

Conversational systems and multimodal coordination dynamics

Stefan Kopp

Social Cognitive Systems Group
Faculty of Technology, Bielefeld University
Center of Excellence „Cognitive Interaction Technology“ (CITEC)



Face-to-face conversational interaction

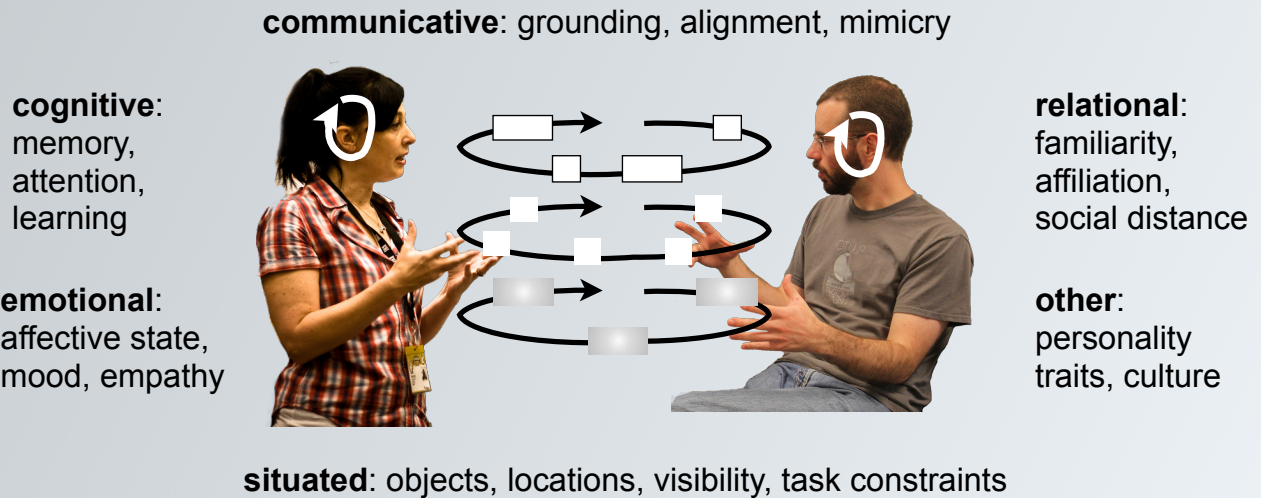


Communication — Emotion — Rapport — Situated Cooperation

→ Multimodal behavior

speech, intonation, facial expression, body and hand gesture, gaze, ...

Multimodal coordination dynamics



Interactive view — focus on dynamic coordination processes

- ▶ multimodal behavior arises from, and steers forward multiple, interrelated coordination processes between and within interactants
- ▶ crucial role of prediction, planning, perception, adaptation

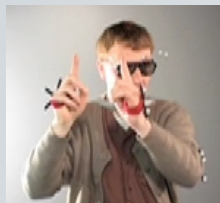
3

Example

„it has two tall towers, to the right and to the left“



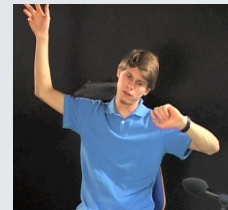
„two tall towers“



„with two towers“



„church towers“



unfolding interaction

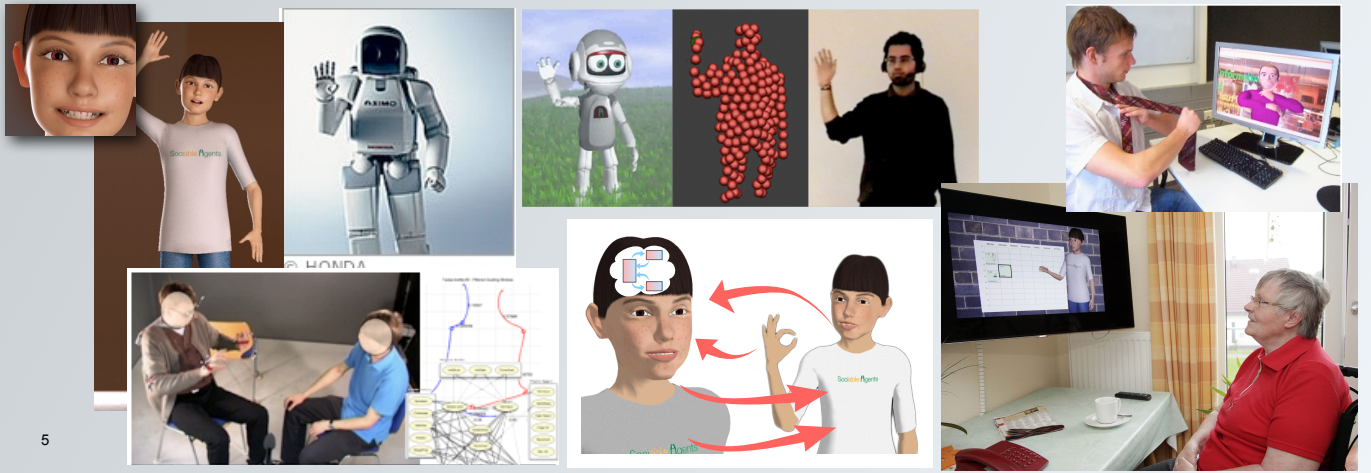
Interactive, incremental reduction of multimodal signals

→ interplay of communicative and cognitive coordination and adaptation

(Clark & Wilkes-Gibbs 1986; Hoetjes et al. 2011; Galantucci & Steels 2008)

4

- ▶ Assistive systems and companions with abilities for conversational social interaction and cooperation
- ▶ Learning from human behavior and embodied cognitive mechanisms of social perception, reasoning, and action
- ▶ Focus: prediction, adaptation and learning in social interaction



Dialogue coordination— communicative feedback

Giving feedback as listener

- incremental interpretation
- assessment of own understanding
- FB selection and placement

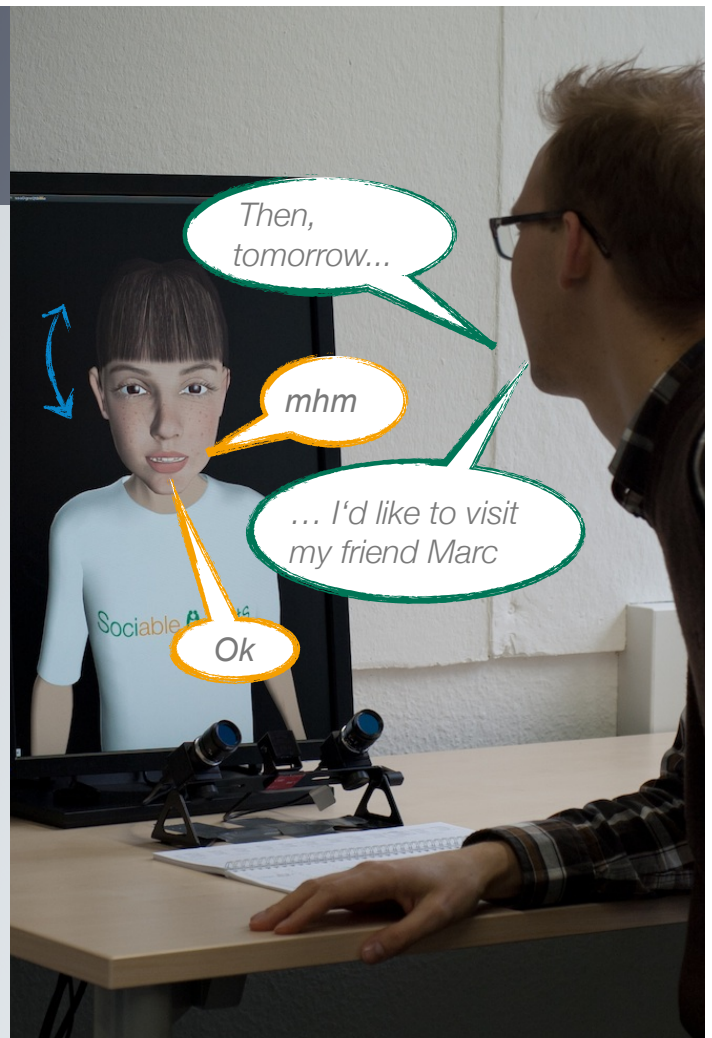
(Kopp et al. 2007)

Related systems

rapport agent (Gratch et al. 2006)

backchannel prediction (Morency et al. 2008)

EU project SEMAINE (2008-2010)



Dialogue coordination— communicative feedback

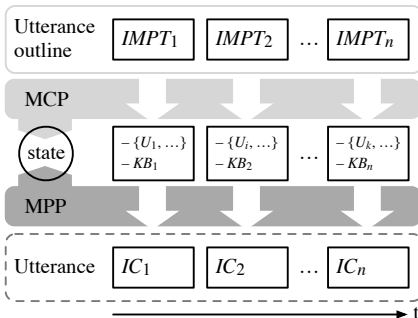
Adapting to the dialogue partner as speaker

- elicit, recognize and interpret listener feedback
- infer mental state
- online behavior adaptation
- incremental generation and synthesis

(Buschmeier & Kopp, 2011, 2012)



Incremental Natural Language Generation



Behavior Markup Language (extended)

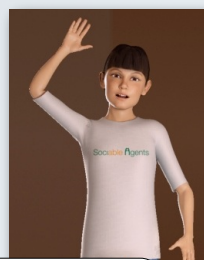
```

<bml id="speechchange">
  <bmla:parametervaluechange
    target="bml1:speech1"
    paramId="volume"
    start="bml1:speech1:sync1"
    end="bml1:speech1:sync1+1">
    <bmla:trajectory type="linear"
      targetValue="90"/>
  </bmla:parametervaluechange>
</bml>

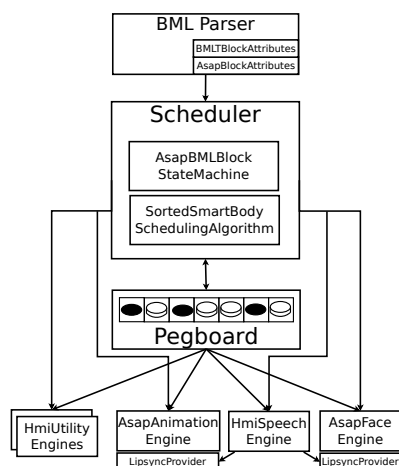
<speech id="speech1" start=
  "anticipators:speechStopAnt:stop-0.1">...
</speech>
  
```



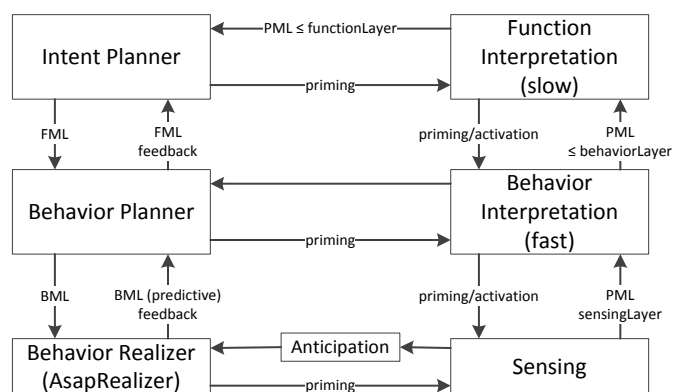
© HONDA



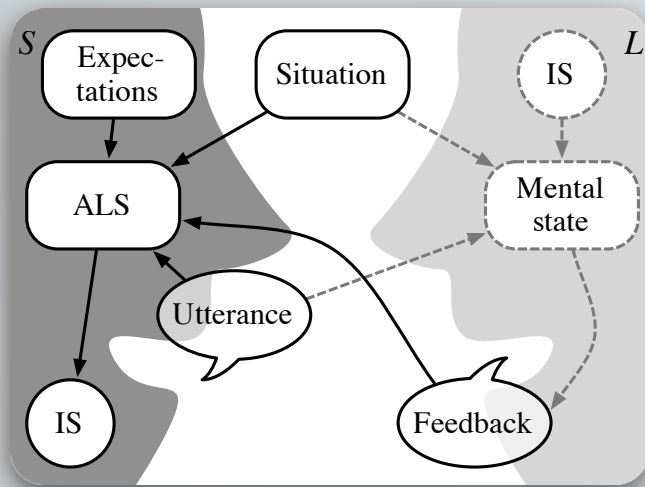
Realtime Behavior Realizer



Incremental Processing Architecture (ASAP)



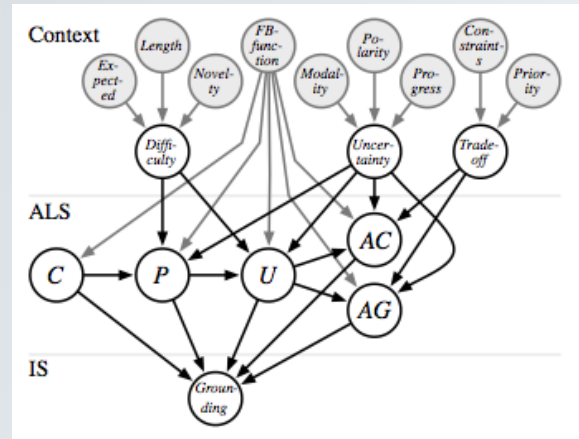
Dialogue coordination — minimal mentalizing and predictive adaptation



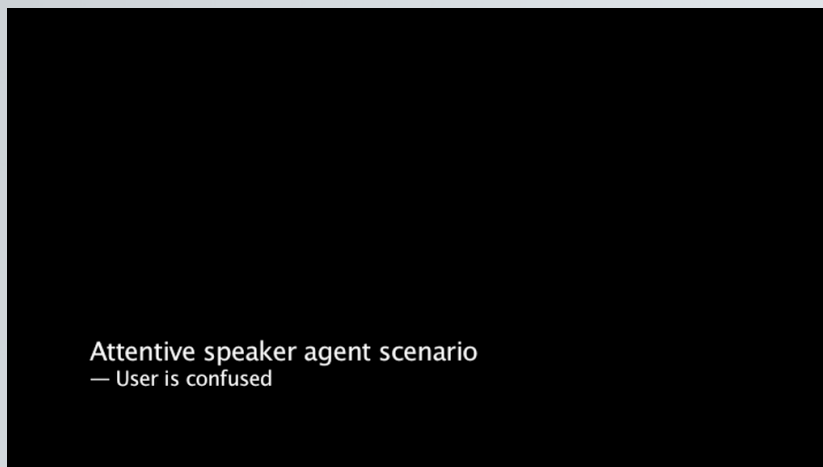
(Buschmeier & Kopp 2012, 2013)

Bayesian model construction and simulation

(Allwood et al. 2000)



Dialogue coordination — minimal mentalizing and predictive adaptation



ECAs as virtual assistants in daily life for elderly or mentally handicapped users

- ▶ schedule management, video communication
- ▶ robust and adaptive dialogue, grounding and repair



Study with WoZ-based systems and different repair strategies

11

(Yaghoubzadeh et al., 2013, Kramer et al. 2013)

Cognitive dynamics — intermodal coordination speech and gesture

Speech and gesture interaction

- ▶ timing, form, semantics, pragmatics
- ▶ develop together, break down together

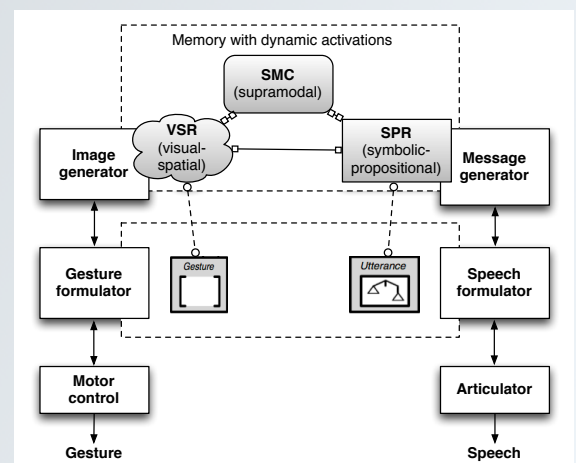
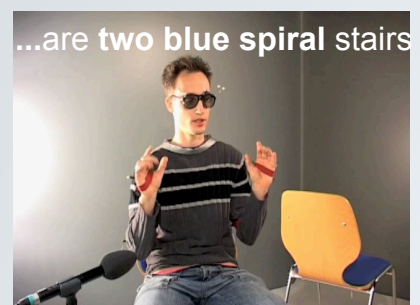
Computational production model

- ▶ from conceptualizing to overt behavior
- ▶ model-based and data-based

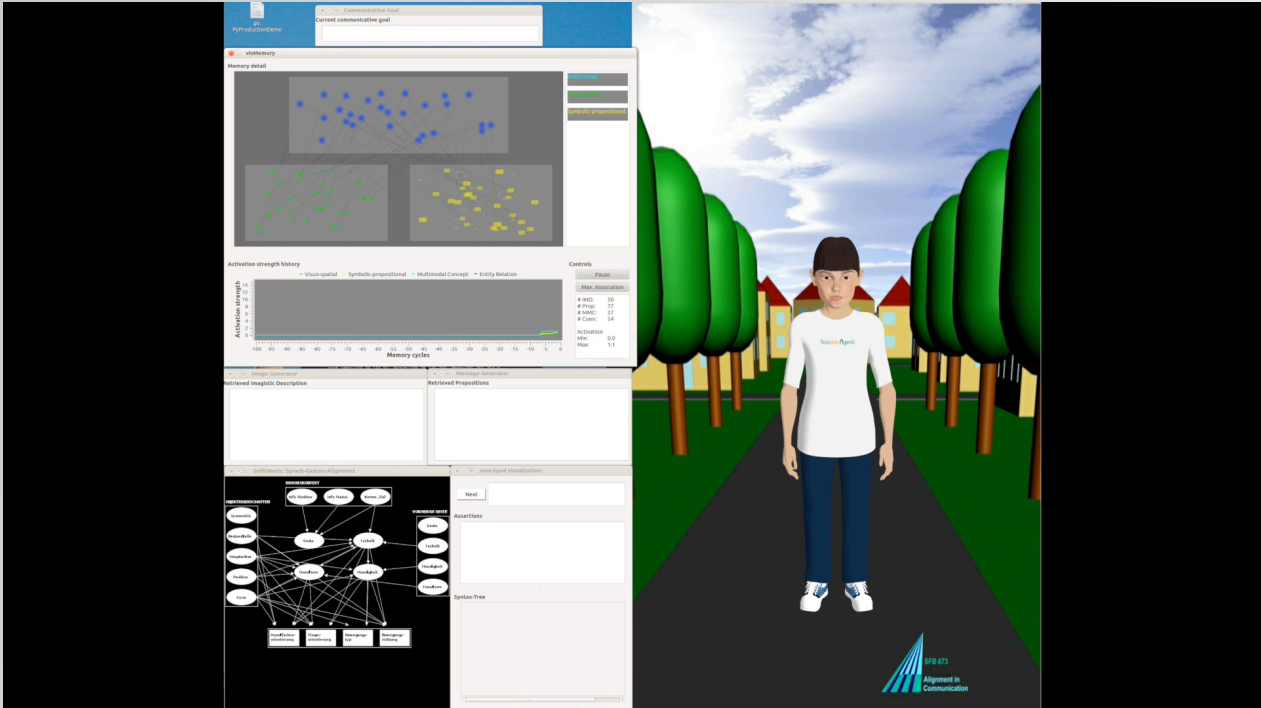
Underlying cognitive dynamics?

„non-redundant gesture-speech combinations occur because mental images are more active in speakers minds at the moment of speaking than are verbal codes“ (Hostetter & Alibali 2011, p.45)

(Bergmann & Kopp 2009; Lücking et al. 2013)

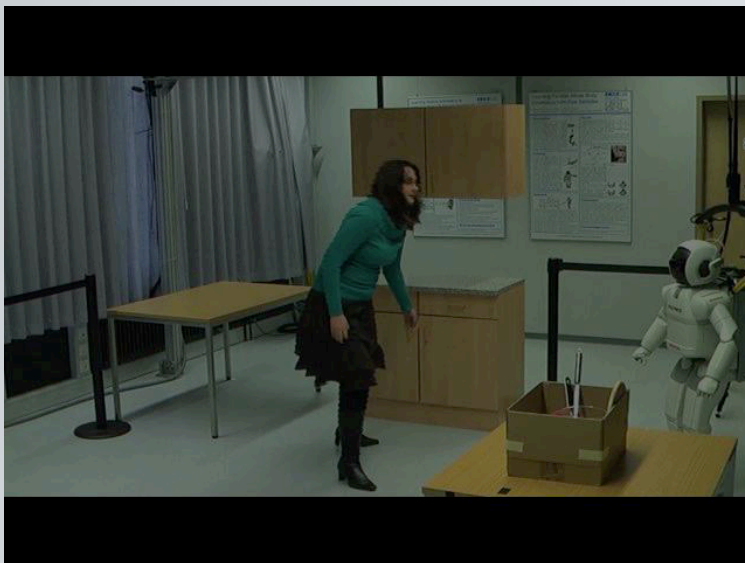


12

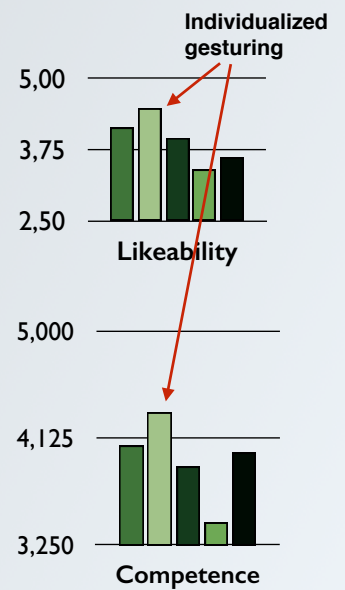


13

On virtual and robotic agents



(Salem et al., 2011, 2012)



(Bergmann et al., 2010)

14

Linguistic

- ▶ prosodic, dialect, lexical, syntactic, semantic
- ▶ audience design, priming, two-stage accounts
(Brennan et al. '10, Pickering & Garrod '04, Keysar et al. '98)

Gestural

- ▶ between co-narrators (Kimbara '06, '08)
- ▶ in re-tellings, especially of meaningful features
(Parrill et al. '06, Mol et al. '11)
- ▶ in face-to-face dialogue (Holler & Wilkin 2011; Bergmann & Kopp 2012)

Social, cognitive and communicative effects

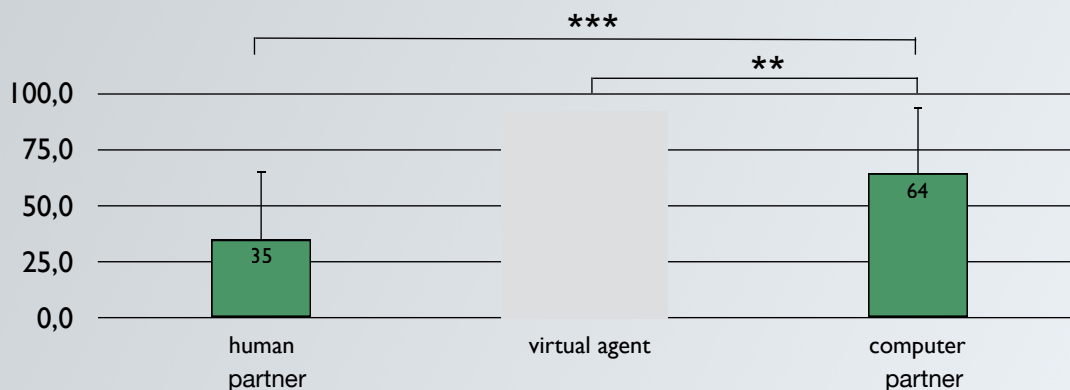


Interactional dynamics — lexical alignment with virtual humans

Alignment strength



(Branigan et al. 2010)



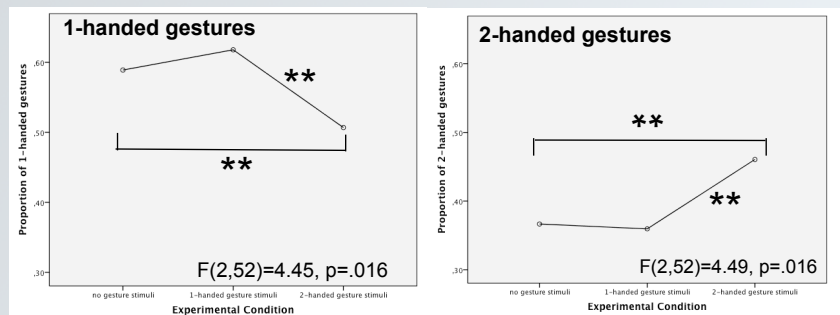
$F(2,87)=8.04, p=.001$

Interactional dynamics — gestural alignment with virtual humans

Tangram description game, human and virtual agent (WoZ) alternating turns

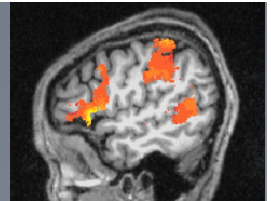


Three conditions:
2-handed vs. 1-handed vs. no gestures

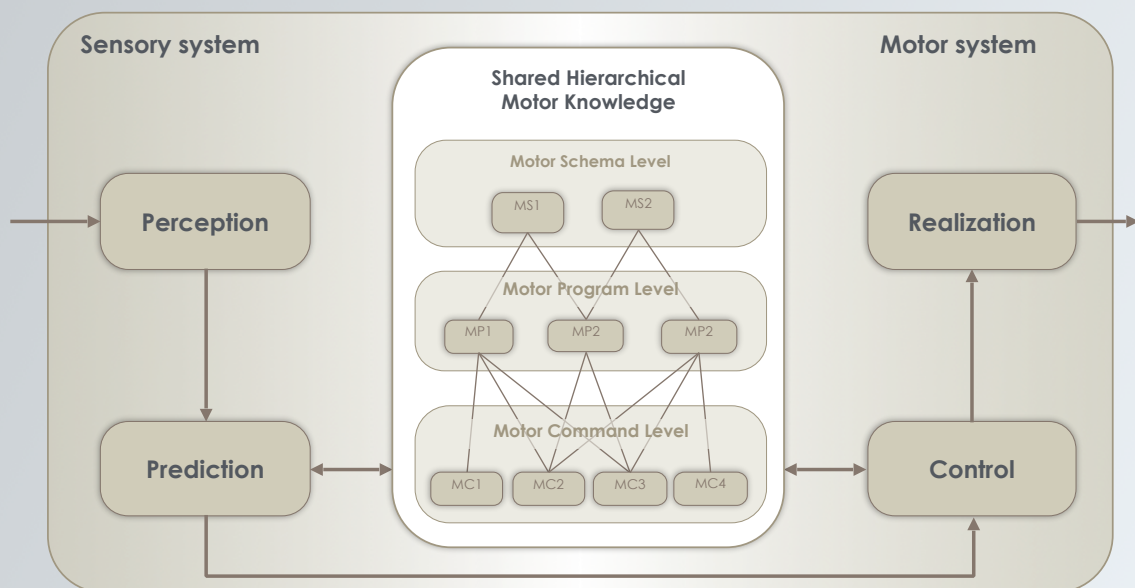


17

Interactional and cognitive dynamics — perception and action coupling



(Montgomery et al. 2007)



Using internal simulation based on own sensorimotor expertise for prediction-based perception and understanding

18

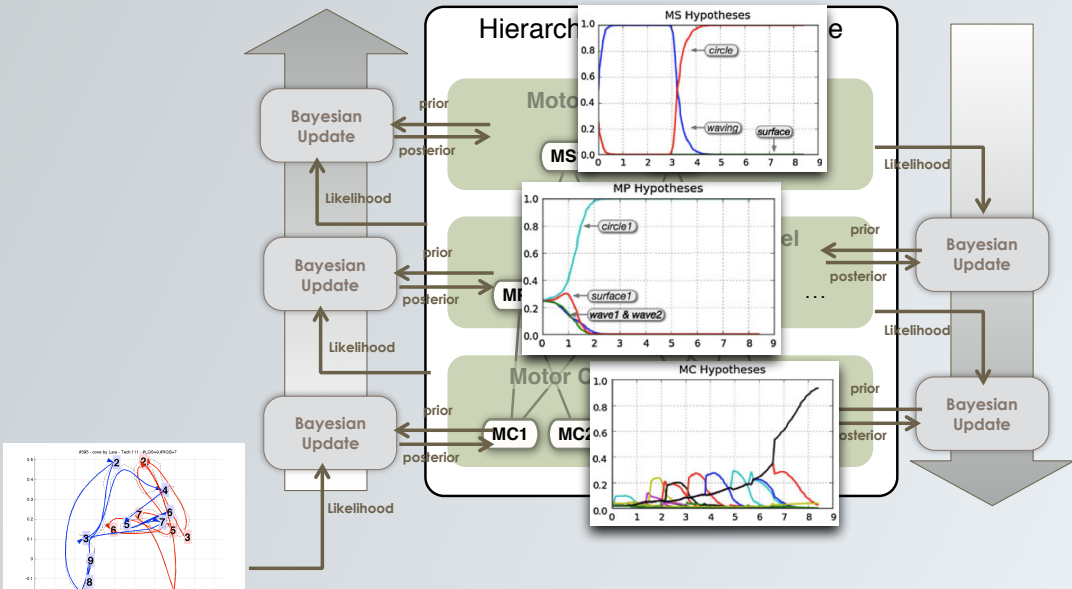
(Sadeghipour & Kopp 2011, 2014)

Interactive and cognitive dynamics — perception and action coupling

Hierarchical Bayesian Belief Update

- ▶ incremental online processing — auto-completion, imitation, ...
- ▶ from perception to understanding, concurrent and interacting

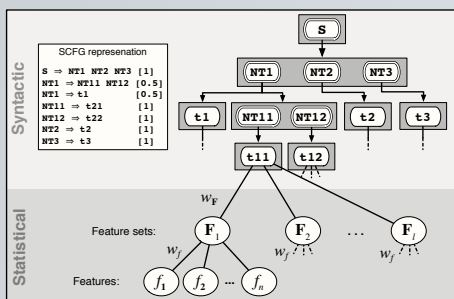
Bottom-up update Top-down prediction



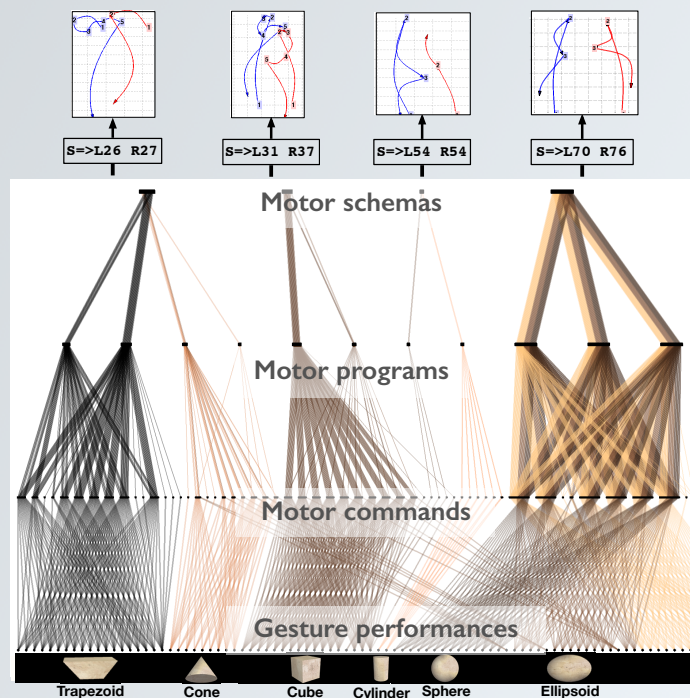
Interactive and cognitive dynamics — learning schema representations

Feature-based Stochastic Context-free Grammar (FSCFG)

- ▶ clustering (unsupervised)
- ▶ classification (supervised)



(Sadeghipour & Kopp 2014a, 2014b)



Intelligent Coaching Space

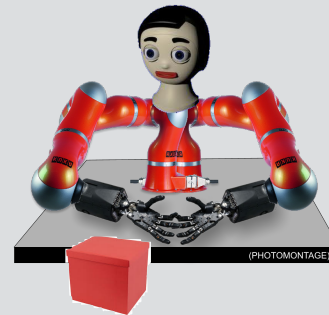
VR-based closed-loop interaction and training system for motor skill learning or rehabilitation



→ Incremental online feedback through mirror or virtual coach

Deep Familiarization Grounded in Manual Action and Language

Autonomous guided familiarization with novel objects and affordances



→ Mental simulation as embodied language processing context

21

Thank you very much for your attention.

Acknowledgments

Kirsten Bergmann
Amir Sadeghipour
Hendrik Buschmeier
Ramin Yaghoubzadeh
Herwin van Welbergen
Sebastian Kahl
Maha Salem
Matthias Priesters
Hannes Rieser

scs.techfak.uni-bielefeld.de

22