



## *Divergent senses and sensors in multimodal addressee design: Avoiding gestural mismatches in human-robot interaction scenarios*

*Ellen Fricke and Jana Bressemer*



Foto: Tobias Naumann



[www.gesten-im-museum.de](http://www.gesten-im-museum.de)

Chair of German Linguistics,  
Semiotics, and Multimodal  
Communication

Chemnitz University of Technology

CRC “Hybrid Societies”  
Project D01: Intentionality and Joint  
Attention in Multimodal Interaction

## Research questions

- How can hybrid group settings consisting of participants with different senses and sensors be communicatively designed?
- What kind of mismatches occur and how can these be resolved – technically as well as semiotically?

## Goals

- Initial steps for an extended concept of addressee design (based on recipient design in CA) including hybrid settings of multimodal and multisensorial human-robot interaction
- Illustration of its potential in combination with the complementary concept of accountability



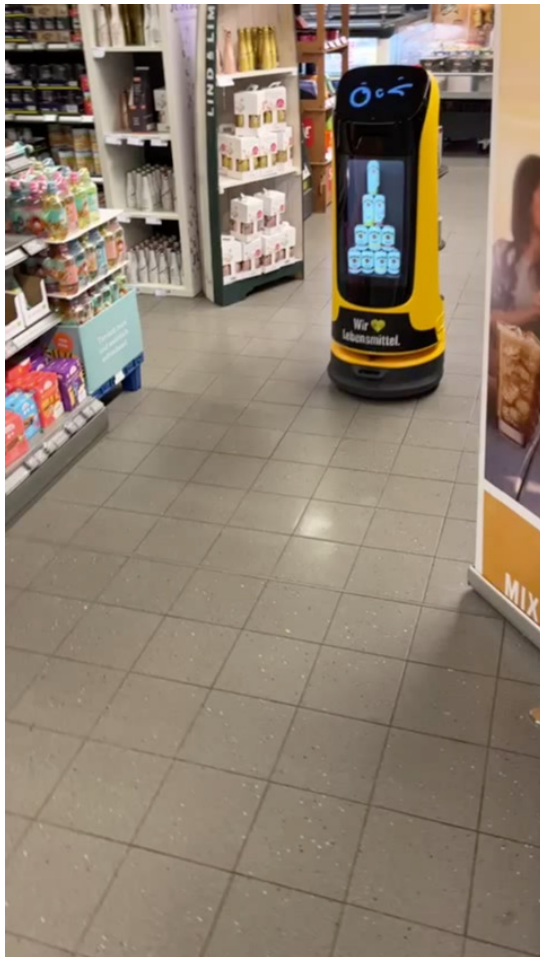
# Hybrid Societies (definition)

“Specifically, we define a Hybrid Society as a group of embodied agents (including humans, partly human actors such as cyborgs, and non-human embodied technologies such as robots) with a capability to engage in intelligible encounters, who interact and communicate, who can meaningfully reference each other as members of society, and who vary in terms of autonomy, agency, and responsibility (cf. S. Meyer et al., 2023).” (CRC Proposal)

 Requires adaptation of concepts for analyses

Meyer, S., Mandl, S., Gesmann-Nuissl, D., & Strobel, A. (2023). Responsibility in Hybrid Societies: concepts and terms. *AI and Ethics*, 3, 25-48. <https://doi.org/10.1007/s43681-022-00184-2>





## Hybrid Society?

## Kettybot in an EDEKA supermarket in Chemnitz.

Thanks to Erik Zimmermann.  
(Edeka supermarket  
Weststraße 77 in Chemnitz)

## Freie Presse

Chemnitz

🏠 | Chemnitz | Warum ein Roboter im Edeka durch die Gänge fährt

Erschienen am 13.05.2023

### Warum ein Roboter im Edeka durch die Gänge fährt



Von [Christian Mathea](#)



Der Roboter namens "Kettybott" rollt in diesen Tagen durch den Edeka an der Weststraße und wirbt für Pralinen

## Robots in the wild: delivery robot in Tel Aviv

Don't pet me!

Instructive example  
of how not to design  
a delivery robot.

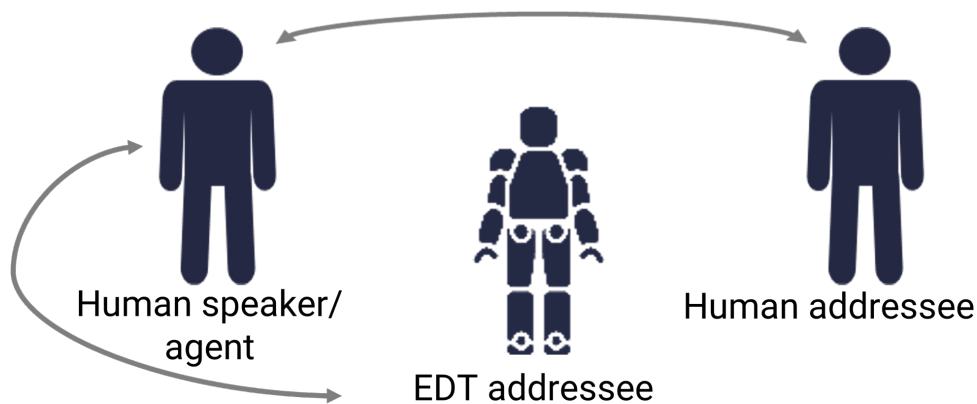
No capability  
to interpret gestures.



Thanks to my colleague Bertolt Meyer.

[https://www.youtube.com/watch?v=S2fmZAU6gSY&ab\\_channel=TheHolyLandChannelbyZahiShaked](https://www.youtube.com/watch?v=S2fmZAU6gSY&ab_channel=TheHolyLandChannelbyZahiShaked)

## Two complementary concepts: Accountability and Addressee Design



### Accountability

The extent to which a robot signals to incidental users what it wants, what it is capable of and how to interact with it.

Müller, Michael R. (2023). "Social Displays. Creating Accountability in Robotics". *Österreichische Zeitschrift für Soziologie*, 49(2).

### Addressee design

(based on recipient design in CA) describes the adaptation of communicative behavior to the abilities, knowledge, and conditions of an addressee (humans as well as EDTs).

Sacks, H. (1992). *Lectures on Conversation*. Blackwell.

Winner, T., Selen, L., Oosterwijk, A. M., Verhagen, L., Medendorp, W. P., van Rooij, I., & Toni, I. (2019). Recipient Design in Communicative Pointing. *Cognitive Science*, 43(5). <https://doi.org/10.1111/cogs.12733>

## Addressee design in human-robot interaction

- (1) attributions of goals, intentions, abilities, knowledge, etc. to addressees, followed by a corresponding adaptation of one's own behavior;
- (2) **attribution of sensory capabilities**: e.g., with regard to a blind person or a robot that does not have optical sensors, one will have to use other types of signs as, for example, visible pointing gestures to localize an object;
- (3) assignment of particular context conditions;
- (4) attributions of linguistic ability or other semiotic competences;
- (5) establishing joint attention;
- (6) adaptation to inter-cultural differences and further aspects.

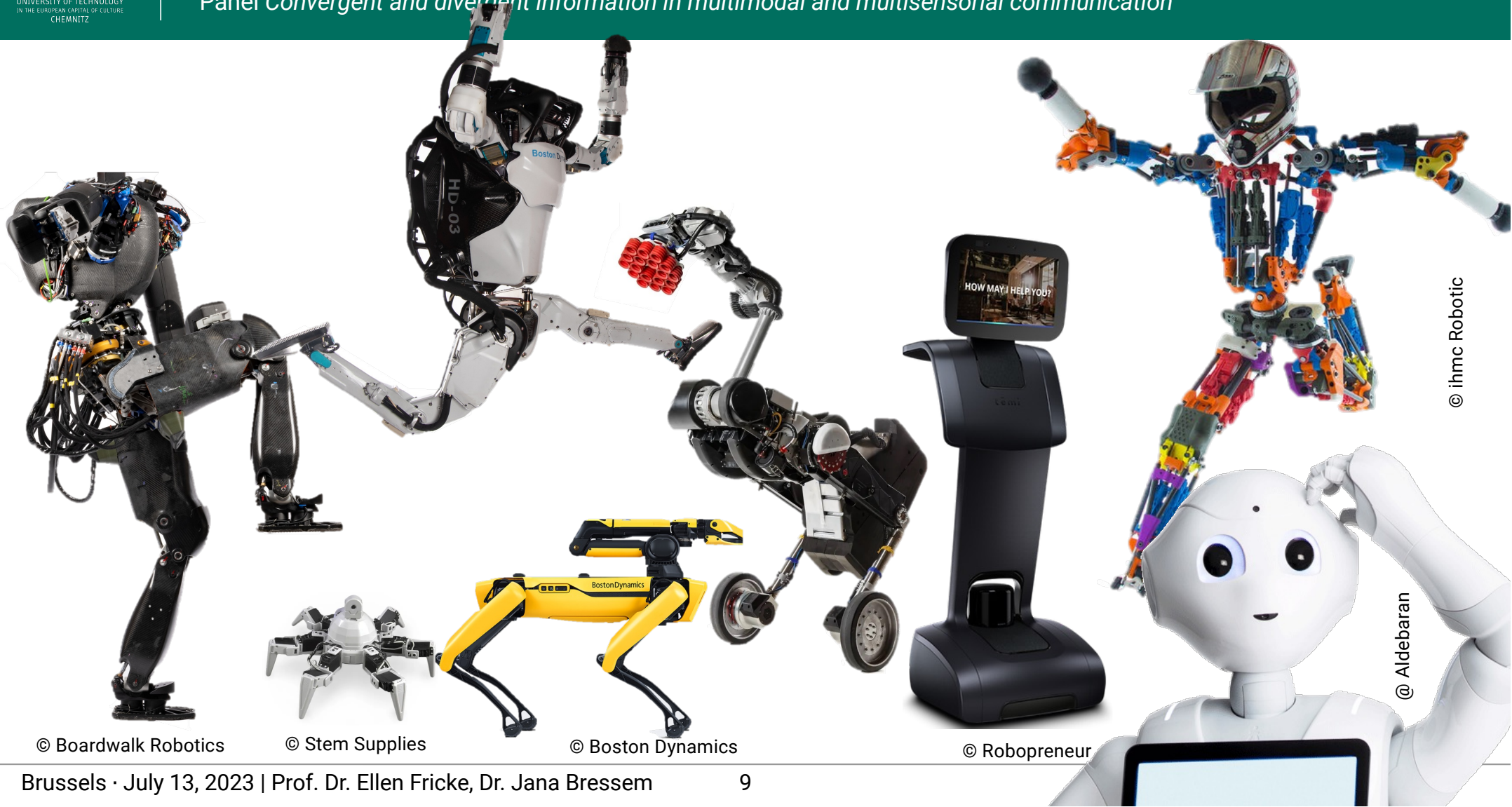
(e.g., Avgustis et al. 2021; Pitsch 2020; Winner et al. 2019)





UNIVERSITY OF TECHNOLOGY  
IN THE EUROPEAN CAPITAL OF CULTURE  
CHEMNITZ

IPrA Conference 2023, July 9–14, in Brussels,  
Panel *Convergent and divergent information in multimodal and multisensorial communication*



© Boardwalk Robotics

© Stem Supplies

© Boston Dynamics

© Robopreneur

@ Aldebaran

© ihmC Robotic

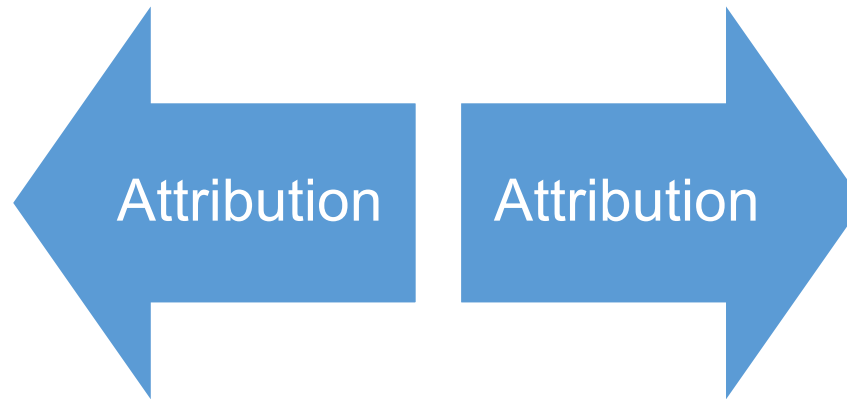


Foto: Tobias Naumann



**Who would you most likely address gesturally?  
And why?**

Foto: Tobias Naumann

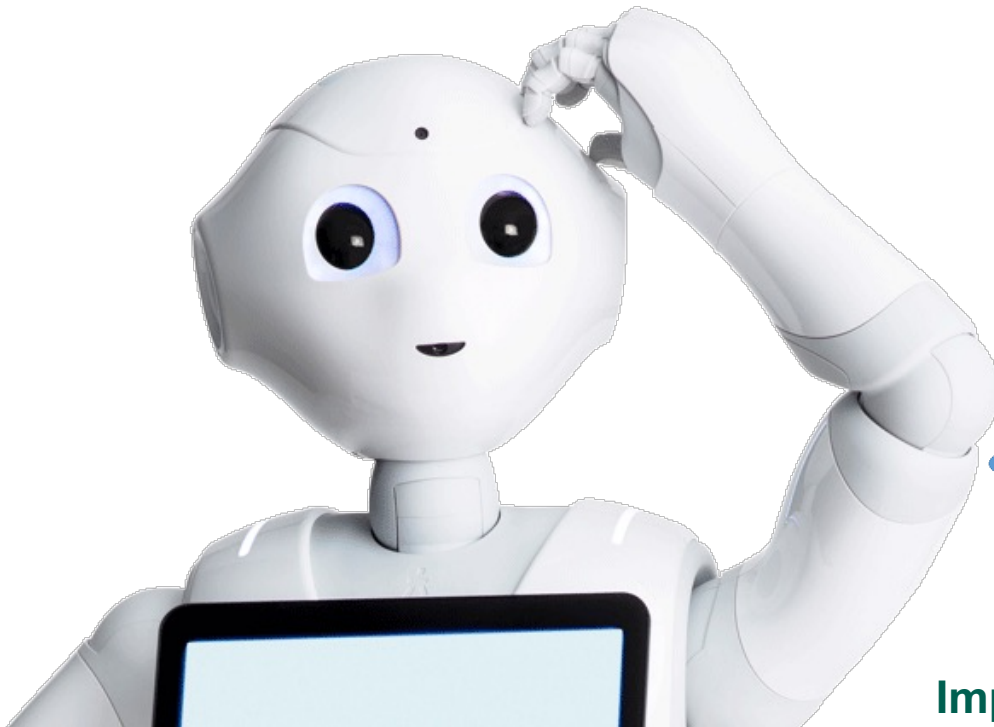


What expectations do we have of a robot compared to human addressees?

What competences do we attribute to it? And how do we adapt our own behaviour accordingly?

How can we tell that a robot can understand gestures, for example?





salute\_left\_b001\_android

salute\_right\_b001\_android

saxophone

saxophone\_a001\_android

scratch\_back\_of\_head\_left\_b001\_android

scratch\_back\_of\_head\_right\_b001\_android

scratch\_top\_of\_head\_left\_b001\_android

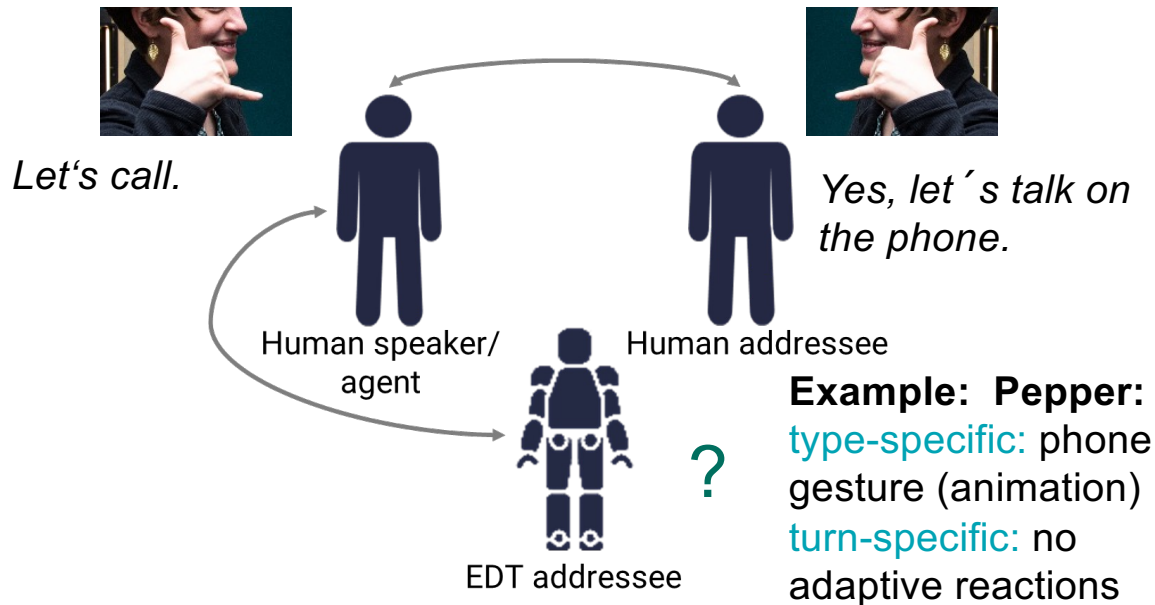
scratch\_top\_of\_head\_right\_b001\_android

scratchbottom\_1

**Important: A robot must signal that it is a robot and what it can do and is capable of.**

<https://www.aldebaran.com/en/pepper>

## DIMENSIONS OF ADDRESSEE DESIGN



### (1) Type-specific:

**Which basic abilities and characteristics does a robot have as a type?**

For example, can it only execute pre-programmed gestural animations? (Pepper)  
 Are different sensor types used (optical vs. sensor-based tracking)?

### (2) Turn-specific:

**How are these abilities and properties instantiated in the interaction process during turn-taking?**

For example, how does a robot indicate its specific intention to a human? Can it indicate it at all? (e.g., constraints to the level of semiotic complexity).

**Addressee Design** describes the adaptation of communicative behavior to the abilities, knowledge, and conditions of an addressee (humans as well as EDTs).

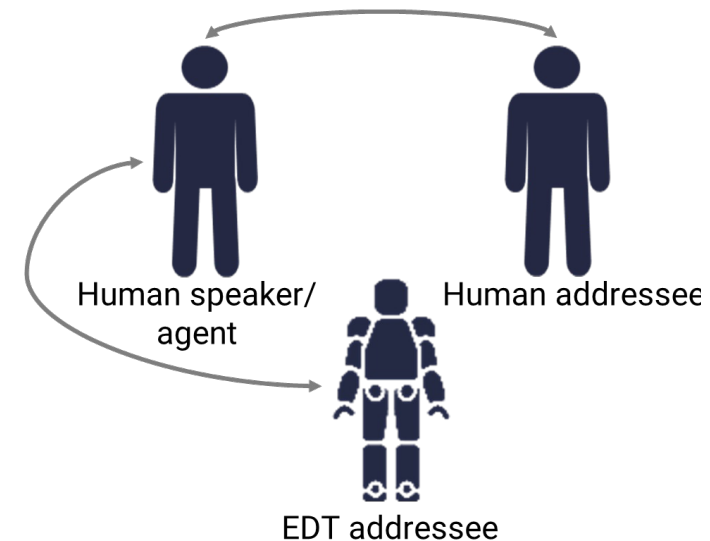
## Potential conflicts of senses and sensors – Creating a gesture set for mixed addressees

**Question:** What can a sensor-based gesture set for gesture control look like that can address a mixed team of humans and robots?

**Problem:** different sensory equipment  
(humans: visual perception,  
robots: sensor data from arm muscle movements)



**Example for illustration:**  
pointing gesture





# Wearable Sensors (muscle movements)

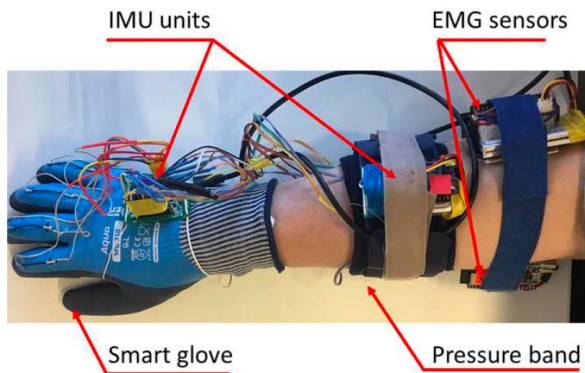


Fig. 4: Designed multi-sensing measurement system

## Classical Vision Based Systems



Leap Motion™

- Sensitive to light and reflective objects
- Limitation of color and surface resolution for skin recognition
- Obstacles hinder the vision from hidden moving object

- Not limited by visual obstacles
- Customizable sensitivity
- Response independent of light exposure
- Attachable to irregular surfaces

What are possible sensor based methods for HGR?  
 What are the advantages and disadvantages of each method?

## Sensor based systems



## RESULTS – GESTURE SET

### Gesture Based Symbiotic Robot Programming for Agile Production

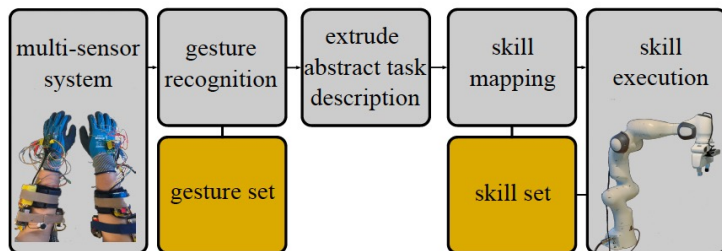
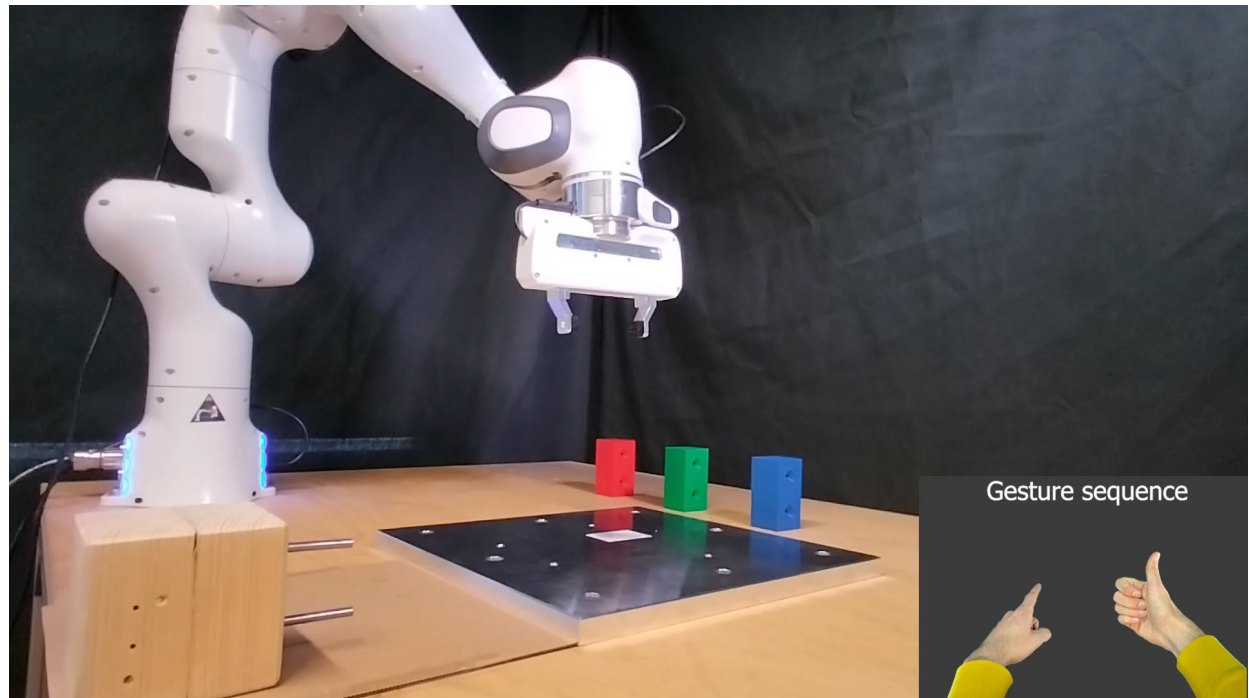



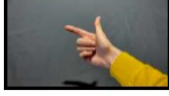










Fig. 1: System architecture



Gäbert, C., Djemal, A., Hellara, H., Atitallah, B. B., Ramalingame, R., Barioul, R., Salzseiler, D., **Fricke, E.**; Kanoun, O., & Thomas, U. (2022). Gesture Based Symbiotic Robot Programming for Agile Production. *2022 IEEE 9th International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, Chemnitz, Germany, 2022 (pp. 1–6). <https://doi.org/10.1109/CIVEMSA53371.2022.9853686>

	Left Hand	Right Hand	Example
Identification	  classifier: select type (object or hole)	  identify object (1-3) or hole (1-2)	 select object 2  hole 1 at selected object
Spatial Relations	  spatial relation (reference or new object)	  spatial relation (reference or new object)	 new block on bottom left  new block on top right

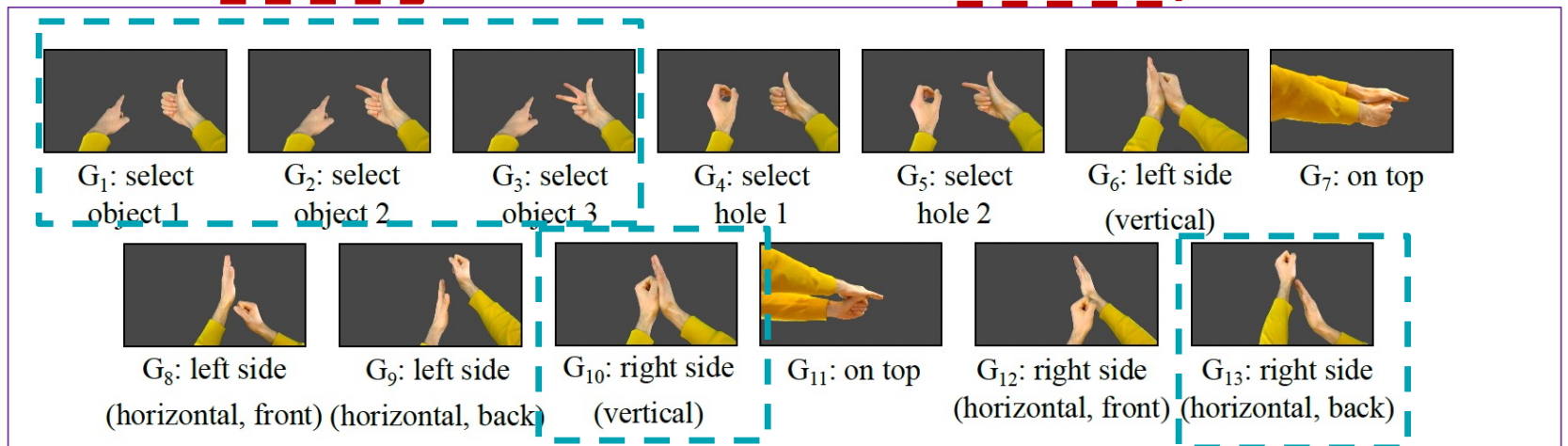
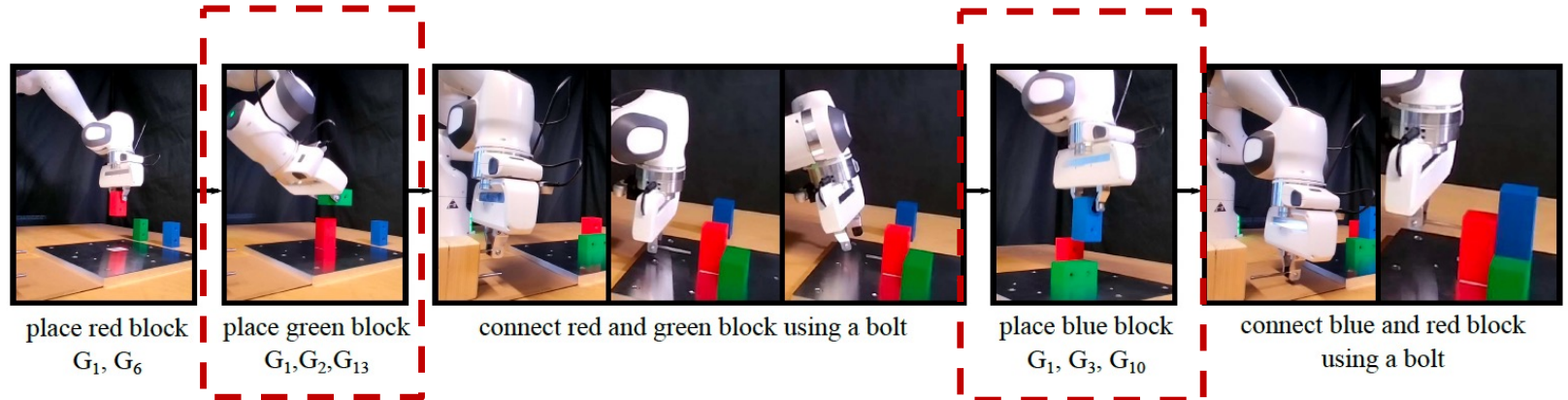
**Novel idea:** simultaneous distribution of different functions to both hands (e.g., selection of particular object types (hole) vs. selection of individual objects (specific hole in cuboid))

**Starting point:** Communicative gestures as a model for gesture control in human-machine interaction

**Semiotic principles** based on

- Basic distinctive form parameters: hand shape, palm orientation, position in gesture space, movement, handedness
- Basic gesture classes such as iconic gestures or pointing gestures
- Efficient construction of a code on the basis of minimal combinatorics
- Distinctive communicative functions, e.g., a) object identification (reference), b) representation of instructions (spatial relations) b) evaluation (e.g., stop, okay)
- Extensibility of the gesture set with regard to more complex tasks

# Gesture set Video stills





# Outlook and conclusion

In hybrid societies, technologically diverse agents interact and share public space.



- Emerging hybrid societies require an **extended concept of addressee design** (type-specific, turn-specific) complementing that of accountability for mixed hybrid group settings.
- Different **sensory capabilities** play a crucial role and have therefore to be considered as dimensions of addressee design and accountability.
- **Mismatches in interaction** can be circumvented by extended design principles also based on linguistic and semiotic expertise.
- Conversely, if there are **advances in sensor technology**, it will be a great advantage for **automatic gesture recognition** in linguistic and semiotic gesture research.



# DATA ANNOTATION – ELAN CODING SCHEME

- Speech - transcription
- Speech - levels of intentionality
- gesture - levels of intentionality
- gesture - type
- gaze - type
- gaze - levels of intentionality
- body - type
- body - levels of intentionality

ELAN 6.5 - ELAN\_2019\_12\_09\_PPaar1\_Kugellabyrinth\_Interaktion.eaf

Future impact of sensor wristbands for automatic segmentation and gesture analysis can be expected. → progress in miniaturization

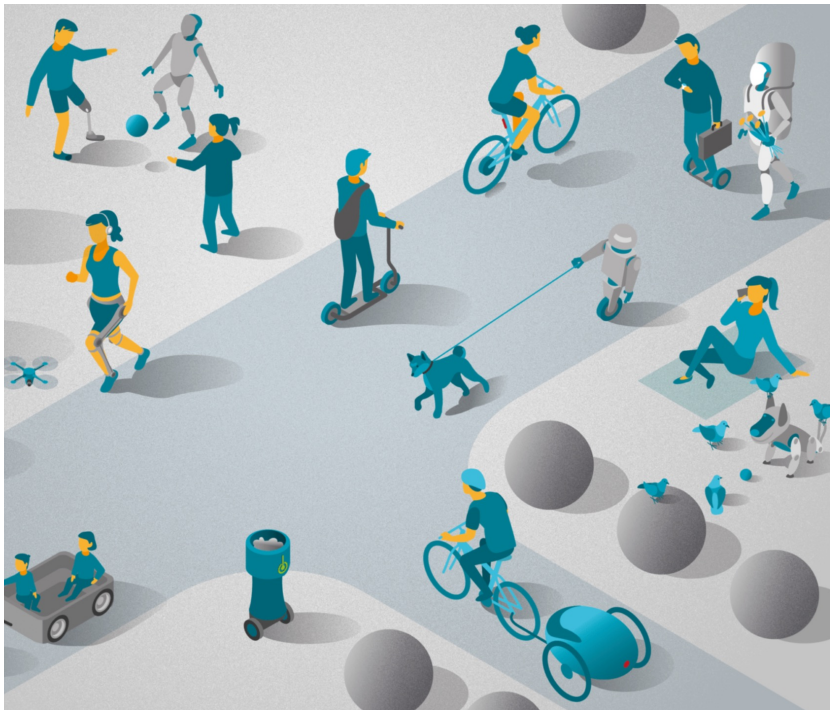
IMU units

EMG sensors

Smart glove

Pressure band

Fig. 4: Designed multi-sensing measurement system



Ellen Fricke



Jana Bressemer

Thank you very much for  
your attention!

Funded by German Research Foundation  
– CRC 1410 – Project-ID 416228727



## Selected References

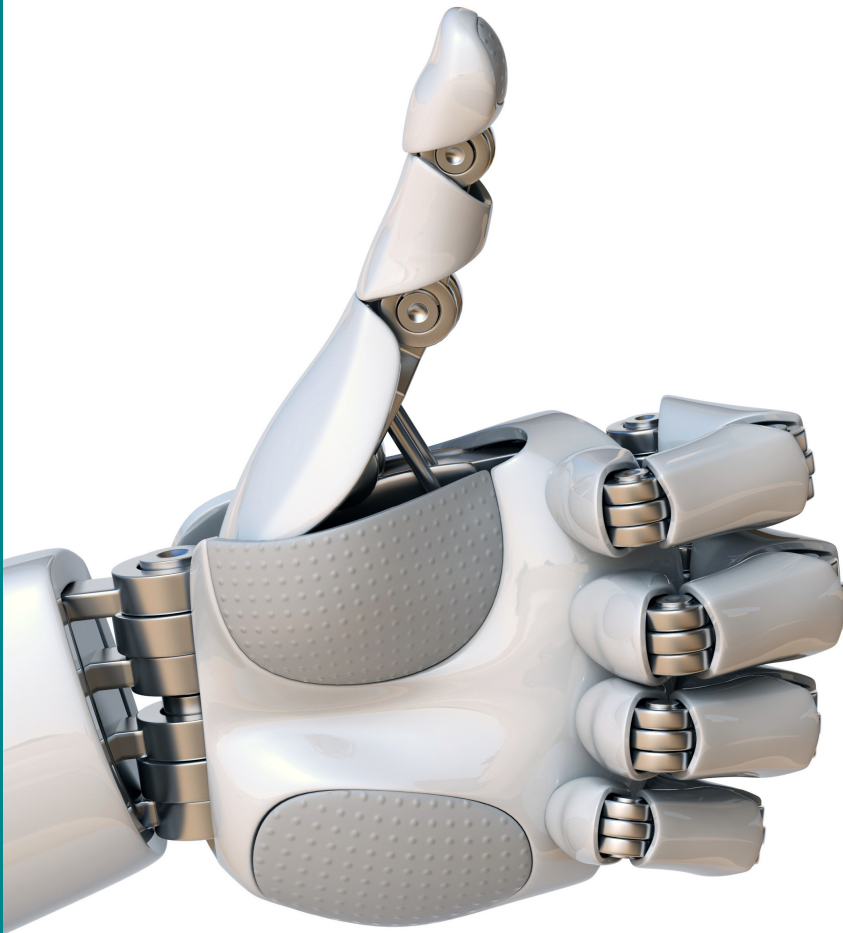
- Djemal, A., Hellara, H., Barioul, R., Atitallah, B. B., Ramalingame, R., **Fricke, E.**, & Kanoun, O. (2022). Real-Time Model for Dynamic Hand Gestures Classification based on Inertial Sensor. *2022 IEEE 9<sup>th</sup> International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, Chemnitz, Germany, 2022 (pp. 1–6). [10.1109/CIVEMSA53371.2022.9853648](https://doi.org/10.1109/CIVEMSA53371.2022.9853648).
- Fricke, E. (2007). *Origo, Geste und Raum: Lokaldeixis im Deutschen*. De Gruyter. <https://doi.org/10.1515/9783110897746>
- Fricke, E. (2012). *Grammatik multimodal: Wie Wörter und Gesten zusammenwirken*. De Gruyter. <https://doi.org/10.1515/9783110218893>
- Fricke, E. (2014). Deixis, gesture, and embodiment from a linguistic point of view. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & J. Bressemer (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction, (Handbooks of Linguistics and Communication Science 38.2)* (pp. 1803–1823). De Gruyter Mouton. <https://doi.org/10.1515/9783110302028.1803>
- Fricke, E. (2021a). Multimodality and Semiotic Complexity from a Linguistic Point of View: Processes of Code Integration and Code Manifestation. *OBST, 99 (Osnabrücker Beiträge zur Sprachtheorie, 99, Special Issue Linguistik und Multimodalität / Linguistics and Multimodality)*, 57–86.
- Gäbert, C., Djemal, A., Hellara, H., Atitallah, B. B., Ramalingame, R., Barioul, R., Salzseiler, D., **Fricke, E.**; Kanoun, O., & Thomas, U. (2022). Gesture Based Symbiotic Robot Programming for Agile Production. *2022 IEEE 9<sup>th</sup> International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, Chemnitz, Germany, 2022 (pp. 1–6). <https://doi.org/10.1109/CIVEMSA53371.2022.9853686>

## Selected References

- Hellara, H., Djemal, A., Barioul, R., Ramalingame, R., Atitallah, B. B., **Fricke, E.**, & Kanoun, O. (2022). Classification of Dynamic Hand Gestures using Multi Sensors Combinations. *2022 IEEE 9<sup>th</sup> International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, Chemnitz, Germany, 2022 (pp. 1–5). <https://doi.org/10.1109/CIVEMSA53371.2022.9853694>
- Müller, Michael R. (2023). "Social Displays. Creating Accountability in Robotics". *Österreichische Zeitschrift für Soziologie*, 49(2).
- Pitsch, K. (2023). Mensch-Roboter-Interaktion als Forschungsinstrument der Interaktionalen Linguistik. In M. Meiler, & M. Siefkes (eds.), *Linguistische Methodenreflexion im Aufbruch. Beiträge zur aktuellen Diskussion im Schnittpunkt von Ethnographie und Digital Humanities, Multimodalität und Mixed Methods*. De Gruyter. <https://doi.org/10.1515/9783111043616-005>
- Pitsch, K., Dankert, T., Gehle, R., & Wrede, S. (2016). Referential practices. Effects of a museum guide robot suggesting a deictic 'repair' action to visitors attempting to orient to an exhibit. *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 225–231). <https://doi.org/10.1109/ROMAN.2016.7745135>
- Sacks, H. (1992). *Lectures on Conversation*. Blackwell.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A simplest systematics for the organization of turn taking for conversation. *Language*, 50(4), 696–735. <https://doi.org/10.1016/B978-0-12-623550-0.50008-2>
- Schmitt, R., & Deppermann, A. (2009). „damit sie mich verstehen“: Genese, Verfahren und recipient design einer narrativen Performance. In J. G. Schneider, M. Buss, F. Liedtke, S. Habscheid, & S. Jautz (eds.), *Theatralität des sprachlichen Handelns: Eine Metaphorik zwischen Linguistik und Kulturwissenschaften* (pp. 179–112). Brill Fink. [https://doi.org/10.30965/9783846747933\\_007](https://doi.org/10.30965/9783846747933_007)
- Schmitt, R., & Knöbl, R. (2014). *Recipient Design: Zur multimodalen Repräsentation des Anderen im eigenen Verhalten*. Institut für Deutsche Sprache. <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/3239>
- Winner, T., Selen, L., Oosterwijk, A. M., Verhagen, L., Medendorp, W. P., van Rooij, I., & Toni, I. (2019). Recipient Design in Communicative Pointing. *Cognitive Science*, 43(5). <https://doi.org/10.1111/cogs.12733>



TECHNISCHE UNIVERSITÄT  
IN DER KULTURHAUPTSTADT EUROPAS  
CHEMNITZ



# GESTERN

gestern  
heute  
übermorgen

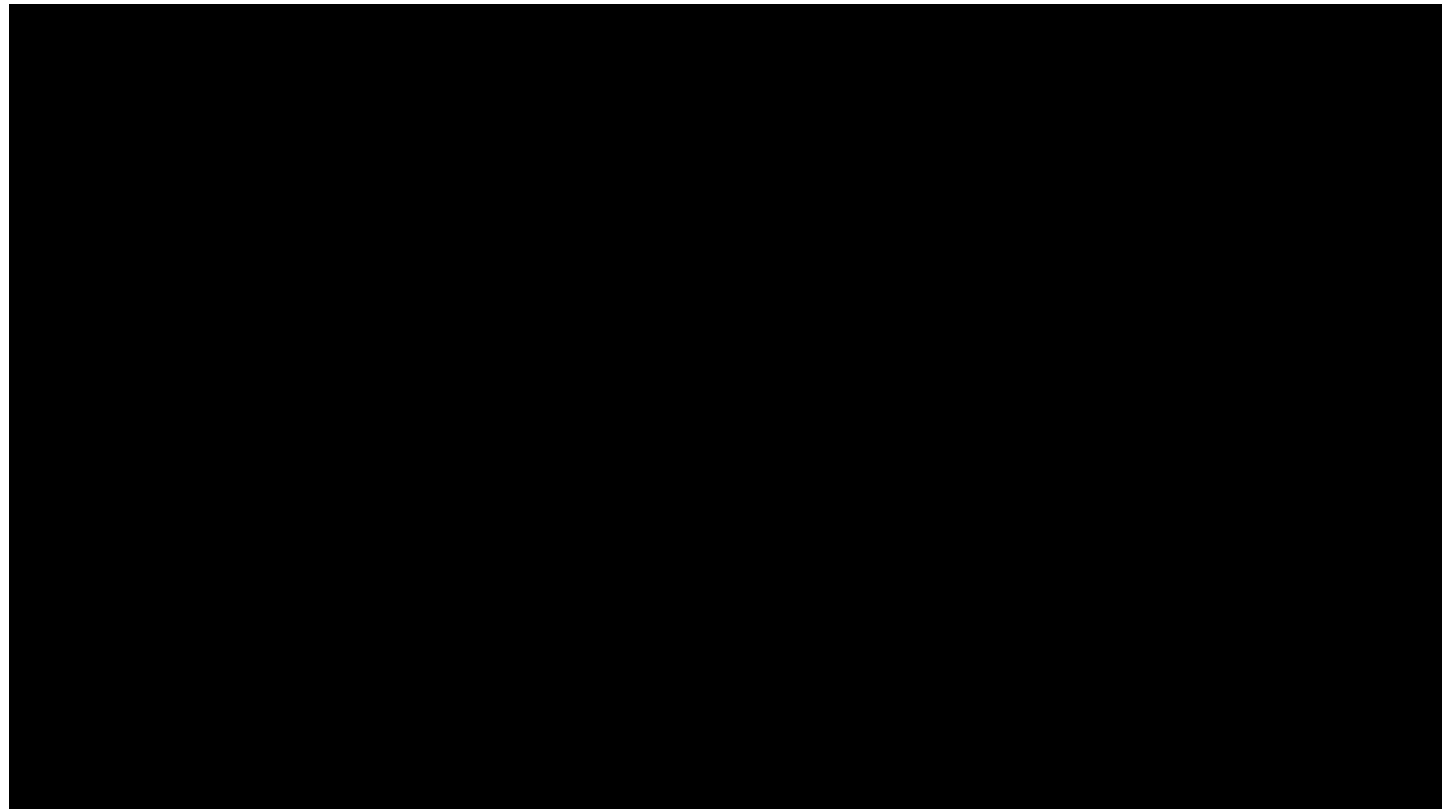


## Gesture exhibition

Chemnitz Museum of Industry

[www.gesten-im-museum.de](http://www.gesten-im-museum.de)

[www.manuact.org](http://www.manuact.org)





TECHNISCHE UNIVERSITÄT  
IN DER KULTURHAUPTSTADT EUROPAS  
CHEMNITZ



# GES TEN

gestern  
heute  
übermorgen