When I first heard about the North American Computational Linguistics Olympiad (NACLO), in eighth grade, I knew next to nothing about linguistics. But I did love solving logic puzzles and other brainteasers, and that’s what the sample problems on the NACLO website reminded me of.

One of the first problems to catch my attention was called “Tenji karaoke.” The basic idea was to match Japanese words with their tenji (Japanese braille) representations. Of course, it would have been trivial had I already known tenji, but the point of NACLO problems is to figure them out without knowing the language, and that’s what makes them interesting. For instance, in this problem, I could use the given example of how “karaoke” is written in tenji to match up the remaining words correctly. To finish up the problem, I had to translate some tenji that contained patterns of dots I had not seen before, which meant figuring out how the system of tenji works on a deeper level. I won’t spoil the answer, but the system was simple and quite elegant!

Excited to see more problems of that sort, I signed up for the open round of the NACLO. (Although the contest is geared to high school students, middle school students can take the tests.) And so one day in January I left school for three hours to decipher a made-up slang, translate a message of peace from one alien tongue to another, and learn to count to 10 in Waorani. That year, the 150 top scorers were given the opportunity to move on to the invitational round—and I was overjoyed to be one of them. Our performance on the second round would determine who would get to represent the United States (or Canada, depending on citizenship) at the International Linguistics Olympiad, or ILO.

One day in March, I took time off again from school to participate in the invitational round. When the scores from the invitational round came in, the top 18 students were invited to train for the ILO. I placed 58th. I was pretty proud of myself, considering that I was one of the youngest people taking the test. Next year, I thought, I bet I can do better!

After that, I looked forward to taking the NACLO every year, in part to see if I could do better, but mostly because I found it fun to spend a few hours solving puzzles.

Making the Cut
Meanwhile, in large part due to the NACLO, I was paying more attention to linguistics. I discovered the works of Steven Pinker. I wondered about Google Translate, about how it works and how it could be fixed when it messes up. I didn’t really “study” for the competition, though; participating in NACLO was always just a matter of going in and figuring things out.

Fast forward to 2016. I had barely qualified for the invitational round and was feeling very grateful for the opportunity to participate in one last NACLO contest before graduating from high school. I went in determined just to have fun and enjoy the experience, but to my surprise, I ended up qualifying for the ILO! I had to check the list of scores three times before I could believe it.

My parents were somewhat less than thrilled at the prospect of sending me off to India, where the ILO would be held that year, but after seeing how excited I was to go, they were really supportive. For my birthday, they gave me some books on linguistics, including The Cambridge Encyclopedia of Language. “If you’re going to do this,” they said, “don’t do it half-heartedly. Prepare!”

And prepare I did. I read the books my parents gave me. I worked on practice problems. I attended weekly online training sessions with the other members of the U.S. and Canadian teams.

Practice with a Purpose
One practice introduced us to a practical application of linguistics in the real world. Some of our team leaders were participating in another competition called LoReHLT (short for Low Resource Human Language Technology) and decided to use the practice to show us what they were working on for this challenge. Basically,
the idea of the competition was to see who could develop the best language technology to help a country deal with a natural disaster. When people want to provide supplies to areas hit by a natural disaster, figuring out how much to send and where to send it can be difficult. However, if we had a computer program that could identify information about what was needed from the millions of texts and messages people send, we would have a better idea of how to allocate the resources. For some languages, such as English, a lot of progress has been made in creating computer programs that can extract information from human language. LoReHLT focuses specifically on languages for which such progress has not been made.

During that practice, we used the limited information we had been given (an incomplete Uzbek dictionary, some Uzbek sentences with their English translations, and a summary of Uzbek morphology) to generate information that programmers could use to train their computer programs. It was really cool, because we got a taste of how the skills we had been using for NACLO and ILO problems could be applied to an actual linguistics problem, which in turn could be applied to do useful things in the world.

During some practices, we learned about features of language, such as phonology, the study of sounds. But most practices were spent solving problems from past competitions. Unlike in biology or chemistry competitions, knowledge of specific content in linguistics competitions often isn’t as valuable as the ability to use whatever information you’re given to derive the most logical conclusions—and the best way to improve that is to practice. However, despite all my preparations, I felt unprepared. In practices, my teammates had seemed to understand instantly what took me a while to figure out, coming to conclusions before I could finish processing all the data. Maybe my placement on the team had been a mistake? Such doubts niggled at the back of my mind as the date of the competition neared.

During that practice, we used the limited information we had been given (an incomplete Uzbek dictionary, some Uzbek sentences with their English translations, and a summary of Uzbek morphology) to generate information that programmers could use to train their computer programs. It was really cool, because we got a taste of how the skills we had been using for NACLO and ILO problems could be applied to an actual linguistics problem, which in turn could be applied to do useful things in the world.

Taa Time

In July, I traveled to India for the international competition. I was nervous about representing my country and anxious about meeting my teammates in person for the first time, but the next few days quickly put my worries to rest. I got to know the U.S. and Canadian teams and spent many engaging hours playing card and word games with them. I got to explore India and play Mafia with people from all over the world. And as worried as I initially had been about
honorably representing the U.S., once I flipped to the first problem of the individual round, it was just like another NACLO. Certainly, the problems were more difficult than the average NACLO problem, but I approached them with the same mindset: These are interesting puzzles, and I want to solve them.

One of my favorite problems had us match a list words in Iatmul with their English translations. To figure it out, I had to look at the structure of the words and see how they related to each other. If I saw an Iatmul word that seemed to be made of two other Iatmul words on the list, I would look for an English word on the other list whose meaning was in some way made out of two other English words on the list. This type of problem is interesting because it asks you to look at the world in a new way, from the perspective of a different culture. For example, in Iatmul, the direct translation for “rifle” is “white people's spear,” which makes sense when you consider how the speakers of Iatmul, an ethnic group in Papua New Guinea, would have first encountered rifles.

Other problems asked us to match Luwian words with their hieroglyphs, develop a system to explain Jaqaru suffixes, and figure out how speakers of Aralle-Tabulahan talk about going between places.

Afterward came the team round, which was a new kind of challenge. I distinctly remember the scene: We were waiting in our room for the contest to begin when our proctor came in, pointed to the television in the room, and said, “Turn it on, and plug in this flash drive.” I exchanged a worried glance with my teammates. There had not been an audio component to the ILO since 2005, and we had not prepared for this at all.

After a 30-minute training session, we set about our actual task: to listen to 114 audio files of words in Taa, a language with clicks and tones and other interesting phonetic features, and match them with their transcriptions. It was taxing to say the least, but I’m proud of how well my team adjusted to this unexpected challenge. We kept each other focused when we heard yet another audio file that sounded, to our untrained ears, more like a groan than a word. Through this experience, I felt a lot closer to my teammates. After all, there’s nothing like shared frustration over Taa to bring people together.

Looking back, I’m so glad I decided to do the NACLO all those years ago. Without it, I might not have learned how exciting and multifaceted a field linguistics is. It’s not just about cataloguing words and their meanings in different languages, it’s about finding patterns and understanding phenomena, and then using what you’ve learned to impact the world.

Margarita Misirpashayeva is a freshman at MIT, currently majoring in the exciting field of Undecided. In her free time, she enjoys music, playing board and word games, and solving puzzles.