**Introduction**

Human communication is rife with figurative language, ranging from hyperbole, to irony, to metaphor. How do people go beyond the literal meaning of an utterance to infer the speaker's intended meaning? In this paper, we model people's interpretations of nonliteral language using an extended version of Rational Speech Act (RSA) models, a family of computational models that formalize language understanding as recursive reasoning between speaker and listener. Using a series of behavioral experiments, we show that an RSA model extended to incorporate uncertainty about the question under discussion (QUD) is able to predict people's interpretations of hyperbole, irony, and metaphor. We argue that despite apparent differences among these subtypes of figurative language, the same computational framework can flexibly produce fine-grained interpretations for a range of nonliteral uses. We use this as evidence suggesting that the rich and often affectively-laden meanings expressed by figurative language can be explained by basic principles of communication.

From ``Juliet is the sun" to ``That woman is a bombshell," nonliteral language is, quite literally, everywhere. Metaphor, hyperbole, and sarcasm are ubiquitous in human communication, often creating poetic or humorous effects that add rich dimensions to language \citep{glucksberg2001understanding, pilkington2000poetic, lakoff2009more, roberts1994people}.

%People often use language to convey information that goes far beyond a sentence's literal meaning. Hyperbolic statements communicate speakers' attitudes and opinions by exaggerating the truth (``It took a million years to write this paper"); ironic or sarcastic statements communicate meanings contrary to what is encoded in the literal semantics (``Paper-writing is my favorite activity"); and metaphors highlight hidden similarities between distinct categories by equating them (``Paper-writing is a marathon").

While figurative statements are often false under their literal meanings, people are highly adept at inferring relevant and true information from these utterances. How do our linguistic, cognitive, and social faculties work together to allow us to fluently and accurately understand figurative language?

An ocean of ink has been spilled on this topic across many areas, including psychology, linguistics, philosophy, computer science, and literary theory \citep{glucksberg2001understanding, papafragou1996figurative, li2010using, kreuz1993empirical}. Some researchers focus on the cognitive mechanisms that enable particular types of figurative language, such as the process of aligning shared properties and analogous relations to understand metaphor \citep{gentner1997alignment, bowdle2005career, glucksberg2003psycholinguistics}. Others focus on the communicative principles that guide interpretation, such as using conversational maxims to select the appropriate meaning of an utterance \citep{grice20134, searle1979metaphor, sperber2008deflationary, ortony1993metaphor, tendahl2008complementary}. In my dissertation, I will adopt an approach closer to the latter, with the goal of proposing a general pragmatic framework that explains the basis of figurative communication.

%%%

%\section{What is figurative language?}

\section{Background}

%However, most theories of figurative language are informal and not fully precise, leaving room for discrepancies in how researchers interpret empirical data. %TODO: references.

%Furthermore, different theories focus on disparate aspects of figurative language (e.g. the standard pragmatic view focuses on recognition; the structure-alignment model focuses on interpretation of relational metaphors, etc.), making it difficult to compare predictions about the same phenomenon. Figurative language understanding involves many moving pieces---literal meaning, common ground, contextual information, speaker intention, feature attribution, subjective attitude and affect, etc. A model that explicitly describes how all of the pieces interact may help resolve some of the discrepancies among theories and provide a unified framework for examining various types of figurative language. In particular, several important questions are left unanswered by the theories reviewed above. I will explore a few of the questions in this section and introduce a formal computational framework in Chapter 3 that may shed light on the answers.

In this section, I will first review some classic ideas in pragmatics. I will then describe two main bodies of work that specifically examine the pragmatics of figurative language understanding. Finally, I will suggest that a complete theory of figurative communication should more carefully account for certain important factors that affect general language understanding.

\subsection{Language in general}

One of the most important insights in pragmatics is that listeners tend to assume speakers to be rational and cooperative agents who aim to be informative, known as the Cooperative Principle \citep{grice20134, clark1996using, levinson2000presumptive}. When interpreting an utterance, a listener uses these assumptions of rationality and informativeness to reason about what meaning a speaker could want to convey that would lead him to choose a particular utterance.

This reasoning between listener and speaker is responsible for many phenomena in pragmatics and language understanding, such as various types of conversational implicatures \citep{horn2006implicature, levinson2000presumptive}.

%A listener who reasons about a speaker can make many more powerful inferences about an utterance by thinking about \emph{why} he chose a particular utterance and what communicative goal it satisfies \citep{gibbs2012interpreting}.

%It seems plausible that listeners also need to reason about the speakers' intent in order to interpret figurative utterances in various communicative contexts.

When speakers and listeners reason about each other to communicate, they also consider and make use of their shared background knowledge \citep{clark1996using}. Suppose Liz asks, ``Is Bob an honest person?" and Sam replies, ``He's a politician." Although Sam did not directly answer Liz's question, it is likely he means ``no."

%His utterance not only asserts Bob's profession (the literal, dictionary meaning of ``politician"), but also attributes features associated with that profession to Bob (e.g. dishonesty, corruption). An encyclopedic approach to meaning proposes that the meaning of a word includes a network of background knowledge shared among people in a community, often reflecting stereotypes, conventions, and the community's beliefs and practices \citep{taylor2003linguistic, langacker1987foundations}. While the dictionary meaning of ``politician" is ``a person who is professionally involved in politics, especially as a holder of or a candidate for an elected office," the encyclopedic meaning of the word can encompass many more features and connotations.

Liz is able to successfully interpret Sam's utterance, and Sam is able to successfully use this utterance, because they both have access to the relevant \emph{encyclopedic meaning} of ``politician"---the network of background knowledge shared among people in a community, which includes the stereotype that politicians may be dishonest or corrupt \citep{taylor2003linguistic, langacker1987foundations}. Naturally, Liz's interpretation is sensitive to the contents of the background knowledge they share. If Liz and Sam belong to a community where all politicians are believed to be honest, then Liz would interpret Sam's reply to mean that yes, Bob is an honest person.

Besides background knowledge, Liz's interpretation also relies on her prior beliefs about specific aspects of the state of the world, in this case Bob's profession. Suppose prior to her exchange with Sam, Liz did not know what Bob does for a living. She will have learned two facts about Bob from Sam's utterance: Bob is a politician, and Bob is not an honest person. Suppose Liz knew \emph{a priori} that Bob is a politician. She will not have learned anything new about Bob's profession from Sam; however, even though she already knows that politicians in general are believed to be dishonest, Sam's utterance makes her more certain that Bob \emph{in particular} is dishonest, because that is the most informative and relevant meaning given her initial question about Bob's honesty. Finally, suppose Liz knows that Bob is a concert pianist and not professionally involved in politics at all. How will Liz interpret Sam's utterance? Instead of updating her beliefs about Bob using the dictionary meaning of ``politician," she will rely on the encyclopedic meaning to conclude that Bob is dishonest (but still a concert pianist). In other words, she will interpret the utterance non-literally. These examples show that interpretation of the same utterance in the same local context can vary in a rich and graded manner based on the speaker and listener's prior beliefs.

Listener's interpretations are also often sensitive to the local context. Suppose instead of inquiring about Bob's honesty, Liz had asked, ``Is Bob a persuasive speaker?" Sam's utterance may now be interpreted as a compliment: Bob is indeed a persuasive speaker. The local context---whether constructed by an explicit question, a situation, or a salient goal---sets up a ``question under discussion" (QUD) to which the information conveyed by the speaker is intended to be maximally relevant \citep{roberts1996information}.

\subsection{Figurative language in particular}

Two main bodies of work examine the pragmatic principles governing figurative language understanding. The first is the standard pragmatic view, which analyzes figurative utterances using standard Gricean Maxims \citep{grice20134, searle1979metaphor}. This view proposes a three-step process to understand figurative utterances: (1) determine the literal meaning of the utterance (2) determine whether the literal meaning violates the quality maxim (3) reanalyze the utterance to identify implied or metaphorical meanings that would allow the utterance to adhere to the conversational maxims. While the standard pragmatic view is appealing in that it fits naturally within Grice's general theory of communication, it has met with several criticisms. For one, many figurative statements do not violate the quality maxim because their literal meanings are also true. ``No man is an island," for example, is a literally true statement in addition to a figurative one \citep{gibbs1992metaphor}. By relying on the violation of the quality maxim to arrive at a figurative interpretation, the standard pragmatic view does not provide a satisfying explanation for these types of utterances. Another common criticism of the standard pragmatic view is that it requires the listener to first access the literal meaning of the utterance, verify that it is contextually inappropriate, compute potential figurative interpretations, and then select the interpretation that best satisfies conversational maxims. Given the number of extra steps involved, a figurative utterance should take longer to understand than a literal utterance. However, many experiments have shown that the figurative meanings of utterances are often accessed as quickly or even more quickly than their literal meanings given certain supporting contexts \citep{glucksberg2003psycholinguistics, gildea1983understanding, gibbs1992metaphor}. These empirical results suggest that literal meanings do not have to be explicitly rejected to make way for figurative interpretations.

Instead of appealing to the maxim of quality as the guiding principle for interpreting figurative utterances, relevance theory proposes that the principle of relevance is key to explaining a range of phenomena in communication, including figurative language

\citep{sperber1986relevance}. Relevance theorists view figurative language as employing the same comprehension processes as literal language---to maximize the relevance of the interpretation to the speaker's communicative goal \citep{sperber2008deflationary}. Suppose Liz is a chemist who needs to make a solution using water that is exactly $100$ degrees celsius. She asks, ``How's the temperature?" and Sam replies, ``The water is boiling." Liz will interpret this to mean that the water is at boiling point. Now suppose Liz is about to give her baby a bath and asks her husband Sam to test the temperature. Sam replies, ``The water is boiling." Liz will interpret this to mean that the water is much too hot, but quite unlikely that it is $100$ degrees celsius. Now consider the utterance: ``Bob is boiling." If what is relevant is Bob's temperature, this utterance means that Bob has a fever. If the topic is Bob's emotional state, this means that Bob is angry. A relevance theory account of metaphor is that all meanings, not just figurative ones, are selected based on relevance to the question under discussion.

While both the standard pragmatic and relevance theoretic view provide useful frameworks for understanding the pragmatics of figurative communication, certain factors that are important in general communication remain overlooked. For example, neither the standard pragmatic nor relevance theoretic view explicitly take into account the speaker's rationality or desire to be informative. There is also little consideration of the shared encyclopedic knowledge associated with different utterances, or representation of the listener's prior beliefs. While relevance theory considers which interpretations are maximally relevant to the question under discussion, they do not provide a clear operationalization of relevance. We believe that an adequate model of figurative language understanding should flexibly integrate these components to determine an appropriate interpretation. In addition, figurative language is often used to express subjective experiences and emotional attitudes \citep{riloff2005exploiting, roberts1994people}.

Since these emotional subtexts contribute greatly to the appeal of figurative language, it is important for a theory of figurative language understanding to include analyses of these effects.

\section{A proposal}

In my dissertation, I will propose a formal modeling framework for figurative language understanding that includes the following components: reasoning about the speaker's choice of utterance, with assumptions of rationality and informativeness; the literal meaning of utterances; shared background knowledge between speaker and listener; specific prior beliefs; local contextual information; affective subtext. While many researchers have suggested that the construction of meaning involves an interplay of these components \citep{coulson2005blending, gibbs1984literal, clark1996using, stalnaker2002common}, to our knowledge there is no formal model that explicitly describes the relationships among these components and integrates them to produce concrete, fine-grained predictions that can be evaluated against empirical data. In the next chapter, I will describe such a formal model and argue that it may illuminate our understanding figurative communication.