

Stimuli For each participant, we generated a reference game with an underlying matrix description identical to that shown in the left of Figure 4. We generated the reference game by choosing a base item randomly from among six possible options: boat, friend (shown in Figure 4), pizza, snowman, sundae, and Christmas tree. Each of the possible items had three features that were plausible additions to the base item. For example, the friend item has a hat, glasses, and mustache, while the boat item has a sail, cabin, and motor. For this experiment, we randomly chose two features, randomly assigning them to rows in the underlying matrix (ensuring, e.g., that glasses was not always the target feature), and we randomly assigned targets to positions in the display (ensuring, e.g., that the target was not always in the middle).

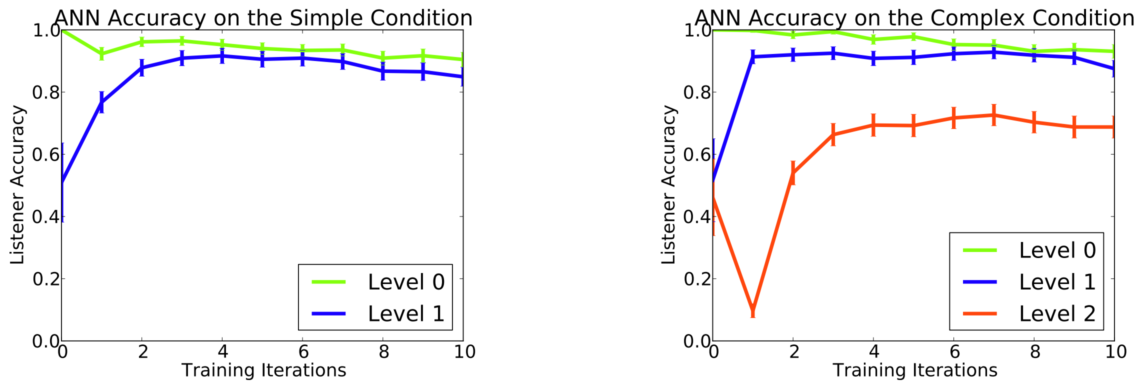
This reference game supports two possible inference types, which we refer to as level 0 and level 1, following our usage in the previous section. In the case of the display shown in Figure 4, the message "hat" unambiguously refers to the face with the hat and glasses; since there is no inference necessary, we call this a level 0 problem. In contrast, the message "glasses" could logically refer to the face with a hat and glasses, or the face with just glasses. A pragmatic inference is required to conclude that the message refers to the face *without* a hat; we refer to this as a level 1 problem.

Procedure Participants saw a webpage that first introduced them to an interlocutor, "Bob", who routinely engaged in some action (e.g., visiting his friends for the friend item). They then saw a scene like that shown in Figure 4 and read that "Bob can only say one word to communicate with you and he says: [target]", where [target] indicates the message relating to the particular condition they had been randomly assigned to (e.g., "glasses" for the level 1 inference in Figure 4). Participants were instructed to indicate which item they thought was Bob's target via a 3-alternative forced-choice. Afterwards, they completed a simple check question (provide the interlocutor's name), which we used to exclude non-compliant participants.

Results

Human performance is plotted in the bar graph in Figure 4. The level 0 utterance was trivial, with 95% of participants choosing correctly. Participants also made a substantial portion of implicature-consistent responses (75%) in the level 1 condition, replicating Stiller et al. (2011).

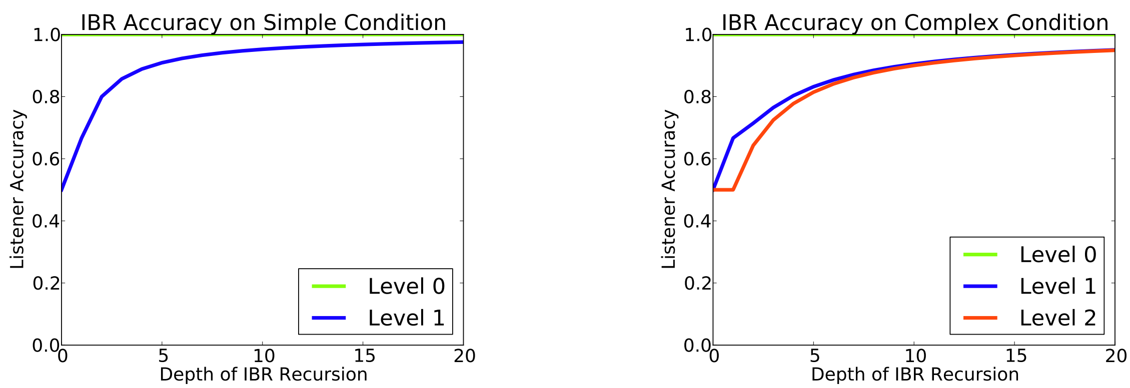
We next compare human performance to the ANN listener model, which was trained on $\mathbb{G}_{sep}$. The accuracy of an ANN listener with 50 hidden nodes is shown in Figure 5(a), for a



(a) ANN Simple

(b) ANN Complex

Figure 5: Evaluation of the ANN listener with 50 hidden nodes on the simple (a) and complex (b) settings.



(a) IBR Simple

(b) IBR Complex

Figure 6: Evaluation of IBR listeners with different depths of recursion.

variety of training iterations. The case of 0 training iterations corresponds to the literal listener L_0. We see that the literal listener gets all of the level 0 (unambiguous) problems correct, and 50% of the level 1 problems. The ANN slightly outperforms the humans but is well aligned with them: 91% accuracy on the level 0 problems, and 86% accuracy on the level 1 problems.

Lastly, we evaluated the iterated best response (IBR) model on this condition, shown in Figure 6(a). The IBR model is always correct on the level 0 problems, and quickly increases in accuracy on the level 1 problems as the depth of recursion increases. This evaluation uses the IBR probabilities to compute accuracy. If we instead evaluate by predicting the target with highest probability for each message, it gets all of the level 1 problems correct after one level of recursion.

Experiment 2: Complex Scalar Implicature

Methods

Participants We recruited 180 participants from Mechanical Turk, with 55 receiving the level 0 stimulus, 60 receiving level 1, and 65 receiving level 2.

Stimuli Stimuli were generated identically to those in Experiment 1, except that we used the base matrix shown on the right in Figure 4. With the setting shown there, the message "hat" is level 0, "mustache" is level 1, and "glasses" is level 2.

Procedure The procedure was identical to Experiment 1.

Results

Figure 4 shows human performance on the more complex scalar implicature task. Similar to the simple task, 95% correctly identified the level 0 referent, and 77% correctly picked the level 1 referent. However, humans had much more trouble with the level 2 inference, with just 52% selecting the correct referent given the utterance.

These results too align with the ANN model (Figure 5(b)). We see similar performance on the level 0 and level 1 problems as in the previous experiment. Importantly, the model also has much more difficulty with the level 2 problems (70% accuracy), which persists across number of training iterations. Although the model is more accurate than the human subjects, the results are qualitatively similar.

This stands in marked contrast to the IBR model's performance, summarized in Figure 6(b). This model always gets the level 0 problems correct. As the depth of recursion increases, the IBR confidence asymptotes to 1. In the same way as in the simple condition, if we instead evaluate the IBR model by choosing the highest probability target for each message, it correctly identifies the level 1 targets after one step of recursion, and all of the level 2 problems after two steps of recursion, thereby vastly outperforming our human subjects.

Conclusion

We presented a discriminative model of communication that embodies the central insights of recursive generative models of pragmatic reasoning but is more computationally efficient, particularly at decision time. Using experiments involving simple and complex reference games, we showed that the model displays human-like pragmatic behavior. In closing, we note that the models have additional advantages that we were unable to explore here, including (i) the ability to reason in terms of partial, heterogeneous representations of the environment, (ii) a decoupling of inferential power (depth of iteration) from memory (dimensionality of the hidden representations), and (iii) a level of computational efficiency that makes them scalable to truly massive problems involving language and action together.

Acknowledgments

We gratefully acknowledge the support of the Nuance Foundation and ONR grant N00014-13-1-0287.

References

Bergen, L., Goodman, N. D., & Levy, R. (2012). That's what she (could have) said: How alternative utterances affect language use. In *Cogsci 2012.*

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

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39.

Degen, J., Franke, M., & Jäger, G. (2013). Cost-based pragmatic inference about referential expressions. In *Cogsci 2013.*

DeVault, D., Kariaeva, N., Kothari, A., Oved, I., & Stone, M. (2005). An information-state approach to collaborative reference. In *Proceedings of the ACL interactive poster and demonstration sessions.*

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

Franke, M. (2009). *Signal to act: Game theory in pragmatics*. Institute for Logic, Language and Computation, University of Amsterdam.

Geurts, B. (2009). Scalar implicatures and local pragmatics. *Mind and Language*, *24*(1), 51–79.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3: Speech Acts).

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116*(1), 42–55.

Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 11–42).

Huang, T. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology*, *58*(3), 376–415.

Jäger, G. (2007). Game dynamics connects semantics and pragmatics. In *Game theory and linguistic meaning.* Amsterdam.

Jäger, G. (2012). Game theory in semantics and pragmatics. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 3). Berlin: Mouton de Gruyter.

Jurafsky, D. (2004). Pragmatics and computational linguistics. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 578–604). Oxford: Blackwell.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA.

Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press. (Reprinted 2002 by Blackwell)

Rohde, H., Seyfarth, S., Clark, B. Z., Jäger, G., & Kaufmann, S. (2012). Communicating with cost-based implicature: A game-theoretic approach to ambiguity. In *The 16th workshop on the semantics and pragmatics of dialogue.* Paris.

Rosenberg, S., & Cohen, B. D. (1964). Speakers' and listeners' processes in a word communication task. *Science*, *145*, 1201–1203.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536.

Schaul, T., Bayer, J., Wierstra, D., Sun, Y., Felder, M., Sehnke, F., et al. (2010). Pybrain. *Journal of Machine Learning Research*, *11*, 743–746.

Smith, N. J., Goodman, N. D., & Frank, M. C. (2013). Learning and using language via recursive pragmatic reasoning about other agents. In *NIPS 2013.*

Stiller, A., Goodman, N. D., & Frank, M. C. (2011). Ad-hoc scalar implicature in adults and children. In *Proceedings of the 33rd annual meeting of the Cognitive Science Society.* Boston.

Vogel, A., Bodoia, M., Potts, C., & Jurafsky, D. (2013). Emergence of Gricean maxims from multi-agent decision theory. In *NAACL 2013.*