Learning about Word Meaning from Pragmatically Enriched Data

Anonymous ACL submission

Abstract

The meaning of a natural language utterance can vary greatly depending on the context of the communication. An artificial agent interpreting natural language needs to be able to integrate models of the human speaker and the communicative goal in order to arrive at the correct interpretation. This paper introduces an approach integrating pragmatic reasoning about the conversational partner while learning representations from scratch. This leads to significant improvements over prior work that only considers pragmatics during inference or builds on fixed representations of literal meaning. Our artificial language learner is situated in a referential game about images, where we show that equipping the agent with explicit reasoning about the speaker and the shared observations, leads to faster learning, higher communicative success, and better generalization to changes in the environment.

1 Introduction

Everyday conversations are rich with implicatures. Being able to infer the intended meaning of an utterance beyond its literal content allows us to communicate efficiently. The process of how people attain interpretations including implicatures using a model of the speaker's intentions has long been studied (Grice, 1975; Horn, 1984; Fox and Katzir, 2011; Levinson, 2000).

In recent years, deep learning methods learning from large natural language corpora have led to great advances in natural language understanding and generation. Large scale datasets however, rarely allow for the study of language in its most natural function: as a tool of communication between agents who have a communicative goal and are surrounded by extra-linguistic context.

While the communicative context is often not recorded in large natural language corpora, it is justified to assume that any natural language data by humans was generated using some form of pragmatic reasoning. In this work, we investigate how to best learn from data where the intended meaning contains implicatures. Our main hypothesis is that inferring the literal meaning from contextual language use creates more robust representations across contexts and leads to faster learning.

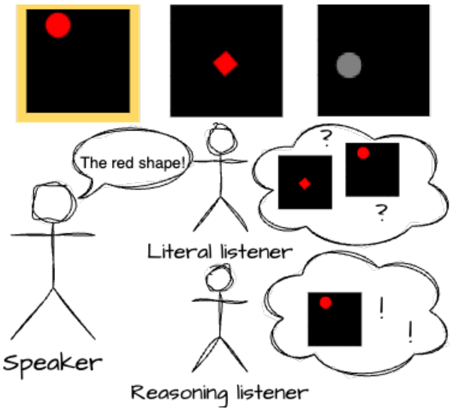


Figure 1: The speaker is asking for the red object. For a literal listener, this is ambiguous. A reasoning listener can conclude that the speaker is asking for the red circle, as "square" would have been a more informative message for the other red object.

When listeners derive the meaning of an utterance, they engage in counterfactual reasoning about alternative sentences that the speaker could have uttered (Davies et al., 2022; Frank and Goodman, 2012). The interaction in Figure 1 depicts an instance of such pragmatic reasoning about alternatives within our simple environment. According to pragmatic theory (Grice, 1975) the same process accounts for the interpretation *"They are in the office for the rest of the week"*, when we hear the sentence *"We are not in the office on Mondays"*.

Previous work has investigated how to adjust the meaning of an utterance using the context and communicative goal (Vedantam et al., 2017; Monroe et al., 2017; Vogel et al., 2013), with the assumption that the context-independent literal interpretations are already given. It is less well studied how language users acquire these literal interpretations that provide the starting point for further reasoning. This is especially interesting if the observed data already contains the result of reasoning that depends on the context.

In our work, we model an agent who learns about language from speakers using varying complexity of pragmatic inference. We show that it is possible for the language learner to differentiate through the speaker's production process and this way directly learn about literal meaning. We investigate how explicit models of the other agent influences learning and communicative success in a referential game about synthetic images. We show that a reasoning learners equipped with pragmatic reasoning during training:

- can successfully learn to interpret messages in a communicative task.
- learn faster in the initial stages as opposed to models that do not integrate explicit models of the speaker.
- achieve higher task accuracy at evaluation time compared to models that only apply reasoning during evaluation.
- learn interpretation functions that are more robust to changes in the environment.

2 Background

We situate our listener in an image-based version of Lewis's signaling game (Lewis, 1969). Image-referential games are commonly used to study the benefit of speakers and listeners reasoning about each other in context (Lee et al., 2018; White et al., 2020; Andreas and Klein, 2016). These works however, do not provide a principled way of learning the literal interpretations that provide the starting point for pragmatics. Image-captioning datasets are commonly used to learn the base speaker (Nie et al., 2020) or listener (Liu et al., 2023). Image-captions however, were not produced in the context of other images so they do not contain conversational phenomena found in referential games, such as choosing the most informative or shortest description that still leads to successful communication. Therefore, pragmatic models building on datasets that were not produced in context, do not address the problem of learning literal interpretations from contextualized language use.

Other works do build on datasets that were generated in the context of referential games. These collections might include human-human game-play or synthetically generated language (Monroe et al., 2017; White et al., 2020). However, when it comes to learning the literal interpretations, they resort to the methods mentioned earlier: they treat the referring expressions provided for each image as a caption, ignoring the fact that these captions were produced in the context of other distractor images.

In our version of the game, at each turn a collection of N images $(o_1, ..., o_N)$ is provided, with the speaker having knowledge of a specific target image o_t, where $1 \leq t \leq N$. The listener's objective is to correctly identify the target image index t given the speaker's message w. The messages may contain multiple words by combining words from a fixed vocabulary.

2.1 Literal meaning and the base listener

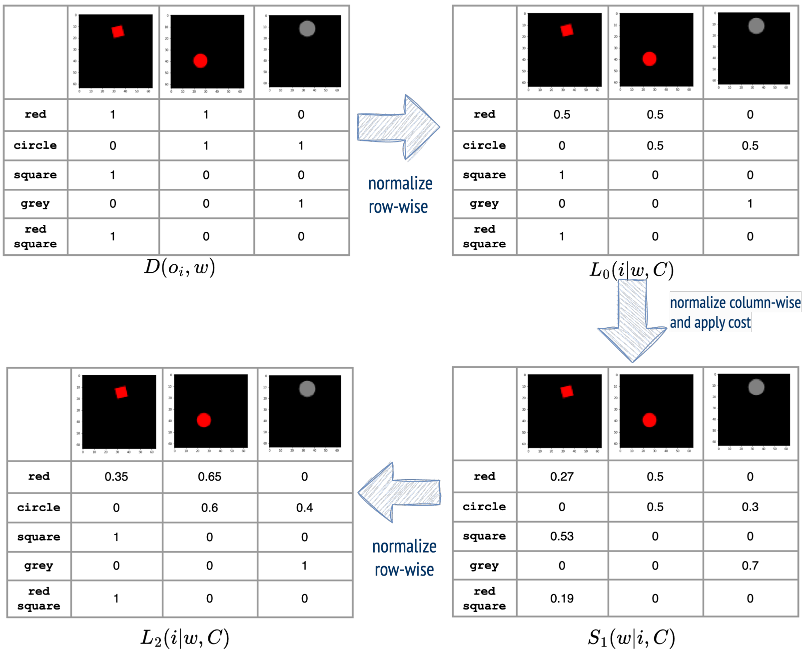
Frank and Goodman (2012) provide a concise model for how speakers and listeners reason about each-other when sharing referential content called the Rational Speech Act model. As a starting point, the model assumes an underlying literal interpretation. This is a function $D(w, o)$ of an utterance w and an observation o, in our case an image. In the formulation of Goodman and Frank (2016) the base interpretation function is a 0-1 valued indicator of the set of messages that are true of the image o. In line with other work, we replace this binary function with a real-valued similarity between the observed image and text.

$$D(o_i, w) = \phi_\theta(o_i)^T \gamma_\theta(w) \qquad (1)$$

Each image is individually embedded with a CNN following the ResNet architecture (He et al., 2016). The message embedding is computed by an RNN with Gated Recurrent Units (Cho et al., 2014). The listener models the distribution over the indices in an ordered set of images. This expresses the likelihood of each image given the message and all the observed images. The simplest listener distribution is produced by normalizing the score assigned by the literal interpretation function over all the images in a given context C.

$$L_0(i|w, C) = \frac{e^{D(o_i, w)}}{\sum_{j=1}^{|C|} e^{D(o_j, w)}} \qquad (2)$$

In previous work, the literal interpretations are initialized by functions learned outside of the context of a referential game. This is most commonly



red	1	1	0
circle	0	1	1
square	1	0	0
grey	0	0	1
red square	1	0	0

$D(o_i, w)$

red	0.5	0.5	0
circle	0	0.5	0.5
square	1	0	0
grey	0	0	1
red square	1	0	0

$L_0(i|w, C)$

red	0.35	0.65	0
circle	0	0.6	0.4
square	1	0	0
grey	0	0	1
red square	1	0	0

$L_2(i|w, C)$

red	0.27	0.5	0
circle	0	0.5	0.3
square	0.53	0	0
grey	0	0	0.7
red square	0.19	0	0

$S_1(w|i, C)$

Figure 2: The iterative process of creating higher level speakers and listeners. The listener's distribution has lower entropy with more iterations.

done by learning a function that maximizes similarity between an image and a corresponding description (White et al., 2020; Lazaridou et al., 2020; Andreas and Klein, 2016).

2.2 Recursive reasoning in speaker-listener games

Here we explain how to add recursive reasoning on top of a base L_0 listener model. The speaker produces a message that maximizes the probability that the listener chooses the right image and also considers the cost of each message w. This means that the speaker has an internal model of the listener.

$$S_n(w|C,i) = \frac{e^{\lambda(log(L_{n-1}(i|C,w))-cost(w))}}{\sum_{w' \epsilon V} e^{\lambda(log(L_{n-1}(i|C,w'))-cost(w'))}} \tag{3}$$

The speaker's probability of producing a message given an image and the context is proportional to the listener's probability of identifying that image based on the message and the same context. The formulation in Equation 3 introduces a parameter λ that controls how rational a speaker's choices are given an internal listener distribution. In this work, we only consider fully rational speakers with $\lambda = 1$.

The speaker equation in 3 also adds a cost for each message, this means that communication is not for free. If we consider the cost to be the same as the negative logarithm of the prior over messages, then we recover the Bayes formula. In this work, we consider a cost function that assigns a constant weight to each word, which gives a linear relation between message length and cost.

Now we define a more complex listener that can also have an internal model of a speaker:

$$L_n(i|C,w) \propto S_{n-1}(w|C,i)P(C,i) \tag{4}$$

By applying Equations 3 and 4 in an alternating fashion, we can produce higher level speakers and listeners. Figure 2 shows the process of altering the underlying literal representations to create the L_0 and then L_2 distributions. This example only considers a fixed set of five messages to illustrate the effect of longer messages being more expensive. The first column of S_1 shows that "red square" is not likely to be produced because of its cost.

Contrasting L_0 and L_2 in this example shows how the iterative process of applying Equations 3 and 4 creates implicatures in context: the message "red" is unambiguous for L_2 while it is ambiguous for L_0. Frank and Goodman (2012) show that humans reliably interpret messages produced by a S_3 speaker consistent with a L_2 listener.

3 Reasoning while Learning

In the previous section we saw how to perform recursive reasoning on top of a given literal representation $D(o, w)$. However, the optimal initial representations are likely influenced by the reasoning itself. In this work, we would like to learn the parameters of the literal representations in the presence of a communicative goal and context. Therefore, we propose to already apply recursive reasoning during training.

Reasoning learners that use models of the speaker's production process seek to update the weights of the literal interpretation function $D(o, w)$ but they need to do so by considering the repeated application of Equations 3 and 4. We would like to derive the gradients of the reasoning process with respect to the literal representations so we can update these parameters through stochastic gradient descent. We achieve this by repeated application of the chain rule through the hierarchical reasoning.

First, the reasoning listener has to differentiate the listener equation 4. The listener's distribution is computed by normalizing the speaker scores from the previous level over all images in the context. For simplicity, we abbreviate $L_n(i|C, w)$ as L_n^{iw} and $S_{n-1}(w|C, j)$ as S_{n-1}^{jw}.

$$\frac{\partial L_n^{iw}}{\partial S_{n-1}^{jw}} = L_n^{iw}(\delta_{ij} - L_n^{jw}) \tag{5}$$

This is the derivative of the softmax function that is applied to the scores assigned by the speaker from the previous level.

The speaker equation is also a computation of normalisation of the listener scores from the previous level, with the additional application of the messaging cost.[1]

In the case of the speaker, the normalization happens over all possible messages $w \in V$. This is the most expensive step in the hierarchical reasoning process. In many natural language applications it is even prohibited by the fact that the set of all possible utterances is infinite. While exact inference is intractable, there are many papers discussing approximations (Cohn-Gordon et al., 2018; Liu et al., 2023; Lazaridou et al., 2020; White et al., 2020). Due to the small number of messages in our game, we can compute the gradients exactly:

$$\frac{\partial S_n^{iw}}{\partial L_{n-1}^{iw'}} = S_n^{iw}(\delta_{w'w} - S_n^{iw'}) \tag{6}$$

This allows us to examine the effect of recursive reasoning during learning without the confounds from approximations.

Depending on the depth of the listener, learning chains Equations 5 and 6 an appropriate number of times. Reasoning learners backpropagate through the hierarchical reasoning and update the weights of the image- and utterance-embedding models with stochastic gradient descent. The representations learned this way take pragmatics into account, as opposed to the prevalent approach to pragmatic reasoning, where the literal interpretations have not been optimised to match the reasoning process.

4 Related work

A growing number of works consider adding explicit pragmatic capabilities to language understanding and generation systems. The first group of these works upgrades models with additional reasoning only during evaluation time. Such models have been used to generate better image-descriptions in context (Nie et al., 2020; Andreas and Klein, 2016; Lazaridou et al., 2020), informative referring expressions about location (Golland et al., 2010) and visual objects (Mao et al., 2016).

Cohn-Gordon and Goodman (2019) integrate a listener model in their machine translation system in order to find better translations to potentially ambiguous sentences. Similarly, models of the speaker have been used in instructions following agents to resolve unclear utterances (Fried et al., 2018b,a).

Closer to our interest, the second group of models already consider pragmatic reasoning during learning. Smith et al. (2013) provides a sampling-based inference for learning the "some and maybe all" literal semantic content for the word "some". Vogel et al. (2013) investigate heuristics for finding the underlying literal interpretations in a navigation game between humans where the referring expressions for the location are rich with scalar implicatures. Liu et al. (2023) learn a speaker in an image-referential game that approximates a model of the listener during training. Their base listener on the other hand is initialised with a literal model that was trained outside of the context of the communicative game.

Most similar to our work, Monroe and Potts

---

[1] The cost of the messages is proportional to their length and is kept fixed during training. The value of the cost is shared knowledge between the speaker and the listener. This corresponds to the intuition that they both experience the same environmental pressures and difficulty of production. We note it though, that it would be an interesting direction to make the cost of communication a learnable parameter, that the players also need to adjust during training.

(2015) learn a literal lexicon of words and symbolic image features by modeling a speaker with an internal listener in an image-referential game. They also derive gradients by differentiating the speaker equation to update the listener model. Unlike the flexible neural network formulations in this paper, they use log-linear model with a hand-crafted feature space for images and words.

5 Experiments

5.1 Data

We create a new environment based on the ShapeWorld dataset (Kuhnle and Copestake, 2017). This dataset was designed for the study of multimodal language understanding and has been previously used to investigate messaging strategies (White et al., 2020; Wang et al., 2021).

Each game consists of a target image and a variable number of N distractor images. Images are described by one out of six different colors and a shape that can take four different values. The shape and size of the objects is randomized on a 64x64 grid which creates a large variation of candidate pictures. In the original dataset, the shapes and colors as well as the index of the target image are drawn from a uniform distribution. If the target is identical to one of the distractors, the sample is rejected, as there is no unique referring expression that would identify the target.

First, we change the way the speaker chooses the message. Instead of the rule-based method of Kuhnle and Copestake (2017), we use an exact implementation of the rational speaker defined in Equation 3. This way we can create speakers with different depth of recursive reasoning. Our speakers are not learned, they are knowledgeable users of the language: they have access to the underlying true literal meaning representations which indicates the mapping between words and image properties.

We parameterize the process that generates the image tuples for each game by three sets of probability distributions: the prior over the shapes $P(S)$, the probability of colors given a shape $P(C|S)$ and a third distribution the controls the co-occurrence of shapes $P(S|S)$. This enables us to introduce correlations between object properties, making it more difficult to find the right referring expression.

We sample $P(S)$, the four conditionals for $P(S|S)$ and for $P(C|S)$ from different Dirichlet distributions. $P(S)$ has four concentration parameters $(\alpha_{ellipse}, ..., \alpha_{square})$. $P(C|S)$ has six concentration parameters, one for each color $(\alpha_{red}, ..., \alpha_{blue})$ repeated over all shapes to create four different conditional distributions. Each of the four conditionals characterizing $P(S|S)$ has four parameters belonging to each shape.

5.2 Experimental setup

The fact that we have full control over the speaker's messaging strategy and the data generating process behind context images allows us to create interesting learning environments for the listeners. For example, we can alter the level of the speakers that the listeners learn from, or modify the environment between training and evaluation time.

We train listeners of three different levels. The level 0 listener, L_0 has no internal model of the speaker. This is the model that we will upgrade to higher level reasoning during evaluation to create a similar setup that was explored in previous work. The level 2 listener, L_2 models a level 1 speaker, and the level 4 listeners applies a further step of reasoning by modeling a level 3 speaker.

We create two different levels of speakers to pair them with our learning listeners: S_1 is level 1 and S_3 is level 3. In our experiments, we create learning scenarios with all possible pairings of listener and speaker levels. This allows us to examine how the learning is influenced if there is a mismatch between the actual level of the learner and the level of the learner that the speaker internally models.

We also design two further versions of the game: one where communication is free, and one where each word has a cost 0.6. This cost was derived to make sure that the speakers of all levels will send the shorter message, where it is possible, while also guaranteeing that communication is not too expensive for speakers to send ambiguous messages that cannot be resolved.

Generating image tuples We set all alphas that belong to $P(S)$ and the four $P(C|S)$ conditionals to 100. This results in sampling distributions that are close to a uniform distribution.

For $P(S|S)$, we construct two different sets of parameters. In the first version, we set all concentration parameters to 100, just like in the case of $P(S)$ and $P(C|S)$. In the second version, we create correlation between shapes by multiplying the α that controls the likelihood of a shape given the same shape by 10. This leads distribution samples where the co-occurrence of the same shape within a game becomes likelier than random chance. For example, a circle target shape will more often appear with other circle distractors. A speaker sending unambiguous messages, will have to more often use color as the distinguishing feature, and the learners will see less evidence of what shape-words mean.

For training, we sample only one instance of $P(S)$, $P(S|S)$ and $P(C|S)$. At test time, we sample different $P(S)$, $P(S|S)$ and $P(C|S)$ instances 10 times. From each of these constellations we sample 3,200 games. This results in a test set of 32,000 examples.

The random seed is fixed across all experiments and is reset for the learning and evaluation of each learner. This ensures that each listener sees the exact same examples in all environments.

Model training and implementation All model-parameters are trained from scratch. Weights are updated with the AdamW optimizer (Loshchilov and Hutter, 2017) which we initalize with a learning rate of $1e-5$.

For each training step, we use a batch of 32 games and the listeners are trained for 25,920 training steps. We train one instance of each listener. The implementation code will be publicly released upon acceptance of the paper.

5.3 Results

In this section, we present the most insightful results from our extensive set of experiments. First, we look into how listeners with different depths behave during training and which parameters of the environment pose the biggest challenges for them.

Next, we turn to the main interest of this paper which is contrasting models that apply reasoning during training with ones that upgrade a base model to higher levels during deployment. For clarity, we call the first group **reasoning learners** and the second group **upgraded listeners**.

We report accuracy for all results, which simply shows the proportion of games where the listener guessed the right image based on the speaker's message. We perform Fisher's exact test for significance testing. For the sake of brevity we do not report significance results for all pairings, but only for the ones most relevant to the research question. We note $p < 0.05$ with one asterisk * and for $p < 0.01$ we put ** next to the results.

Learning dynamics We first examine the learning dynamics of reasoning learners with different depth of recursive reasoning. Figure 3 shows that when we keep all parameters of the learning environment constant, and only vary the listener's depth, we observe that listeners with higher levels, learn to perform the task with good accuracy faster. The gap in performance is especially large in the initial learning stages. This suggests that applying pragmatic reasoning might be particularly useful in case of limited data.

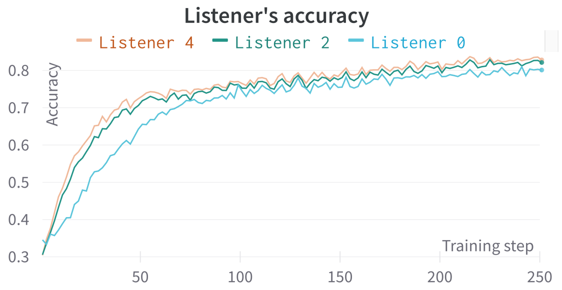


Figure 3: Higher level listeners learn quicker. In this comparison all other parameters such as speaker level, number of distractors, correlation between shapes are left constant.

Here we show results for a level 1 speaker with $N = 2$ distractor images and uniform distributions over all image features. These results however, hold across all training scenarios.

	Listener	Easy	Hard
a)	0	91.9	80.8
b)	2	93.7**	84.7**
c)	4	94.1**	85.4**

Table 1: Accuracy in the easiest and hardest environments. The easy environment has no messaging cost, each game has N=2 distractor images and the same shapes co-occur frequently. In the hard environment N=4, each word has a cost 0.6 for the speaker, the image features are uniformly distributed. In both environments higher level reasoning learners perform better. The speaker is S_1 for training and evaluation.

Easiest and most challenging learning environments We now present results of the models' performance after the reasoning learners finished training. We would like to see what parameters in the environment make learning more challenging. In this set of evaluations, we keep the parameters of the environments the same across training and testing. Table 5.3 shows the accuracy of the reasoning learners of three different levels in the least and most challenging settings.

The cost of messaging has the biggest impact on accuracy, followed by the number of distractors. The fewer distractors the environments have, the easier the task becomes. Having correlation between the the target and distractor shapes also makes the task easier, as in these cases the speaker uses messages about the color more frequently and the learners can achieve a high accuracy with only understanding the color terms.

	Listener	Speaker train	Accuracy
a)	0	1	80.8**
b)		3	79.2
c)	2	1	84.7**
d)		3	83.4
e)	4	1	85.4**
f)		3	84.1

Table 2: For each level of reasoning learner, learning from a lower level listener results in significantly better accuracy. Here we have a cost of 0.6, and a S_1 speaker at evaluation, $N = 3$. The same trends hold for a S_3 at evaluation time.

Learning from speakers with different depth The level of the speaker who the listener learns from is also of interest. S_1 has an internal model of a competent L_0, while S_3 anticipates L_2-behavior. We would like to see how reasoning learners of different levels are impacted by learning from different speakers.

Table 2 shows that reasoning learners that learned from lower level speakers always achieve higher accuracy at evaluation. This can be explained by the fact that lower level speakers send longer messages on average, see Table 3, because their internal model is of a simpler listener who needs longer descriptions for success. Higher level speakers on the other hand, model higher level listeners who can successfully resolve shorter messages. As communication has a cost in this setup, speakers of higher levels will resort to shorter communication if possible.

Learners benefit from longer messages as this provides more data about both color and shape of the objects. This behaviour nicely aligns with the intuition that language learners benefit from simple, verbose communication and teachers should not assume challenging patterns of communicative competence early on in the learning process (Nguyen, 2022).

	Speaker	
Distractors	1	3
2	1.05	1.01
3	1.13	1.02
4	1.23	1.06

Table 3: Average message length over 5000 samples for different number of distractors and speaker levels. Higher level speakers send shorter messages and more distractors result in longer messages.

	Listener		
	Training	Evaluation	Accuracy
a)	0	0	85.4
b)	0	2	85.7
c)	2		87.6**
d)	0	4	85.8
e)	4		88.2**

Table 4: Reasoning learners perform significantly better than upgraded listeners of the same level at evaluation time. The environment parameters are $N = 2$, $cost = 0.6$ with S_1 speaker.

Accuracy impact of reasoning during leaning One of the main interests of this paper is to see the impact of applying pragmatic reasoning during learning as opposed to only upgrading a L_0 listener to higher levels during evaluation. In Table 4 row a) we show the performance the simplest 0-level listener. Then we compare two versions of each higher level listener: rows b) and d) show listeners that were only upgraded to higher levels during evaluation. The corresponding results in row c) and e) show listeners that applied the same level of reasoning already during training. Reasoning learners are significantly better than upgraded listeners of the same level at evaluation time.

These results confirm our hypothesis that it is worth making the extra effort to differentiate through pragmatic reasoning when learning the underlying literal representations. As most real-world examples of natural language use also involve pragmatic reasoning, this result is important to consider when designing artificial language learners in a conversational setting.

Reasoning learners are more robust We would also like to understand how robust our listeners are to environmental changes. As reasoning learners make an explicit effort to model how the context contributes to the observed distributions, we expect that these listeners will generalize better to out-of-domain data.

In order to test this idea, we take listeners that were trained in the easiest setting still including communication cost: two distractor images and frequent co-occurrences of the same shapes. Having correlated shapes results in the speaker using more messages about the color, which means that the learners observe less evidence about the meaning of shape-words.

We first evaluate in-domain, in the same environment, then we change the environment to the most challenging set of parameters. The out-of-domain evaluation set has 4 distractor images and the shapes come from a uniform distribution.

	Listener		Accuracy		
	train	eval	ID	OOD	Δ
a)	0	0	92	73.5	-18.5
b)	0	2	92.2	74.5	-17.7
c)	2		93.5	77.3	-16.2*
d)	0	4	92.1	74.8	-17.3
e)	4		93.9	78.3	-15.6 *

Table 5: Models evaluated in the easy in-domain (ID), and challenging out-of-domain (OOD) environment. The last column shows the difference in performance between the two environments for the same model.

In Table 5 we report the relative changes in accuracy between in- and out-of-domain data. Reasoning learners and upgraded listeners degrade in accuracy when they are evaluated in a more difficult environment. In the last column, we indicate the drop in performance for all models. We observe that the reasoning learners suffer significantly less from environmental changes.

As here we are interested if the difference in differences is significant, we perform binomial logistic regression as significance test, and find that reasoning learners are significantly more robust than upgraded listeners of the same level at evaluation time.

6 Discussion

A lot of the analysis presented in this paper was made possible by the fact that we used an artificial setup where the data-generating process is fully known. In most real world data sets the natural language data is not labelled with the level of hierarchical reasoning that the participants applied during production, and this factor remains a latent variable. Curriculum learning that would like to integrate the knowledge about speaker-depth would first have to infer this from the data.

We proposed to use stochastic gradient descent for learning language while integrating pragmatic reasoning. This was permitted by the fact the hierarchical reasoning process we presented is fully differentiable. The most unrealistic aspect of this process is the normalisation over all possible messages in the speaker's production. This assumes that listeners and speakers are able to consider all competing sentences during reasoning.

Obviously, the set of natural language sentences is not finite and an infinite number of sentences can be composed from a fixed vocabulary. This is a known problem in models that integrate counterfactual reasoning about alternative utterances. The most common resolution is sampling a set of relevant sentences from a proposal distribution (Andreas and Klein, 2016; Liu et al., 2023). This of course breaks differentiability. As a result, models using sampling often turn to reinforcement learning. Any model that integrates hierarchical reasoning during training for natural-language datasets, also needs to consider this problem.

7 Conclusions

For artificial agents to understand humans, it is critical for them to correctly interpret context. By recursively modeling the conversational partner, reasoning listeners can improve their interpretation in context. In this work we introduced artificial language learners that have an explicit model of the speaker's production process *during* learning.

We have shown that learners that reason about how the context influences the data they observe, learn to perform the task with high accuracy faster. They also learn literal utterance representations that lead to higher task success within the training domain and lead to better generalization when there are changes in the environment. Finally, we also provided evidence that speakers that assume high level reasoning from the learners actually are worse teachers. Overall, we conclude that the widely used practice of applying pragmatic reasoning during evaluation is not sufficient. Instead, it is beneficial to model pragmatic processes already during learning.

8 Limitations

While the conversational phenomena we model in this paper have been widely attested to in linguistic theory and psycho-linguistic research, our experiments are limited to an artificial sandbox scenario with a small vocabulary and simple observations. As discussed in Section 6 reasoning about all possible utterances is intractable with larger vocabularies.

Real world conversations contain a wide range pragmatic inferences, not all of which can be accounted for by the recursive reasoning presented in this paper.

References

Jacob Andreas and Dan Klein. 2016. Reasoning about pragmatics with neural listeners and speakers. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1173–1182, Austin, Texas. Association for Computational Linguistics.

Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar. Association for Computational Linguistics.

Reuben Cohn-Gordon and Noah Goodman. 2019. Lost in machine translation: A method to reduce meaning loss. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 437–441, Minneapolis, Minnesota. Association for Computational Linguistics.

Reuben Cohn-Gordon, Noah D. Goodman, and Christopher Potts. 2018. Pragmatically informative image captioning with character-level inference. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 2 (Short Papers)*, pages 439–443. Association for Computational Linguistics.

Catherine Davies, Vincent Porretta, Kremena Koleva, and Ekaterini Klepousniotou. 2022. Speaker-specific cues influence semantic disambiguation. *Journal of Psycholinguistic Research*, 51(5):933–955.

Danny Fox and Roni Katzir. 2011. On the characterization of alternatives. *Natural language semantics*, 19:87–107.

Michael C. Frank and Noah D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science*, 336(6084):998–998.

Daniel Fried, Jacob Andreas, and Dan Klein. 2018a. Unified pragmatic models for generating and following instructions. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1951–1963, New Orleans, Louisiana. Association for Computational Linguistics.

Daniel Fried, Ronghang Hu, Volkan Cirik, Anna Rohrbach, Jacob Andreas, Louis-Philippe Morency, Taylor Berg-Kirkpatrick, Kate Saenko, Dan Klein, and Trevor Darrell. 2018b. Speaker-follower models for vision-and-language navigation. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc.

Dave Golland, Percy Liang, and Dan Klein. 2010. A game-theoretic approach to generating spatial descriptions. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 410–419, Cambridge, MA. Association for Computational Linguistics.

Noah D Goodman and Michael C Frank. 2016. Pragmatic language interpretation as probabilistic inference. *Trends in cognitive sciences*, 20(11):818–829.

Herbert P Grice. 1975. Logic and conversation. In *Speech acts*, pages 41–58. Brill.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Laurence Horn. 1984. Toward a new taxonomy for pragmatic inference: Q-based and r-based implicature. In *Meaning, form, and use in context: Linguistic applications*, pages 11–42. Georgetown University Press.

Alexander Kuhnle and Ann A. Copestake. 2017. Shapeworld - a new test methodology for multimodal language understanding. *ArXiv*, abs/1704.04517.

Angeliki Lazaridou, Anna Potapenko, and Olivier Tieleman. 2020. Multi-agent communication meets natural language: Synergies between functional and structural language learning. *arXiv preprint arXiv:2005.07064*.

Sang-Woo Lee, Yu-Jung Heo, and Byoung-Tak Zhang. 2018. Answerer in questioner's mind: Information theoretic approach to goal-oriented visual dialog. *Advances in neural information processing systems*, 31.

Stephen C Levinson. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. MIT press.

David Kellogg Lewis. 1969. *Convention: A Philosophical Study*. Cambridge, MA, USA: Wiley-Blackwell.

Andy Liu, Hao Zhu, Emmy Liu, Yonatan Bisk, and Graham Neubig. 2023. Computational language acquisition with theory of mind. *arXiv preprint arXiv:2303.01502*.

Ilya Loshchilov and Frank Hutter. 2017. Decoupled weight decay regularization. In *International Conference on Learning Representations*.

Junhua Mao, Jonathan Huang, Alexander Toshev, Oana Camburu, Alan L Yuille, and Kevin Murphy. 2016. Generation and comprehension of unambiguous object descriptions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 11–20.

Will Monroe, Robert XD Hawkins, Noah D Goodman, and Christopher Potts. 2017. Colors in context: A pragmatic neural model for grounded language understanding. *Transactions of the Association for Computational Linguistics*, 5:325–338.

Will Monroe and Christopher Potts. 2015. Learning in the rational speech acts model. *CoRR*, abs/1510.06807.

Minh Thi Thuy Nguyen. 2022. Interlanguage pragmatics as communicative competence. chapter 8, pages 135–151. Taylor & Francis.

Allen Nie, Reuben Cohn-Gordon, and Christopher Potts. 2020. Pragmatic issue-sensitive image captioning. *arXiv preprint arXiv:2004.14451*.

Nathaniel J Smith, Noah Goodman, and Michael Frank. 2013. Learning and using language via recursive pragmatic reasoning about other agents. In *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc.

Ramakrishna Vedantam, Samy Bengio, Kevin Murphy, Devi Parikh, and Gal Chechik. 2017. Context-aware captions from context-agnostic supervision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 251–260.

Adam Vogel, Christopher Potts, and Dan Jurafsky. 2013. Implicatures and nested beliefs in approximate decentralized-pomdps. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 74–80.

Rose E. Wang, Julia White, Jesse Mu, and Noah D. Goodman. 2021. Calibrate your listeners! robust communication-based training for pragmatic speakers. In *Findings of the Association for Computational Linguistics: EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 16-20 November, 2021*, pages 977–984. Association for Computational Linguistics.

Julia White, Jesse Mu, and Noah D. Goodman. 2020. Learning to refer informatively by amortizing pragmatic reasoning. In *Proceedings of the 42th Annual Meeting of the Cognitive Science Society - Developing a Mind: Learning in Humans, Animals, and Machines, CogSci 2020, virtual, July 29 - August 1, 2020*. cognitivesciencesociety.org.