Research Statement

I am a semanticist who uses experimental, computational, and fieldwork methods to understand meaning. My research focuses on perspectival expressions: words whose meanings are dependent on a point-of-view, or perspective. More broadly, I seek to understand how conversation participants' knowledge about each others' beliefs (mental perspective) and location in time and space (spatiotemporal perspective) shapes discourse. My research explores 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?* These important questions are challenging to answer using traditional linguistic methods because they involve gradient judgments about interactive, dynamically evolving discourse contexts.

One of my strengths as a researcher is the breadth of methodological techniques that I can draw upon to address questions about meaning. My research contributions include a novel semantic analysis of *tomorrow* as a perspectival expression for certain American English speakers, supported by a series of experiments; a computational model of how conversation participants reason about perspectival expressions in which listeners reason jointly about perspectives and meanings; comprehension and production experiments on perspectival motion verbs that test the model's predictions; and an exploration of micro-variation in the semantics of complex motion verb constructions, based on fieldwork data from Western Tlacolula Valley Zapotec. Because my research connects to broader questions about cognition and theory-of-mind, I am also passionate about building connections between Linguistics and other cognitive sciences, such as Psychology and Computer Science.

1 Exploring the Semantics and Pragmatics of Perspective in English

My dissertation research focuses on several dimensions of the semantics and pragmatics of perspective-sensitive phenomena: the semantics of perspectival expressions; the way perspectival expressions are reasoned about in discourse; and the factors that affect how perspectives rise and fall in prominence as the discourse proceeds. I use different methodologies to explore these different aspects of perspective in language, ranging from production experiments to computational modeling.

1.1 Formal and Experimental Methods Applied to Temporal Perspective

One component of my dissertation focuses on variation in the semantics of *tomorrow* in American English. Although *tomorrow* has long been considered a pure indexical (Kaplan 1989), I argue that for a group of American English speakers, it is in fact anaphoric to the *now* of a perspective-holder.

My re-analysis of *tomorrow* as perspectival rather than context parameter-dependent is part of a line of recent work that has muddied the neat division between between time-anaphoric expressions and temporal indexicals: the discovery of shifty indexicals in many languages (Schlenker 2003), and the debate over discourse-anaphoric uses of *now* (Stojnić & Altshuler 2019, Anand & Toosarvandani 2019). My work also engages with the literature on special pragmatic contexts under which indexicals routinely interpreted under a context other than that of the speaker, such as Free Indirect Discourse (Banfield 1982) and Protagonist Projection (Stokke 2013).

My evidence comes from the existence of non-utterance time readings of *tomorrow*, as in (1). Here, the continuation (*but she never did*) forces an interpretation where *tomorrow* refers to the day after Athena made the promise, rather than the day after the sentence is spoken.

1. Last week, Athena said that she would water my plants tomorrow, but she never did.

Such readings are unexpected under a pure indexical view of *tomorrow*, but are available for some American English speakers. Because the non-utterance time readings are only accepted by some speakers, experimental methods are critical for studying the phenomenon. My experimental evidence comes from a series of comic-captioning tasks. Participants were shown a comic like Fig. 1 and were instructed to judge the sentence as a caption for the last panel.

Figure 1: 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 of my experiments 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.

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

I am currently developing a formal pragmatic account of the discourse factors that control how the perspectival variable p gets its value, drawing upon Centering theory (Grosz et al. 1995) and coherence relations (Hobbs 1979), as well as probabilistic discourse models like Kehler & Rohde (2012).

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; and (3) uncovering patterns of interspeaker variation. It also highlights the difficulty in distinguishing among different classes of context-sensitive expressions.

This work appeared in the proceedings of Sinn und Bedeutung and is being revised for *Semantics and Pragmatics*.

1.2 Computational Modeling of Perspective

Another portion of my dissertation focuses on how conversation participants reason about perspectival expressions. When a listener hears a perspectival expression, they must reason about the truth-conditional content of the utterance in tandem with the perspectival component. Understanding whose perspective the speaker is using can help them understand what the speaker is trying to say (and vice versa). In many circumstances, however, the speaker has optionality about whose perspective they adopt, which complicates the listener's reasoning process.

Understanding how listeners reason about perspective involves understanding not just what interpretations are possible, but also, what interpretations are likely. For this reason, I use compu-

Figure 2: Perspectival Rational Speech Acts model

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

tational and experimental methods to explore how listeners decide whose perspective the speaker has adopted. I have developed a probabilistic, context-based model of perspectival reasoning in the Rational Speech Acts (RSA) framework, which views sentence comprehension as a Bayesian inference process (Frank and Goodman 2012). 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.

I extend the RSA model so that listeners reason jointly over pairs of meanings and perspectives. Intuitively, understanding who the perspective-holder is helps the listener understand the speaker's message and vice versa.

For instance, if the listener hears 3, understanding the perspective that the speaker has adopted can help the listener reason about Thelma's intended destination.

3. Thelma is coming.
4. I am coming.

On the other hand, if the listener hears 4, the truth-conditions of the sentence eliminate the speaker from the set of possible perspective holders, since they cannot be both in motion and located at the destination of motion.

My model has three key properties. First, motivated by the literature of egocentricity in psychology (Epley et al. 2004), 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). Second, the model can be run forwards to generate empirically testable predictions both about the behavior of speakers and about listeners. Third, 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. I discuss experimental work testing some of the predictions of the model in Section 1.3.

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

1.3 Production and Comprehension Experiments

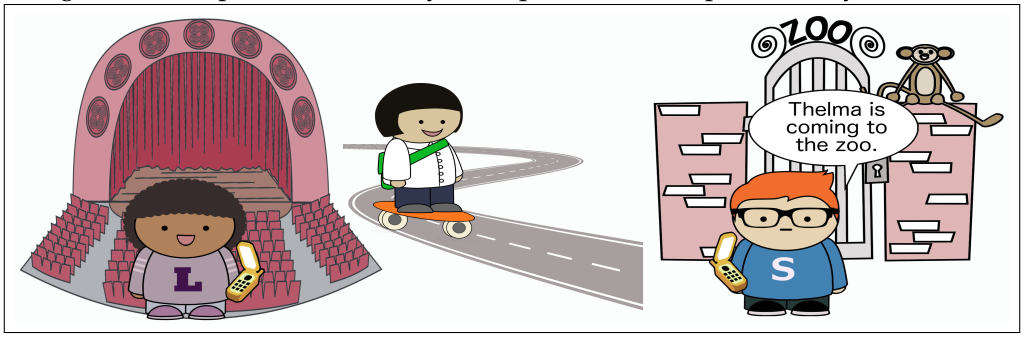
Computational models like the one I propose in Section 1.2 are useful because they can be used to generate experimentally falsifiable predictions. I have carried out comprehension and production experiments to test some of the predictions of the perspectival RSA model. This work will be presented at the 2019 Rational Approaches in Language Science conference.

1.3.1 Comprehension Experiment

A central prediction of my probabilistic model of perspectival reasoning is that listeners entertain multiple perspectives simultaneously. Prior work has posited that listeners assume that the perspective-holder is the speaker by default. This predicts that when listeners hear a sentence like (5), they should assign equal probability to all worlds in which the speaker is at the destination.

5. Thelma is coming to Northampton.

Figure 3: Comprehension study example stimulus, speaker-only condition



By contrast, the multiple perspectives approach that a world with both the speaker and listener at the destination of motion should be the most likely.

I tested this prediction in a comprehension experiment. Participants were told 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; (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 scenes that depict more probable meanings.

The results support 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 validates a central prediction of the Bayesian perspectival reasoning model, and suggests that listeners do consider multiple perspectives at once.

1.3.2 Production Experiment

The computational model that I propose has several parameters: the prior probabilities of a speaker selecting each utterance and perspective, as well as the cost functions. The multiple perspectives hypothesis discussed above is predicted for a wide range of parameter values. However, to test other kinds of predictions, we would like to have empirical evidence supporting the choice of parameter values.

I have run a production study to gather data on the likelihood of speakers selecting each utterance given each world. 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.

The data from this experiment can be used to estimate the parameter settings of the computational model. In addition, the results revealed an unexpected effect in the listener condition: in scenes where only the listener is at the destination, participants avoided using both *come* and *go*.

This is unexpected if the lexical semantics of *go* contain a perspectival component; it is also unexpected if the anti-perspectival implication of *go* arises through pragmatic competition (Sudo 2018). However, this effect is predicted by the computational model if the listener is uncertain about which semantics for *go* the speaker has in their grammar.

While my exploration of the data is still preliminary, 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.

2 Exploring Perspective Further Afield

My interest in perspective is long-running and includes two projects on perspective that are not part of my dissertation research.

2.1 Perspectival Motion Verbs in Western Tlacolula Valley Zapotec

In addition to my dissertation work on perspective in English, I have also studied perspectival expressions in Western Tlacolula Valley Zapotec. I have spent two summers conducting fieldwork on the semantics of perspectival motion verbs in San Lucas Quiaviní, a town in Oaxaca, Mexico.

I produced a formal semantic analysis of a complex motion verb construction known as 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.

6. **E-type Non-Boolean Conjunction**:
   Given two functions $f_{\langle e\langle \epsilon,t\rangle\rangle}$ and $g_{\langle e\langle \epsilon,t\rangle\rangle}$, Non-Boolean Conjunction produces a function $h_{\langle e\langle \epsilon,t\rangle\rangle}$:
   $f_{\langle e\langle \epsilon,t\rangle\rangle} \quad g_{\langle e\langle \epsilon,t\rangle\rangle} \quad \Rightarrow \quad \lambda x.\lambda e''.\exists e,e'[e'' = e \oplus e' \wedge f(x)(e) \wedge g(x)(e')]$

This allows the complex motion verb construction to be formed with both transitive and unaccusative second verbs.

I presented this work at MVC 2017 and SSILA 2018, and the manuscript is in preparation for submission to the *International Journal of American Linguistics*. I have also published work on negation in Valley Zapotec in *Transactions of the Philological Society*, and have presented on another aspect of Zapotecan motion verbs at the 2019 Texas Linguistics Circle.

2.2 Probing Neural Networks for Perspectival Sensitivity

I am also interested in perspective as a linguistic phenomenon that is particularly challenging for artificial intelligence. State-of-the-art approaches for many NLP tasks rely on large amounts of raw text data which is, by nature, ungrounded. Grounded linguistic phenomena therefore are important test cases for whether there are limits on the kinds of linguistic knowledge can be acquired from text alone. I am currently working on a project that uses perspectival motion verbs to probe how well text-trained neural language models can learn spatially grounded terms.

I have compiled a small corpus of examples of *come* and *go*, which I have annotated for linguistic features, that I will use to test the ability of neural language models 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.

3 Automatic Speech Recognition

I have also developed expertise in automatic speech recognition through my affiliation with the speech technology group at BBN Technologies. I have worked on a number of projects involving neural networks for speech recognition. My research has focused on ways of making better use of text data in speech recognition training. Some of the ideas I explored were 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. Much of my work at BBN also focused on techniques for low-resource languages.