

Better Speech Recognition with Vision

BY LYNNE E. BERNSTEIN, EDWARD T. AUER, JR., SILVIO P. EBERHARDT AND NICOLE JORDAN

In these days of digital hearing aids and cochlear implants, it may be easy to underestimate the benefits of looking at people when they are talking. But being able to use their visible speech, particularly when you are in a noisy social setting, can be very helpful. A hearing aid or a cochlear implant can improve speech in noise by about two or three decibels, but many scientific studies show that visual speech combined with hearing speech is equivalent to about 12 or more decibels of noise reduction.

The ability to comprehend speech by watching someone talk is referred to as “lipreading” or “speech-reading.” Lipreading is possible, because when we talk, the motions and shapes of our lips, jaw, tongue and cheeks

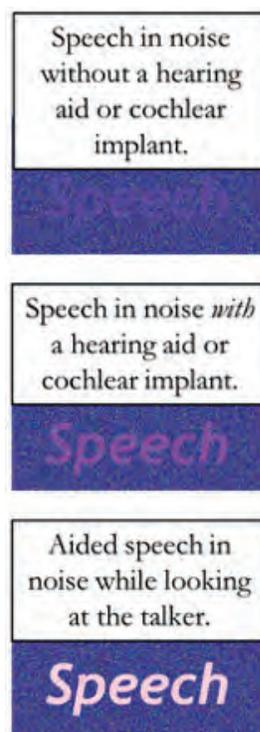
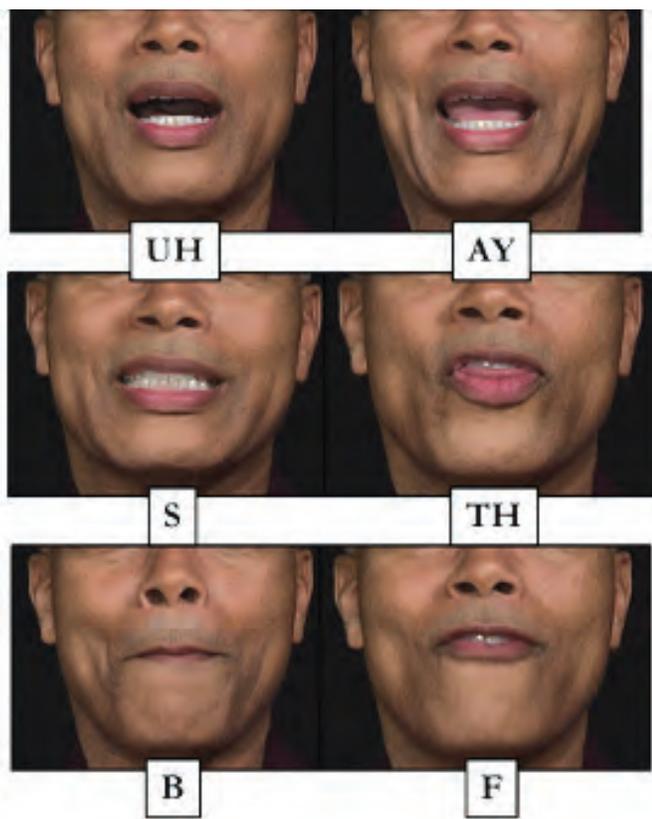
that form speech sounds also form speech “sights.”

Unfortunately, most adults who have had good hearing throughout a large part of their lives are not skillful in recognizing the visible speech sounds. In our experience, many adults with hearing loss would like to improve their lipreading abilities and some have even managed to train themselves to better recognize visible speech. Others might have tried learning through a class or perhaps a website that offers lipreading practice. However, learning to lipread better is notoriously difficult. Previous research efforts suggested that lipreading cannot be trained: You were either born with it or you were not. We suggest otherwise in our research.

Lipreading ability is not an inborn trait. Our research on adults who were deaf before learning language and who did not use cochlear implants has shown that most learned to be very good or excellent lipreaders. This tells us that lipreading is learnable and that there is information to be learned by focusing on the talker’s face.

However, in order to make the most use of visual speech information, adults need more than reminders to look at the face of the person who is talking. They need structured training experience with feedback. In our research, we have found that the type of feedback is critical to learning to lipread more accurately and with better lipreading skills comes improved audiovisual speech recognition in noisy situations. In other words, the correct type of lipreading training can also improve an adult’s ability to use the combination of listening and looking when there is a lot of noise.

Samples of Visible Speech Sounds



What are the keys to successful training? Well, we probably do not know all of them, but one of the keys is giving feedback based on what the lipreader actually thought the talker said, as opposed to merely revealing the correct answer. Another key is doing the correct type of training task.

With the internet and powerful home computers, it is now possible to self-train on lipreading and audiovisual speech recognition at home. The new training methods we are developing are based on the neuroscience of perceptual learning. That science tells us that for adults to improve their lipreading, they need feedback that corrects their lipreading errors on consonants and vowels but not through drills on separate speech sounds such as “b” or “f.” The most effective training we have studied uses words and gives feedback for lipreading errors within words. We have developed software that analyzes errors and focuses training on “near misses.”

Our research tells us that many adults who attempt to lipread make errors but these lipreaders may be surprisingly close to lipreading correctly. A favorite example of ours is a lipreader who saw the sentence, “Proofread your final results,” and thought the talker said, “Blue fish are funny.” Obviously, the lipreader completely misunderstood the words in the sentence. But our software program shows that this lipreader committed several “near misses.” The software used its algorithms to line up the correct and the incorrect sentences to show the problems. The sounds “r” and “l” in “proof” and “blue” are actually visually different, and training with feedback can improve recognition of these sounds.

Talker said:	Proo f	read your final results.
Lipreader reported:	Blue fish are	funny.

There remains much for us to learn about the best methods for guiding adults with hearing loss to become better lipreaders and better at comprehending audiovisual speech in noise. We are presently carrying out research on several different training techniques. If you might be interested in participating, please contact us using one of the methods listed here. **HL**

NOTE: This article is based on two articles: Bernstein, L. E., Auer Jr, E. T., & Eberhardt, S. P. (2022). During lipreading training with sentence stimuli, feedback controls learning and generalization to audiovisual speech in noise. *American Journal of Audiology*, 31, 57-77. doi:10.1044/2021_AJA-21-00034; and Bernstein, Jordan, Auer, Jr., & Eberhardt, S. P. (in press). *Lipreading: A Review of its Continuing Importance for Speech Recognition with a Hearing Loss*.

Lynne E. Bernstein, Ph.D., Edward T. Auer, Jr., Ph.D., Silvio P. Eberhardt, Ph.D. and Nicole Jordan, Au.D., are faculty in the Speech, Language & Hearing Sciences Department at George Washington University, Washington, DC. They have been working together for quite a while to understand lipreading and audiovisual speech recognition and to develop approaches to ameliorate speech recognition difficulties in adults with hearing loss. Reach out to Lynne at lbernste@gwu.edu.

If you are interested in participating in research funded by the National Institute on Deafness and Other Communication Disorders (NIH) on lipreading and audiovisual speech recognition in noise, please email study@seehear.us, or use your browser to go to seehear.us/study or use this QR code.





HLAA MISSION CIRCLE

A monthly online giving program offering convenience while supporting work for people with hearing loss.

hearingloss.org/make-an-impact