3Qs: Researcher studies how human listeners process speech in order to recognize spoken words

June 4 2014, by Angela Herring

Joanne Miller, Matthews Distinguished University Professor and chair of Northeastern's Department of Psychology, was recently recognized for her pioneering research on human language processing in the field of speech perception. The 167th Meeting of the Acoustical Society of America, held last month in Rhode Island, honored Miller by hosting a special session that featured talks by researchers whose work complements and builds on Miller's contributions over four decades. Here, we asked Miller to discuss her research in more detail.

What primary research question did you focus on?

Our main focus was on the way in which human listeners process speech in order to recognize spoken words. When people speak, they generate a complex and continuously varying acoustic signal. Somehow the listener's brain analyzes this signal and recognizes the individual consonants and vowels—the phonetic units—that make up the words of the language. This seems effortless—listeners don't even have to think about it. But it's really a very intricate process. The question is: How does the process work?

The central problem we worked on is variability. Each time you meet someone new, you encounter a unique voice; each person pronounces the consonants and vowels that make up a given word a bit differently. Moreover, people speak at different rates of speech, with different

accents, and in different dialects. All of these factors modify the speech signal in systematic ways. Even the same person will produce a word slightly differently each time it's spoken. In order to recognize which word was spoken—which sequence of phonetic units was produced—our perceptual system has to contend with all of these sources of acoustic variability, and more.

What were the major contributions of your research?

The issue of variability has been at the core of speech perception research for decades. Very early on in the field, people thought that our perceptual system mostly ignored all the fine variation in the speech signal and only focused on the categorical identity of the specific consonants and vowels that make up the words. So, for example, the idea was that listeners focus on identifying "b" versus "p" to distinguish the words "bath" and "path," and largely ignore the fine-grained differences in various pronunciations of "b" or "p."

Over many years, we and other research groups found that this isn't the case. Instead, our perceptual system is exquisitely sensitive to fine phonetic detail—to exactly how a word is pronounced. And, critically, we now know that when listeners process this fine-grained variation in speech, they are continually adjusting, in very systematic ways, for the different contextual factors that produce the variability. In our view, this dynamic adjustment plays an important role in listeners' ability to recognize spoken words. Much of our work involved measuring acoustic properties of speech, selectively manipulating these properties using various speech synthesis techniques, and then testing the effects of our manipulations on listeners' perception.

How is the field building on the ideas that you put forth?

On the basis of our work, and that of many other labs, there's been a growing emphasis on how listeners process the fine phonetic detail of speech and the role this plays in spoken language processing overall. One can see this emphasis across a number of domains, including research on voice recognition, language acquisition, speech and language disorders, and the intelligibility of speech in noise. One particularly intriguing domain concerns the neural basis for such processing. When we began our work many years ago, there were very few tools for investigating how the brain analyzes the speech signal and processes spoken language. But this has changed with the development of a host of neuroimaging techniques, and understanding the neural basis for how we perceive speech is now a thriving research area.

Provided by Northeastern University