Research Statement

I am a computational linguist who studies language using a variety of techniques: experimental methods, statistical modeling, machine learning, and formal methods. My research focuses on context-sensitive expressions: words whose meanings change depending on who is speaking and in what context. More broadly, I am interested in how conversation participants use knowledge about each other's mental states. My research investigates questions such as: *Do speakers accommodate their listeners' visual perspectives on the scene?* and *Do listeners use their knowledge of a speaker's opinions to better interpret sentences?* Exploring these questions can help us understand general cognitive behavior such as theory of mind.

My dissertation focuses on natural language expressions that are perspectival: their meaning depends on the point-of-view adopted by the speaker. For instance, motion verbs like *come* describe motion towards a perspective-holder, who could be the speaker, listener, or the subject of an attitude verb like *think*. For example, suppose that Sam Speaker in Seattle calls his friend Lucy Listener, in London. If Sam says, *Thelma is coming*, it can mean two different things. If he is using his own perspective, the destination is Seattle. If he adopts Lucy's perspective, the destination is London. For Lucy to understand Sam, she needs to infer whose perspective he is using.

In addition to being a core component of human cognition, perspective represents a key challenge for artificial intelligence. Perspectival aspects of meaning are some of the most challenging for artificial intelligence to grasp, because they are situational, grounded, and interactive. I seek to model how people use and understand perspectival expressions in situated interactions using a variety of methodologies: simulations of reasoning processes implemented using probabilistic programming languages; crowdsourced linguistic production and comprehension experiments; and type-theoretic analyses of the semantics of perspectival expressions.

1 Computational Modeling of Perspective

A major goal of my research is to understand how conversation participants reason about the mental state of the speaker. Prior work on perspectival expressions have relied on simple heuristics, such as a speaker-default. These approaches are limited because they cannot capture how contextual factors affect interpretation. By contrast, I have developed a probabilistic, context-based model of perspectival reasoning and have verified the predictions of the model experimentally.

I model perspectival reasoning as a Bayesian inference process based on the Rational Speech Acts (RSA) framework. The RSA framework, one of the most important recent advances in pragmatics, applies Bayesian reasoning to sentence production and comprehension. In this framework, listeners infer a probability distribution over a set of worlds representing possible meanings based on a mental model of how the speaker selects an utterance. My contribution is to extend the RSA

Figure 1: Perspectival Rational Speech Acts model

Literal Listener

$$p(w|m, a) \quad \propto [[m]]^{a,w} p(w) p(a)$$

Pragmatic Speaker

$$p(m, a|w) \quad \propto \text{softmax}(p(w|m, a) \textstyle\sum_w [[m]]^{a,w} p(a) - \text{Cost}_m(m) - \text{Cost}_a(a))$$

Pragmatic Listener

$$p(w, a|m) \quad \propto p(m, a|w) p(w)$$

where a is a perspective, w is a world, and m is an utterance model so that listeners reason jointly over pairs of meanings and perspectives.

My model has three key properties. (1) Motivated by findings in psychology, the model's cost functions bias the speaker towards simpler utterances even if they are less precise (Cost_m) and towards their own perspective (Cost_a). (2) Since the model is implemented in a probabilistic programming language (WebPPL), it can be run backwards to fit the parameter settings on human behavorial data. (3) The model can also run forwards to generate empirically testable predictions. I discuss the empirical validation of the model in Section 2.

The model described above was published in a SCiL proceedings paper in 2019.

2 Experimental Validation

I conduct crowdsourced experiments in order to verify the predictions generated by my computational models of language. Computational models like the one I propose in Section 1 are useful for two reasons: they can be trained on behavioral data in order to find empirically validated parameter settings, and they be used to generate experimentally verifiable predictions. I have carried out web-based experiments to gather both kinds of data to support my model. The experimental work discussed below will be presented orally at the 2019 RAILS conference.

2.1 Comprehension Experiment

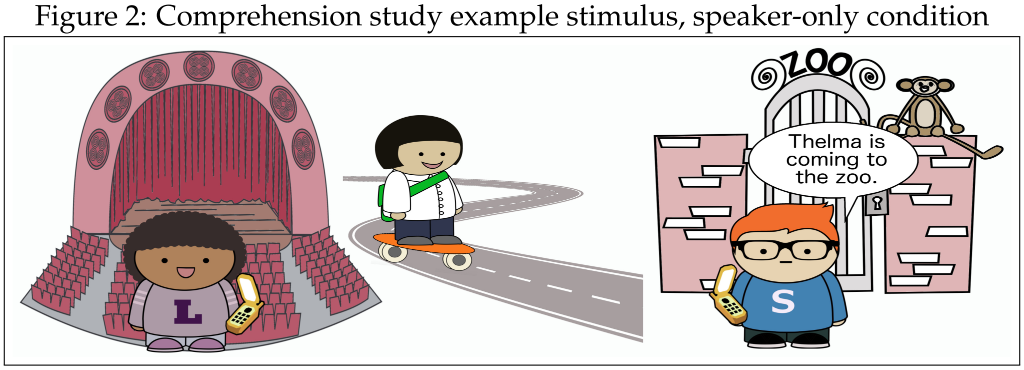
A central prediction of my probabilistic model of perspectival reasoning is that listeners entertain multiple perspectives simultaneously. Prior work used simple heuristics, like assuming that the perspective-holder is the speaker by default. Is the probabilistic version actually necessary? I have verified the multiple perspectives model in a comprehension experiment.

My multiple perspectives approach predicts highest marginal posterior probability for the world with both possible perspective holders at the destination of motion, while the heuristic hypothesis predicts equal probability for worlds where the speaker is a valid perspective holder.

I tested this prediction in a comprehension experiment. Participants were asked to imagine themselves as Lucy Listener. They then saw Sam speak a sentence with a perspectival verb (*Thelma is coming to the zoo*) or manner-of-motion verb (*Thelma is walking to the zoo*), followed by a scene depicting (1) both the speaker and listener at the destination of motion; (2) just the speaker; (3) just the listener; or (4) neither. They then indicated whether the scene and sentence matched. The linking hypothesis is that participants' reaction times are faster for more probable worlds.

My results validate the predictions of the multiple hypotheses model: reaction times decreased for sentences that use *come* relative to sentences that use *walk* for the scenes with both perspective-holders at the destination, but increased for all other scenes. A mixed-effects regression model revealed a significant interaction between the both-scene and *come* condition. This

Figure 2: Comprehension study example stimulus, speaker-only condition



validates a central prediction of the Bayesian perspectival reasoning model, and shows that the multiple perspectives approach is superior to the heuristic-based approaches of prior work.

2.2 Production Experiment

My model has several parameters that determine the prior probability of the speaker selecting certain utterances and perspectives. The prior comprehension experiment validated the model for a wide range of parameter values. However, I have also run a production study to gather data to fit the parameter values. Using the same visual stimuli, I asked participants to complete the sentence spoken by the speaker ("Thelma is ... "). The responses were coded based on linguistic features of interest, such as whether a perspectival motion verb was used.

I ran this experiment in order to gather empirical data to use for learning parameter settings for the computational model. However, the results of the experiment provided an additional bonus: they revealed an unexpected effect in the listener condition. In scenes where only the listener is at the destination, participants avoided using either *come* or *go*. This is unexpected under two previous analyses of the lexical semantics of *go*. However, it is predicted by the computational model if the listener is uncertain about which semantics for *go* the speaker has in their grammar.

This experience illustrates the value of production studies: although they are time-consuming since participant responses are open-ended and must be coded, their results are often useful for answering multiple research questions.

3 Probing Neural Networks and Grounded Expressions

I am also interested in what context-dependent expressions, like perspectival motion verbs, can tell us about the learning capabilities of deep learning models. I am currently working on a probe task to assess the ability of text-trained neural networks to learn grounded linguistic terms.

State-of-the-art deep learning approaches for many NLP tasks rely on large amounts of raw text data which is, by nature, ungrounded. Children have access to **grounded** information when acquiring language: situational, perceptual information. Human use of grounded information could be an artifact of human cognition; however, since such information is always available to humans, our languages might have evolved to depend on it. Grounded linguistic phenomena therefore are important test cases for whether there are limits to what kinds of linguistic knowledge can be acquired from textual data alone.

This project explores what the motion verbs *come* and *go*, which rely on spatial perspective, can tell us about the ability of artificial learners to acquire grounded language understanding from text. I have compiled a diverse dataset of naturally occurring corpus examples; expert-selected corpus examples, annotated for linguistic features; and examples from linguistic analyses.

I will test the ability of context-sensitive language models such as BERT and GPT-2 to predict whether *come* or *go* is more likely to occur. I will then gather human judgments on the annotated examples and the hardest automatically extracted examples. I will compare the human judgments against those of the deep learning models and explore whether they vary in systemic ways using the annotated linguistic features of the manually selected examples.

4 Formal and Experimental Methods Applied to Temporal Perspective

In addition to modeling meaning in conversation, I have also developed new analyses of particular linguistic expressions. As a semanticist with extensive training in formal methods, I develop type-theoretic denotational semantics for linguistic expressions. A particular strength of my work is that I apply high quality experimental methods to validate the semantic analyses that I propose. This has helped me explore pockets of variation in the semantics of American English that have been overlooked by previous intuition-based accounts.

I have developed a novel analysis of the temporal adverbial *tomorrow* as a perspectival expression. I have gathered empirical evidence from a series of crowdsourced comic-captioning tasks in support of my analysis, showing that for a population of American English speakers, *tomorrow* has non-utterance time uses that cannot be explained under previous theories.

For example, consider the sentence *Last week, Athena said that she would water my plants tomorrow, but she never did.* Does *tomorrow* mean the day after the sentence was said, or the day after Athena promised to water the plants? Prior semantic theory predicts that only the former interpretation should exist. However, my experiments show that is a population of American English speakers who find both interpretations acceptable. I put forward a perspectival analysis of the denotational semantics of *tomorrow* for these speakers.

My experimental evidence comes from a series of comic-captioning tasks. Participants were shown a comic like Fig. 3 and were instructed to judge the sentence as a caption for the last panel.

Figure 3: Example stimulus for non-utterance time *tomorrow* experiments

Kevin is angry because Kate said that she would water his plants $\left\{\begin{array}{l}\text{tomorrow / the next day}\\ \text{Friday / Saturday}\end{array}\right\}$.

The results support the following claims: (1) that such non-utterance time interpretations are viable for a portion of American English speakers; (2) that they cannot be explained by indexical shift under Free Indirect Discourse; (3) that, unlike shifty indexicals in other languages, they do not require an embedding verb; and (4) that *tomorrow*, unlike true temporally anaphoric expressions like *the next day*, cannot appear in most quantificational binding contexts.

Based on these results, I have developed an analysis of *tomorrow* as a perspectival expression for the speakers of American English who accept these readings.

1. $[[\text{tomorrow}]]^{c,g} = \lambda Q.\lambda e.\tau(e) \subset \iota t.\text{DAY-AFTER}(\text{NOW}(p), t) \wedge Q(e)$
   where p is a prominent perspective-holder in the Common Ground

This work shows how experimental methods can drive the development of linguistic theory by (1) empirically testing the predictions of different analyses; (2) measuring gradient contextual effects (instead of binary truth judgments); and (3) uncovering patterns of interspeaker variation.

This work appeared in the proceedings of Sinn und Bedeutung and is under revision.

5 Automatic Speech Recognition

As a research intern in the speech technology group at BBN Technologies, I have worked on a number of projects for improving the language model component of speech recognition systems. My research has explored adding dialect ID tags to enhance multi-dialect language models; using recurrent neural networks to rescore the predictions of n-gram language models; training neural networks with a masked prediction training objective to directly rescore n-gram language model lattices; and applying neural machine translation techniques to correct the transcription of numerical terms in order to improve keyword search. I have also developed tools for statistical analysis and visualization of keyword search error rates.

6 Analysis and Documentation of Low-Resource Languages

I have significant expertise in low-resource languages of the Americas. As a Fulbright scholar, I worked on technology for the revitalization of First Nations languages. I also spent two summers conducting fieldwork on Western Tlacolula Valley Zapotec, an endangered language of Oaxaca, Mexico. Zapotecan languages are understudied, and my work involved documenting and analyzing its semantics, with a focus on perspectival motion verbs.

As well as documentation, I produced a formal semantic analysis of the **andative and venitive construction**. These constructions, which are similar to pseudo-coordination in English (*come and get*), are common cross-linguistically, but vary in subtle semantic ways. My analysis proposes that the events of the two verbs are unified via a non-Boolean conjunction that is applied after the external argument is projected but before it is saturated.

I presented this work at MVC 2017 and SSILA 2018, and the manuscript is in preparation. I also published work on negation in Valley Zapotec in *Transactions of the Philological Society*.

7 Future Work

Social knowledge My computational models of conversation improve upon previous theories of perspectival reasoning by taking into account the discourse context. In future work, I would like to incorporate non-linguistic contextual information. For example, speakers may defer to listeners of higher rank by adopting their perspective. Incorporating implicit, unstated domains of information, such as social dynamics and world knowledge, pose a significant challenge for current NLP techniques, because they are difficult to infer from text data.

Conversational adaptation Another challenging aspect of conversation to model is adaptation. The behavior of conversation participants is dynamic: listeners adapt their expectations based on speaker behavior and vice versa. In future work, I plan to model how conversation participants adapt in the face of interspeaker variation in perspective use. A trivial case of adaption is already captured in my model, in the form of the updated probability distribution over perspectives, but I would like to explore how social factors like trust and information asymmetry affect adaptation.

Stance detection My work also connects to stance detection. In the future, I would like to leverage psycholinguistic and cognitive science insights into theory-of-mind and perspectival reasoning to develop better inductive biases for stance detection systems.

MT for low resource languages Low resource languages pose a key challenge for modern data-intensive NLP techniques. In future work, I would like to work on improving neural machine translation approaches for very low-resource languages by making better use of information that is shared within language families. For instance, although there are few resources for any individual Zapotec language, the similarity of languages within the Zapotec family could be leveraged to make better use of these sparse resources.

ASR for low resource languages I also plan to apply my expertise in multilingual speech recognition to low resource languages. Recent research on speech-to-text for indigenous languages has focused on individual speaker models. These models, which are intended to aid transcription for language documentation and revitalization, can be trained on remarkably little data. I am interested in developing techniques to incorporate multilingual information into these models.