UNIVERSITÄT BIELEFELD

Speech and Co-speech Gestures in Interactions

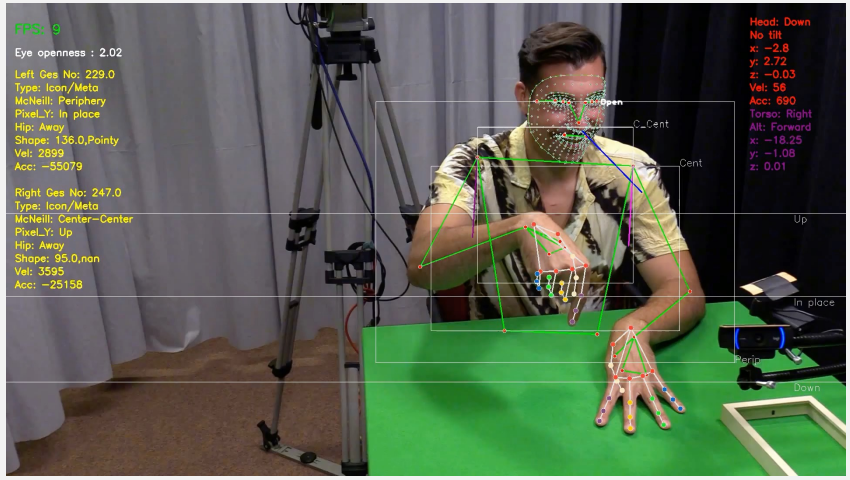
- Co-speech gestures linked to prosodic expressions
- Speakers use their entire bodies to communicate
- This is best investigated in spontaneous speech-based interactions

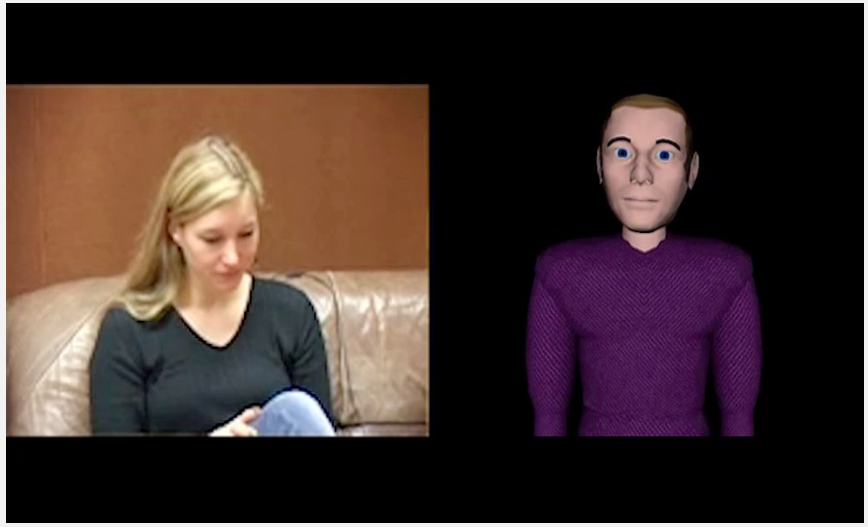
Wagner, P., Z. Malisz & S. Kopp (2014). Gesture and Speech in Interaction. *Speech Communication*, 57, 209-232.

UNIVERSITÄT BIELEFELD

Feedback signals

- How do we signal attentive listening?
- How do we signal (non-)understanding
- Can we apply this knowledge to HMI?
- How language- or culture-specific are these insights?





Türk, O., S. Lazarov, Y. Wang, A. Grimminger, H. Buschmeier, & Wagner, P. (2024). Predictability of Understanding in Explanatory Interactions Based on Multimodal Cues. *Proceedings of ICMI'24*, San José, Costa Rica

Schade, L., Dallmann, N., Türk, O., & Wagner, P. (2024). Understanding "understanding": Presenting a richly annotated multimodal corpus of dyadic interaction. *INTERSPEECH 2024 (S&T)*, Kos, Greece.

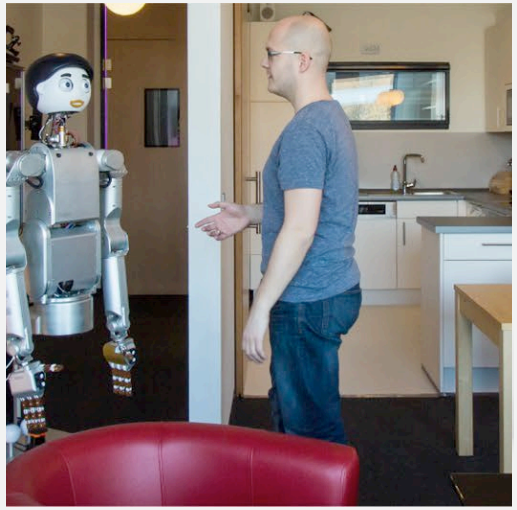
Malisz, Z., Wlodarczak, M., Buschmeier, H., Skubisz, J., Kopp, S., & Wagner, P. (2016). The ALICO Corpus: Analysing the Active Listener. *Language Resources and Evaluation*, *50*(2), 411–442. doi:10.1007/s10579-016-9355-6

Jabeen, F., Bryhadyr, N., & Wagner, P. (2022). *Yes or no: An analysis of Indian head gestures*. International Society of Gesture Studies Conference, Chicago, IL, USA.

UNIVERSITÄT BIELEFELD

Speaking Styles and Situations

- How do speaking styles differ across different situations, und
- How do speaking styles differ when humans interact with machines?
- How human-like should machines be when interacting with humans. Should they be able to laugh, for instance?



Wagner, P., Trouvain, J., & Zimmerer, F. (2015). In defense of stylistic diversity in speech research. *Journal of Phonetics*, *48*, 1-12. doi:10.1016/j.wocn.2014.11.001

Ludusan, B., & Wagner, P. (2021). Knock-Knock! Who's There? The Laughter-Enhanced Virtual Real-Estate Agent. In S. Hillmann, B. Weiss, T. Michael, & S. Möller (Eds.), *Studientexte zur Sprachkommunikation : Vol. 99.* (pp. 281-288). Dresden: TUDpress, 2021.

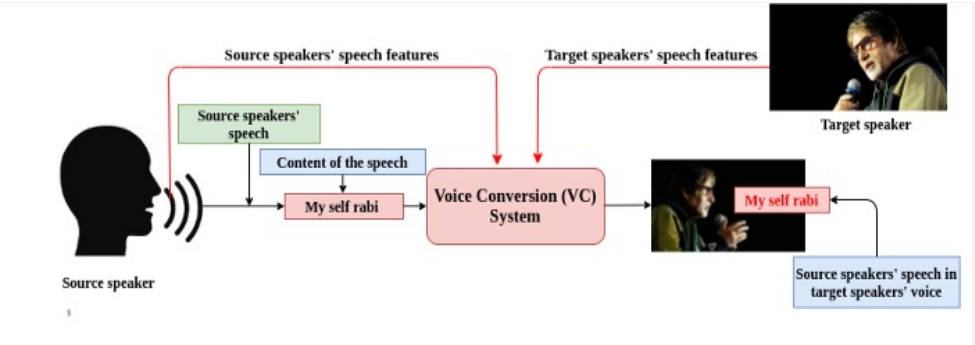
https://50jahre.uni-bielefeld.de/2019/01/11/the-art-of-making-our-speech-assistant-more-human/

UNIVERSITÄT BIELEFELD

Voice quality, voice conversion, TTS quality

- Voices are diverse and express our personalities, but very difficult to measure or to describe or explain
- AI-based voice adaptation systems are able to capture individual voice patterns and superimpose them to a different utterance.

➔ We use such technologies to better understand voice profiles, and use phonetic techniques to evaluate synthesized voices.



Quelle: Dhar, S. (2021). Introduction to Deep Learning-based Voice Conversion: a growing domain of speech synthesis research.

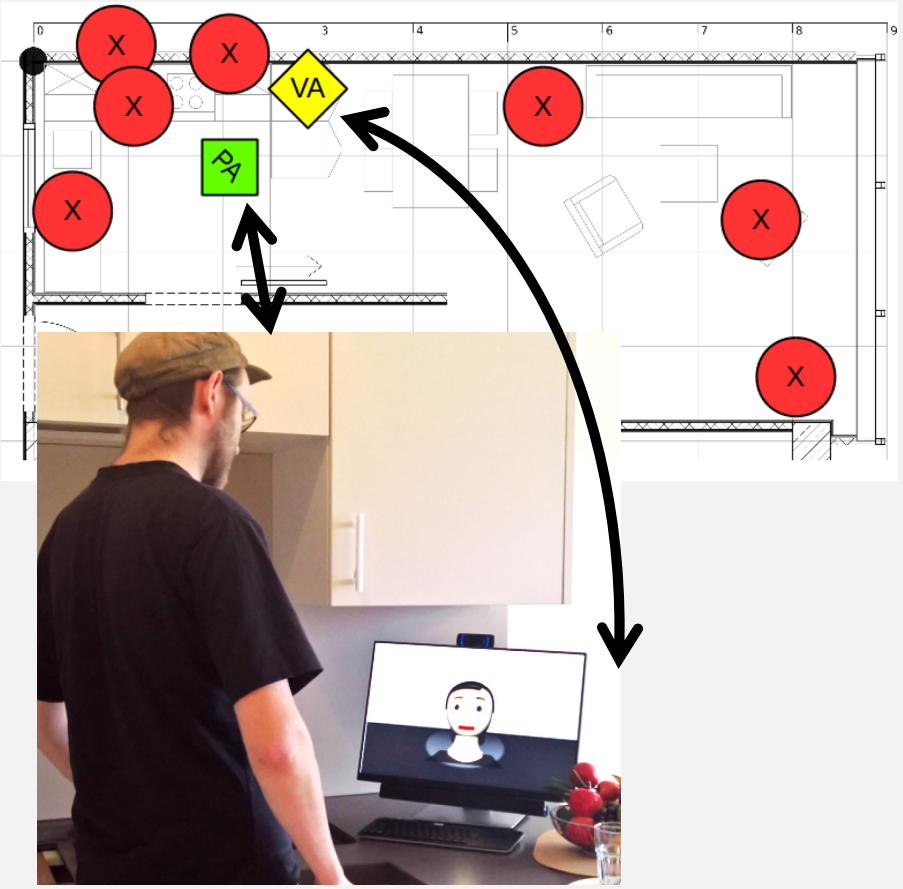
Wiechmann, J., Rautenberg, F., Wagner, P., & Häb-Umbach, R. (2023). *Explaining voice characteristics to novice voice practitioners - How successful is it?*. Presented at the 20th International Congress of the Phonetic Sciences (ICPhS) , Prague, Czech Republic.

Seebauer, F. M., Kuhlmann, M., Haeb-Umbach, R., & Wagner, P. (2023). Re-examining the quality dimensions of synthetic speech. *12th ISCA Speech Synthesis Workshop (SSW2023)*, 34-40. ISCA. https://doi.org/10.21437/SSW.2023-6

UNIVERSITÄT BIELEFELD

Non-verbal vocalizations: hesitations, laughter, breathing

- Hesitations, breathing and laughter structure interactions, are time-buying and signal cognitive load.
- Hesitations increase task performance in HMI, if they guide the listeners‘ attention.
- Laughing Artificial Agents are more likeable, but also perceived as less formal.
- Hesitations, breathing and laughter are perceived multimodally!



Betz, S., Wagner, P., & Wrede, B. (2018). Interactive Hesitation Synthesis: Modelling and Evaluation. *Multimodal Technologies and Interaction*, *2*(1).
Ludusan, B., & Wagner, P. (2022). Laughter entrainment in dyadic interactions: Temporal distribution and form. *Speech Communication*, *136*, 42-52.