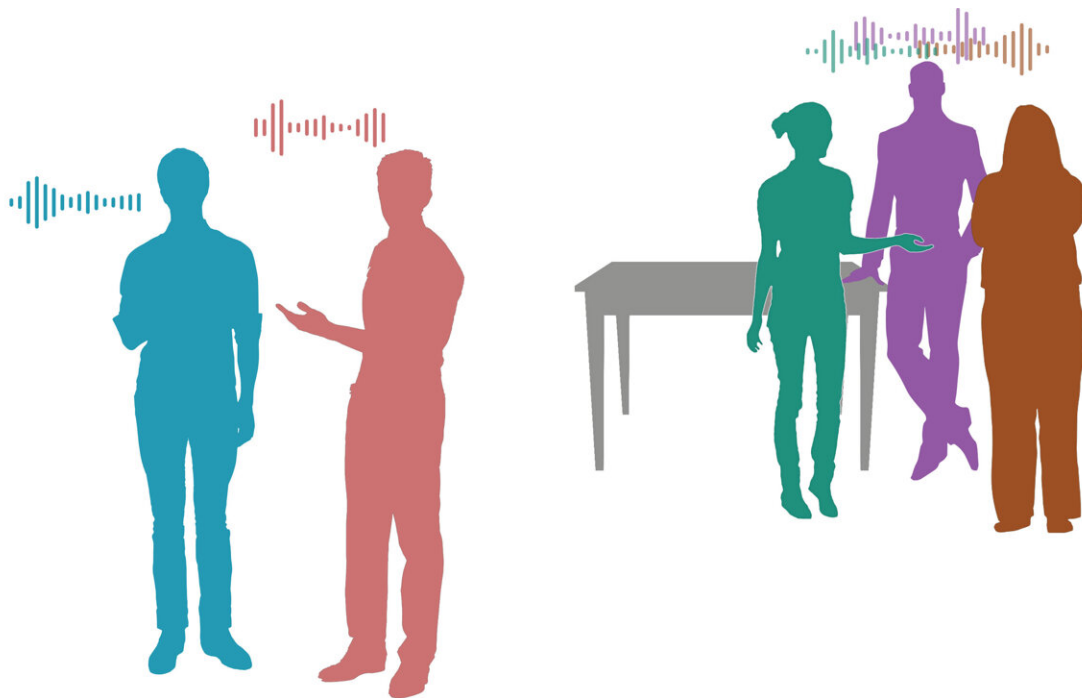


To better understand speech, focus on who is talking

October 26 2021



Seeing a talker's face improves your ability to perceive speech, but only if the face and voice come from the same location in space. Credit: Justin Fleming

Seeing a person's face as we are talking to them greatly improves our ability to understand their speech. While previous studies indicate that the timing of words-to-mouth movements across the senses is critical to this audio-visual speech benefit, whether it also depends on spatial alignment between faces and voices has been largely unstudied.

Researchers found matching the locations of faces with the [speech](#) sounds they are producing significantly improves our ability to understand them, especially in noisy areas where other talkers are present.

In the *Journal of the Acoustical Society of America*, researchers from Harvard University, University of Minnesota, University of Rochester, and Carnegie Mellon University outline a set of online experiments that mimicked aspects of distracting scenes to learn more about how we focus on one audio-visual talker and ignore others.

"If there's only one multisensory object in a scene, our group and others have shown that the brain is perfectly willing to combine sounds and visual signals that come from different locations in space," said author Justin Fleming. "It's when there's multisensory competition that spatial cues take on more importance."

The researchers first asked participants to pay attention to one talker's speech and ignore another talker, either when corresponding faces and voices originated from the same [location](#) or different locations. Participants performed significantly better when the face matched where the [voice](#) was coming from.

Next, they found [task performance](#) decreased when participants directed their gaze toward a voice trying to distract them.

Finally, the researchers showed spatial alignment between faces and voices was more important when the [background noise](#) was louder, suggesting the brain makes more use of audio-visual spatial cues in challenging sensory environments.

The pandemic forced the group to get creative about conducting such research with participants over the internet.

"We had to learn about—and, in some cases, create—several tasks to make sure participants were seeing and hearing the stimuli properly, wearing headphones, and following instructions," Fleming said.

Fleming hopes their findings will lead to improved designs for hearing devices and better handling of sound in virtual and augmented reality. They look to expand on their work by bringing additional real-world elements into the fold.

"Historically, we have learned a great deal about our sensory systems from studies involving simple flashes and beeps," he said. "However, this and other studies are now showing that when we make our tasks more complicated in ways that better simulate the real world, new patterns of results start to emerge."

More information: Justin T. Fleming et al, Spatial alignment between faces and voices improves selective attention to audio-visual speech, *Journal of the Acoustical Society of America* (2021).
doi.org/10.1121/10.0006415

Provided by American Institute of Physics

Citation: To better understand speech, focus on who is talking (2021, October 26) retrieved 11 July 2024 from <https://phys.org/news/2021-10-speech-focus.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.