

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Modeling auditory voice recognition improvements by face simulation

Permalink

<https://escholarship.org/uc/item/84d61188>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Gumbsch, Christian

Butz, Martin V.

von Kriegstein, Katharina

Publication Date

2024

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Modeling auditory voice recognition improvements by face simulation

Christian Gumbsch

TUD Dresden University of Technology, Dresden, Germany

Martin Butz

University of Tuebingen, Tuebingen, Germany

Katharina von Kriegstein

TU Dresden University of Technology, Dresden, Germany

Abstract

Voice identity recognition in auditory-only conditions is facilitated by knowing the face of the speaker. This effect is called the ‘face-benefit’. Based on neuroscience findings, we hypothesized that this benefit emerges from two factors: First, a generative world model integrates information from multiple senses to better predict the sensory dynamics. Second, the model substitutes absent sensory information, e.g., facial dynamics, with internal simulations. We have developed a deep generative model that learns to simulate such multisensory dynamics, developing latent speaker characteristic contexts. We trained our model on synthetic audio-visual data of talking faces and tested its ability to recognize speakers from their voice only. We found that the model recognizes previously seen speakers better than previously unseen speakers when given their voice only. The modeling results confirm that multisensory simulations and predictive substitutions of missing visual inputs result in the face-benefit