Language allows humans to make infinite use of finite means. We can understand situations and ideas we have never encountered from structured messages composed of familiar units. We combine this with reasoning about other agents' communication strategies to understand their beliefs and intentions. This makes language a powerful means of acquiring knowledge about the world and agents in it that we can use to perform other tasks. I am interested in studying how this linguistic structure and communication can be used to learn about the world.

I am drawn to these problems based on my research experience so far. I studied the influence of domain-specific concepts on learning using program synthesis, and wanted to find more flexible ways of acquiring these concepts. I studied how pragmatic reasoning allowed making sense of ambiguous instruction in program synthesis, and saw how resolving ambiguity was crucial to learning from other agents. In an ongoing project, I am examining the processing of tasks specified in text by neural language models, and understanding the different ways linguistic meaning is used in an interaction. Pursuing a PhD at the intersection of **natural language processing** (NLP) and **machine learning** (ML) would provide opportunities to study how linguistic structure and communication can be modeled computationally, and how this can be used to acquire knowledge to solve other tasks.

Program synthesis for linguistic rules. In my first research project, I worked towards solving phonology puzzles from the Linguistics Olympiads. These puzzles involve inferring abstract rules given only a few examples from a language, and pose a challenging problem in learning from very little data. I approached this problem by using program synthesis to learn rules in the form of programs in a domain-specific language (DSL). Not having had any background in the field of programming languages, I taught myself the basics of program synthesis, and adapted algorithms to implement a synthesizer. I designed experiments to understand how access to linguistic concepts helped synthesis, and developed methods to analyze and visualize the results. Through this experimentation and analysis, I found that linguistic concepts led to solutions with fewer, more generalizable rules. This work led to a paper at the ***SIGMORPHON Workshop at ACL 2021*** [W1][1]. I followed up with a more specific study of phonological stress published at ***ICON 2021*** [C1].

Through this project, I became familiar with the research process. I learned how to approach a broad goal by identifying small, concrete sub-problems. I came to appreciate the importance of communicating my findings, and refined my ability to write a research paper. I also saw how having access to the right abstractions (here, from the DSL) could help with learning from very little data. However, specifying these abstractions manually is laborious, and as I came to see, not sufficiently flexible. This is where I see the potential to discover these abstractions from language, which humans use powerfully and flexibly to refer to complex concepts.

Pragmatic program synthesis. A central problem in program synthesis is choosing among multiple programs consistent with the given examples. To learn linguistic rules, I approached this by engineering a preference for linguistic abstractions into the DSL. However, a paper at NeurIPS 2020[2] cast synthesis as a reference game, and showed how principles of pragmatic communication could be used

---

[1]References point to entries in the *Publications* section of my CV

[2]Yewen Pu, Kevin Ellis, Marta Kryven, Josh Tenenbaum, and Armando Solar-Lezama. Program synthesis with pragmatic communication. In *Advances in Neural Information Processing Systems*, 2020.

to choose among programs. While their framing of the problem was insightful, I also saw that they used pragmatic reasoning only after synthesis, leading to an approach that scaled poorly.

I reached out to the authors of the paper, and led a collaborative effort to see if pragmatics could be used during synthesis, towards a more efficient and pragmatic synthesizer. We found that a simple decomposition of the program synthesis problem lent itself to pragmatic specification of intent, and made pragmatic inference over these specifications less expensive. We presented these findings as a talk at the *Meaning in Context Workshop at NeurIPS 2021* [P1].

This project gave me first-hand experience initiating and building a collaboration to improve upon prior work. Starting with an intuitive idea, I was able to discuss, refine, and iterate on it until we got to a formal solution. I also understood how much humans communicate by choosing not to specify something, and how modelling pragmatics can enable understanding such communication as well.

Representations of semantics in NLP models. In addition to providing directions for improving models, I also see the value of frameworks from linguistics and cognitive science in analyzing representations in NLP models. I had a chance to do this during my internship at MILA when I joined an effort to investigate language models trained to complete tasks or instructions specified in text. My attention was quickly drawn to the question of how these models represented a task that was specified in text. Inspired by recent work that used linear probing to examine language models for representations of worlds described in text, I turned to a framework of meaning called dynamic semantics to formalize this question. In this ongoing project, I am working with people with expertise in NLP and text-based games to design experiments in simple text domains, and studying the linguistics literature on how the meaning of tasks or instructions may be represented. Through this, I am gaining experience with research that synthesizes different forms of expertise in answering a question.

Beyond the specific question of how language models represent tasks, I am also interested in broader questions of how we look for linguistic ability in large, learned models. To ponder this question with a broader perspective, I started a reading group with colleagues at MILA to discuss perspectives from machine learning, NLP, and linguistics on what linguistic abilities we look for in models of language, and why these are useful to solving the tasks for which these models were designed.

At LTI, I am interested in working with **Prof. Daniel Fried**, **Prof. Graham Neubig**, and **Prof. Yonatan Bisk**. A central challenge in leveraging language to acquire knowledge, especially when interacting with other agents, is correctly interpreting the meaning in context. I saw in my own work on pragmatic program synthesis how using models of human communication can help resolve communicative ambiguity. I hope to work with Prof. Fried on ways to resolve linguistic ambiguity, and enable learning from what others convey in language. I am also interested in working with Prof. Neubig on modelling linguistic interactions, and creating better interfaces that allow machines to learn from language to perform tasks like program synthesis and question answering. I also hope to work with Prof. Bisk on modelling the semantics of the world, and how this can enable acquisition of world knowledge from language.

More broadly, LTI is home to faculty and students exploring the connection between language and computation in a variety of ways. I hope to engage actively with this range of perspectives on NLP research, and believe the vibrant environment makes CMU an excellent place for me to do my PhD.