

Computational Linguist

“My supervisor at Stanford contacted me to see if I was interested in joining a start-up he was involved in. This company, YY Technologies, was using the English grammar we had developed to create a range of applications include automated customer service e-mail response, and they wanted to branch out into Japanese. I speak Japanese, and as one of approximately two people in the world at that time who both spoke Japanese and was trained in the style of grammar engineering used by the company, I could basically name my price. It was a delightful job, essentially teaching Japanese to a computer, and it gave me the credentials and connections I needed to land the job I have now.”

Emily, 33, Computational Linguist

A lot of people associate linguistics, which is the scientific study of language, with an academic career. Others think of a linguist as someone who is fluent in several languages and travels the world as a translator. One of the coolest careers I’ve heard about in years, computational linguistics, involves a slightly different take on this interesting field. In general, linguists investigate how knowledge of language is acquired, how it interacts with other mental processes, and how it varies depending on the person or the region. They study how the structure of language (such as sounds or phrases) can be represented, and how different components of language interact with each other (such as intonation and meaning). According to *Wikipedia.org*, computational linguistics is an interdisciplinary field dealing with the statistical and logical modeling of natural language. This career originated in the 1950s, when people first tried to use computers to translate foreign language texts into English and sought to develop algorithms and software for intelligently processing language data.

In 2000, the *Wall Street Journal* published a high profile article on the explosion of the computational linguistics field. The immense opportunity of the Internet was leading established and start-up technology companies to hire linguists in droves. These linguists, who had completed various levels of education and had computer science

expertise as well, helped companies use natural-language processing to respond meaningfully to requests for help or information. What were some of the initial applications? The *Journal* provided the example of how an online customer asking about shaving products might be automatically asked whether he also needed razors, blades or shaving cream.

In the years that followed, as the Internet became even more commercially viable, hundreds of companies sprung up around the development of fast and user-friendly products that leveraged linguistic phenomena to get customers the right data in the fastest, most efficient manner possible. Today, a computational linguist working within a technology company or research institution might be responsible for building or testing natural-language processing systems, or engineering linguistic applications (sometimes known as lingware) or related software. Or, her role might involve making it easier for employees to access critical knowledge resources, as is the case with 32-year-old Helen, a computational linguist in Colorado. “My job entails describing the variation of language, and devising ways for programs to deal effectively with that variation,” she says. “On a daily basis, this translates to activities such as searching bodies of text for specific grammatical structures or conceptual constructs, describing the features of variation, and then programming automatic concept recognizers and text processors to deal with the variation.”

Some computational linguists are helping to develop future technologies. “I recently had a meeting with some folks from a local start-up company, VoiceBox, which builds conversational-speech based interfaces for a variety of gadgets, including satellite radio, in-dash navigation systems, and other sources of information you might find inside

a car,” says Emily, 33, a computational linguist in Seattle. “Their technology allows you to control these via natural spoken commands and requests, to which the car can reply via a synthesized voice. While KIT from Night Rider and HAL from 2001 are still a long way off, it's thrilling to be involved in bringing such fantasies one step closer to reality.”

If all this sounds as intriguing to you as it does to me, you're probably wondering how you can score a gig in the field. The good news is, there are several ways in. Your best bet is to complete a bachelor's, master's, or doctoral degree program in computational linguistics, which will teach programming languages such as C++ or Perl and will offer coursework in natural-language processing, artificial intelligence, machine translation, and text mining. Those who have obtained a bachelor's degree in linguistics, and have therefore mastered basics such as phonetics (the study of sounds), morphology (the study of words), syntax (the study of sentences), and semantics (the study of meaning) are pretty marketable too. However, due to the shortage of private-sector candidates with computational linguistics experience, many enter the field purely on the strength of their computer science or computer engineering skills. “What I hear from hiring managers at human language technology companies is that it is rare to find people who are trained in both computer science and linguistics, so they typically end up taking people who are trained in computer science and giving them some linguistics and computational linguistics training on the job,” says Emily.

The contacts you'll be exposed to by doing a formal degree program in computational linguistics may make finding your first job in the field relatively easy. But those data-heads entering from the computer science side of things will likely need a little more assistance. The experts I spoke with recommended getting involved with the

Linguistics Society of America (www.lsadc.org) and the Association of Computational Linguistics (www.aclweb.org), and the North American Association of Computational Linguistics (www.cs.cornell.edu/home/lee/naacl). All of these organizations will point you in the direction of valuable learning resources, networking opportunities, and employment options. You can also make headway by attending their industry conferences. “These events act as venues for the cutting-edge work in natural language processing,” says Helen. “There are usually student rates, and some conferences accept volunteer work at the event in exchange for registration fees.” In addition to your networking activities, you should scout openings advertised by the top companies themselves (Microsoft at www.microsoft.com, IBM at www.ibm.com, Cymfony at www.cymfony.com, Google at www.google.com, AOL at www.aol.com, etc.), or by checking out the following Web sites:

- Computer Jobs (www.computerjobs.com)
- Dice (www.dice.com)
- Jobs for Programmers (www.prgjobs.com)
- Linguist Careers (www.linguistcareers.com)
- Linguist List (www.linguistlist.org)
- Technology Recruiter (www.technologyrecruiter.com)

Finally, be patient, because there is a lot to learn before you’ll be able to be an effective member of a computational linguistics team. “Even if you haven’t landed that job yet, find projects to keep yourself busy so you have more hands on experience with challenges common to the field,” says Helen. “Teach yourself new programming languages, new software engineering practices, practice with existing natural language processing tools that are out there, and stay up to date with current linguistic research and innovation by reading tech journals.” Suggests Emily: “If you live near a university, find out if they have computational linguistics or natural language processing research

going on. Colloquia and other research presentations are usually open to the public, and there are typically e-mail lists that you can join.

It's a rare individual who is blessed with equally strong language interpretation, analytical, and critical thinking skills, and for this reason gifted computational linguists tend to be well-compensated. According to Robert Half International, starting salaries for software engineers ranged from \$63,250 to \$92,750 in 2005, and those with a background in computational linguistics often hit six figure salaries in just a few years. It's far from an easy road, though. Computational linguists are required to develop products in a pressure-filled environment, under tight timeframes and budgets, and their success often depends on the strength of the technologies they're working with. "Everything is done on the computer, and so if a server goes down or a hard drive acts up, my work completely halts," says Helen.

Nevertheless, the future for computational linguistics is certainly bright. "I fell in love with linguistics during my second semester in college, but I felt like I had to make my peace with studying something that I saw as having little practical application," says Emily. "There's still that perception, and now people are starting to say that computer science isn't a profitable field to go into because so many jobs are being outsourced overseas. But in fact, U.S. industries are hungry for graduates trained in human language technology."