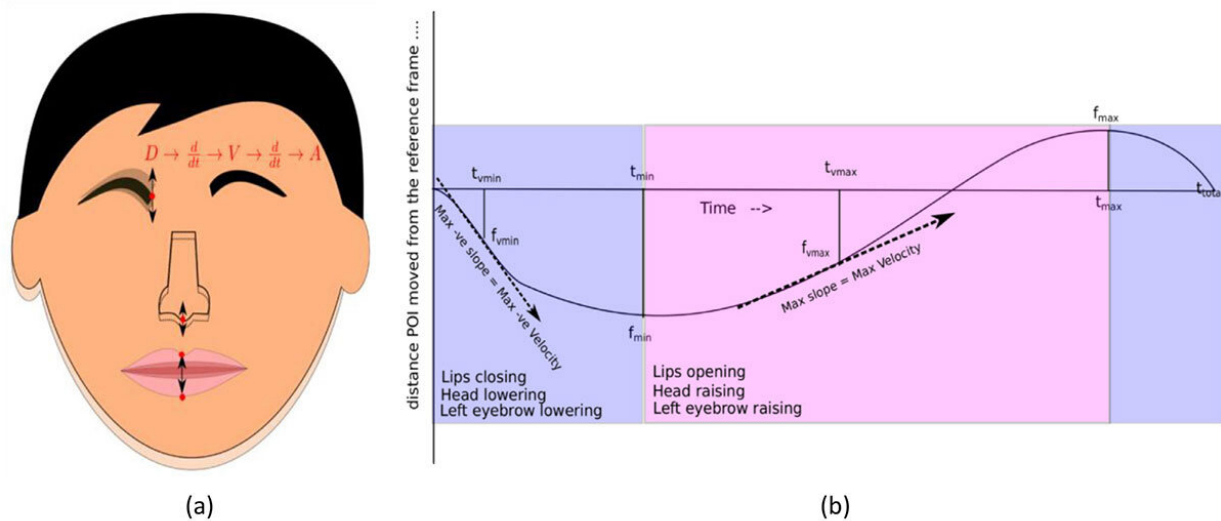


# Study shows facial features track with intonation of words

September 24 2019, by Rick Hellman



In the diagram above, the left side shows the position and direction of movement of the four points tracked in the study. The shadows near eyebrow/jaw/nose indicate motion. At right is a visual summary of the different features extracted from the four tracked key points. POI=Position of Interest. Credit: University of Kansas

Even though they are not needed to make the specific sounds, parts of Mandarin Chinese speakers' faces—their eyebrows and lips—mimic the rising and falling pitch that distinguishes one word spelled exactly the

same from another.

In other words, speakers' eyebrows follow the sounds their throat is making.

This is the finding of a new study co-written by Allard Jongman and Joan Sereno, University of Kansas professors of [linguistics](#).

They say it implies the need to study whether these subtle facial movements help listeners perceive what the speakers are trying to convey.

Along with Canadian university colleagues Saurabh Garg, Ghassan Hamarneh and Yue Wang, Jongman and Sereno wrote "Computer-vision analysis reveals facial movements made during Mandarin tone production align with pitch trajectories" in the October edition of the journal *Speech Communication*.

They say it's the first time these [facial movements](#)' association with vocal tone has been identified and quantified.

In a recent interview, Jongman gave the example of the Chinese word that would be phonetically spelled "ma" in English. As with many, if not most, Mandarin words, it can be pronounced one of four ways: level, rising, dipping or falling. Depending, then, upon intonation, the word can mean "mother," "hemp," "horse" or "scold."

"You've got to get this right," Jongman said. "Otherwise, you're calling your mother a horse."

"There are a lot of languages that have these types of pitch that are associated with different words," Sereno said.

"The majority of speakers in the world are tone-language speakers," Jongman said. "And it all comes from variations in the rate of vocal-fold vibration."

Pointing to his throat, Jongman said, "At high pitch, the vocal folds down here are vibrating at a very fast rate. And when you have a low pitch, they vibrate much more slowly. So it all comes from the action down here."

"Technically," Sereno said, pointing to her face, "there shouldn't be anything up here that differentiates 'ma' with a high pitch versus 'ma' with a falling pitch."

And yet, their study found, there is. People's facial structures mimic the rising/falling intonations they are applying to their words.

"Some of these associations between facial movement and intonation or pitch had been provisionally established," Jongman said. "We went one step further and said, 'OK, let's systematically look at this.'"

The researchers recorded the Chinese speakers' faces as they spoke, then subjected the files to computer analysis to determine the relationship between intonation and facial movement. Jongman said it's a big improvement over previous methods.

"In the past," he said, "people have used what is called opto-track analysis. You glue a couple of light-emitting sensors to the face, and then you track the motion of those sensors. And sometimes you even glue stuff to the lips. So you're sitting there with these sensors, and you are asked to speak. I mean, how natural is that? That limits what people would do in normal circumstances. Number two, you have to have a preconceived idea of what the important regions on the face are going to be, because you can't take a reading of anything that you didn't put a

sensor on.

"That's one big advantage of our approach. We have settled on these couple of features now. But if later on some other group finds that your earlobe is really important, we can go back to our recordings and analyze that."

Now that they have proven that Mandarin speakers' faces track with their intonation, Jongman and Sereno said the next step will be to determine to what extent these visual cues help the listener—their "linguistic significance."

"We're going to computer-generate a face and then manipulate those features that we just found and see whether it makes a difference," Jongman said. "Do listeners respond to it?"

Just as people with hearing impairment can learn to read lips, people might be able to benefit from the more subtle facial cues, the new study has found.

"If you're hearing-impaired, and we can maybe exaggerate some of these cues a little bit or teach people to exaggerate them, or also in language learning do that, that might greatly help," Jongman said.

Sereno said she can imagine a scenario in which a language learner is helped by watching native speakers' facial cues.

"Visual manipulation could help identify that," she said. "Then when they get into a natural situation, when there's maybe less exaggeration of those cues, they're still able to identify them a lot better."

**More information:** Saurabh Garg et al. Computer-vision analysis reveals facial movements made during Mandarin tone production align

with pitch trajectories, *Speech Communication* (2019). [DOI: 10.1016/j.specom.2019.08.003](https://doi.org/10.1016/j.specom.2019.08.003)

Provided by University of Kansas

Citation: Study shows facial features track with intonation of words (2019, September 24) retrieved 10 May 2024 from <https://phys.org/news/2019-09-facial-features-track-intonation-words.html>

<p>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.</p>
--