23 Who will start speaking next and when?
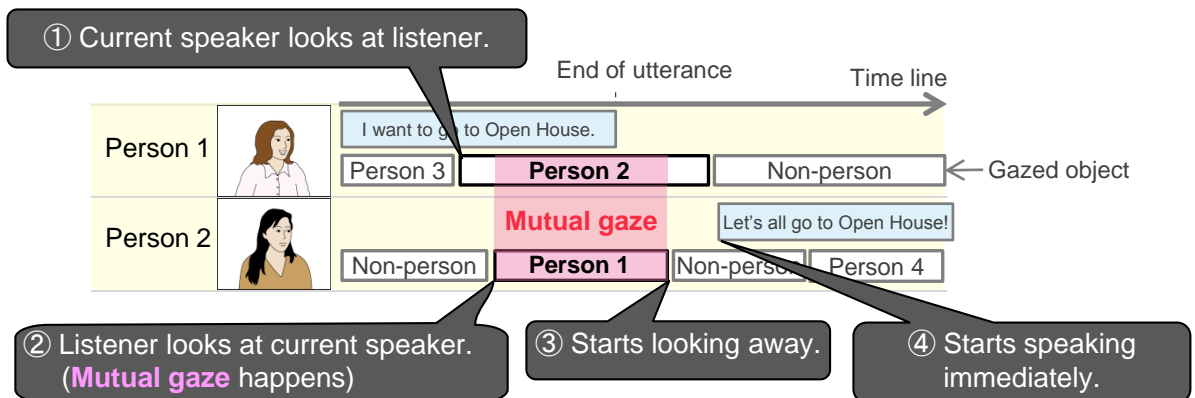## ~Predicting the next speaker and timing from gaze and breathing~

Abstract

People in meetings usually take turns speaking without uttering simultaneously. Our aim is to clarify the kinds of behavior that contribute to such smooth turn-taking and develop a model for predicting the next speaker and the start time of his or her next utterance in meetings. We empirically demonstrated that gaze behavior and respiration are related to the next speaker and the start time in meetings and can be used to predict them. If such a prediction model were developed, it would lay the foundation for the development of natural conversational systems in which conversational agents speak with natural timing and for teleconference systems that avoid utterance collisions by apprising who of the participants will speak next and adding appropriate time delays.
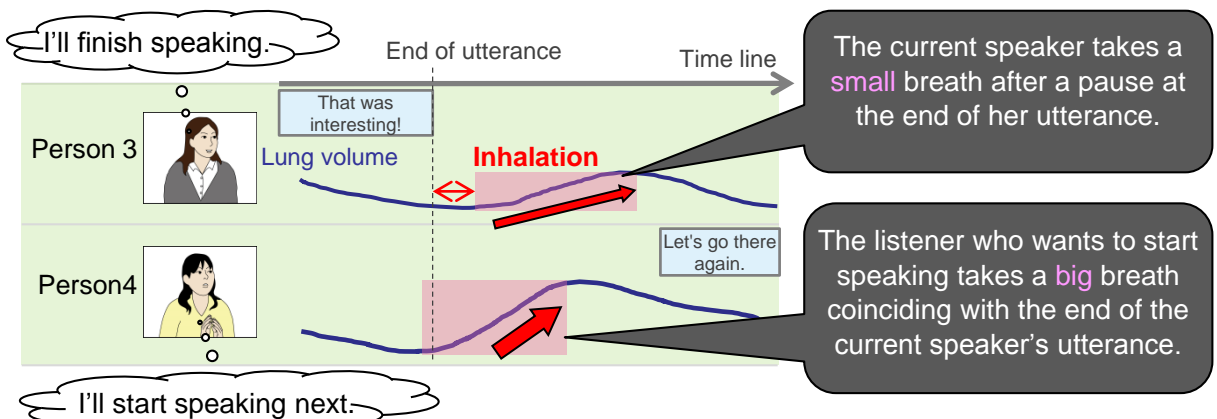
Gaze — We can tell who will start speaking and when from **changes in gaze.**

① Current speaker looks at listener.

② Listener looks at current speaker. (**Mutual gaze** happens)

③ Starts looking away.

④ Starts speaking immediately.

Breathing — We can see who will try to start speaking and when from **inhalation patterns.**

The current speaker takes a small breath after a pause at the end of her utterance.

The listener who wants to start speaking takes a big breath coinciding with the end of the current speaker's utterance.

Related works

[1] R. Ishii, K. Otsuka, S. Kumano, J. Yamato, "Analysis of respiration for prediction of "who will be next speaker and when" in multi-party meetings," in Proc. ACM International Conference on Multimodal Interaction (ICMI), pp. 18-25, 2014. (Outstanding Paper Award)

[2] R. Ishii, K. Otsuka, S. Kumano, J. Yamato, "Analysis and modeling of next speaking start timing based on gaze behavior in multi-party meetings," in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 694-698, 2014.

[3] R. Ishii, K. Otsuka, S. Kumano, M. Matsuda, J. Yamato, "Predicting next speaker and timing from gaze transition patterns in multi-party meetings," in Proc. ACM International Conference on Multimodal Interaction (ICMI), pp. 79-86, 2013.

Contact

Ryo Ishii Sensory Resonance Research Group Human Information Science Laboratory
E-mail : ishii.ryo(at)lab.ntt.co.jp