DESIGNERLY
HRI
KNOWLEDGE

Intermediate-Level Knowledge: A Conversation Analysis Perspective

Hannah Pelikan,

Linköping University, Linköping, Sweden, hannah.pelikan@liu.se

Abstract —When designing for and evaluating interaction with specific robots, we intuitively draw on knowledge about interaction with other humans and with other robots or machines. Much of our preconceptions and lay observations about interaction are rarely spelled out in HRI research. As I will demonstrate in this paper, our knowledge about basic human interaction patterns can be considered intermediate-level knowledge. Using the example of greeting patterns, I introduce multimodal conversation analysis both as a body of research that intermediate-level knowledge can be built upon and as a tool for exploring and critically reflecting knowledge about interactional patterns.

Keywords — conversation analysis, intermediate-level knowledge, human-robot interaction.

1. INTRODUCTION

Interacting with others – humans or machines – is a fundamental aspect of everyday life. In conversation analysis and related social science disciplines, it is well established that humans do not act randomly but rather follow stable patterns [1]. When evaluating and designing for interaction with robots, we draw on our tacit knowledge about such interaction patterns. For instance, in our everyday lives we constantly face interactional challenges like how to open a conversation with a stranger, how to instruct someone in using the dishwasher, and how to behave in a traffic situation that is not covered by the rules.

We smoothly resolve such issues in human-human interaction. For instance, we friendly greet the stranger before moving into more serious discussions and gesture to a pedestrian or other car to pass in front of us. When designing for human-robot interaction, we draw on this knowledge, gained through being competent members of society. While this knowledge about interaction generalizes across more than one particular situation, it remains tacit and often too specific to be considered a full-fledged theory in itself. Höök and Löwgren [2] characterize knowledge in between theory and single cases as *intermediate-level knowledge*. Interaction patterns generalize across contexts, while remaining grounded in specific examples. Located in between abstract theory and specific design or interaction examples, I suggest that designerly knowledge about general interactional practices and patterns can be regarded as a form of intermediate-level knowledge.

Treating knowledge about interactional patterns as intermediate-level knowledge will enable the HRI community to discuss preconceived ideas about interaction that would otherwise remain implicit. Considering our everyday knowledge about interaction as intermediate-level knowledge is simultaneously embracing and challenging our gut feelings about how interaction works. While we may often be right in our general idea of how people do things with other humans and machines, our intuitive knowledge may not always fully apply to the specific context that we are designing for. Considering designerly knowledge about interactional patterns as intermediate-level knowledge does not mean that we blindly accept our gut feelings as scientific concepts but rather stresses that we should critically engage with them. How can we do so?

This paper presents *multimodal conversation analysis* (CA) as a suitable approach for engaging with our intuitive ideas about interaction. Offering a format to systematically describe and discuss interactional patterns, multimodal conversation analysis can support (1) formulating our tacit knowledge and (2) critically reflecting and rectifying it in a scientific way. Conversation

analysis can be applied at a small scale, even using single examples, which makes it particularly suitable for iterative design processes.

In the following, I will briefly introduce multimodal conversation analysis. Subsequently, I will introduce greetings as an example of a basic interactional pattern that has been extensively researched in human-human and human-robot interaction. Finally, I will discuss how designers and HRI researchers can practically engage with conversation analysis when formulating and discussing intermediate-level knowledge.

2. CONVERSATION ANALYSIS

Conversation analysis is often referred to as a micro-sociological method and generally stands in a constructivist tradition. The main objective is the study of social interaction and its organization in fine detail, typically working on video recordings of human interaction. Focus lies on how participants demonstrably make sense of each other's actions. Conversation analysis focuses on the sequential organization of action, studying how participants through a current action display their understanding of a previous action and make particular next actions relevant. Actions are situated and dynamically adjusted rather than indifferently following a pre-determined plan. This concept of situated interaction has been introduced to HCI through the work of Lucy Suchman [3]. Participants constantly react to each other, and new actions always build on the context of the previous ones. In CA, language is seen as action and as tightly intertwined with the body [7]. Actions are studied as holistic, *multimodal Gestalts* that may involve language, gesture, gaze, body postures, movement and embodied manipulations of objects [8]. In treating embodiment as a crucial element in social action, conversation analysis shares theoretical foundations with Paul Dourish's [9] and Kia Höök's [10] perspectives on the body as being deeply entangled with our sense-making practices.

An important conversation analytic tool are detailed transcripts of interaction, in which talk, movement, gaze, gesture and engagements with material objects are annotated. These transcripts enable us to pinpoint the sequential unfolding of interaction and serve as a shared resource for joint analysis. A transcript illustrates and captures phenomena of interest but at the same time supports contrasting of different interaction snippets and stimulates discussion about what can be seen in a particular video snippet. It should be stressed here that transcripts are never final and fixed but they are part of the analysis and serve to stimulate discussion.

To date there have been a few attempts at combining CA and interaction design (see e.g. [11], [12]), and CA is a growing approach in HCI and HRI (see e.g. [13], [14], [15]).

3. INTERACTIONAL PATTERNS

When deciding what a robot (or virtual agent) should say or do, we often draw on our intuitive understanding of language and interaction. In doing so, we include knowledge that we have as members of society of how to behave competently and appropriately in social interaction. Consider an encounter between a human and a robot: How would you start the interaction?

Most certainly the first thing you would let the robot say would not be instructions for a joint task but you would probably start with a "Hello" or "Hi" and some introductory phrases. This is how we often open interaction in human-human encounters [16]. Interactional patterns are often described as *practices* [17] in conversation analytic work. In the following I will present greetings as an example of a basic interactional pattern that has been studied both in human-human and human-robot interaction.

3.1. Greetings in Human-Human Interaction

A number of practices related to the beginnings, or *openings*, of interactions have been described by conversation analytic research (see e.g. [16], [17]). In many cases, interactional openings involve greetings, which can take a variety of formats. The following excerpt (Figure 1) illustrates how two people may greet each other on the phone.

01		((phone rings))	SUMMONS
02	Bee	H'llo:?	ANSWER
03	Ava	(h)Hi:,	GREETING
04	Bee	Hi:?	RETURN GREETING
05	Ava	(h)Howuh you:?	QUESTION

Figure 1. Greetings at the beginning of a phone call, adapted from [17]. Transcription symbols: : lengthening of a sound, ? rising intonation, , level intonation, (h) breathiness.

Note that there is not only one person that is greeting, but Bee and Ava are greeting each other. You might not be very surprised to see this – after all, you have surely done this yourself many times. Conversation analysis categorizes a reciprocal greeting as in lines 03-04 as an *adjacency pair* [17], finding that one greeting typically does not stand alone but if Ava greets Bee, this makes a return greeting relevant. If Bee would instead stay silent, this would be problematic for progress of the interaction. Note that while greetings may often be verbalized, they do not necessarily have to be. For instance, "waving hello" is also a greeting action.

3.2. Greetings in Human-Robot Interaction

Greetings have been explored in robots, for instance with a Nao robot [18] and human greetings to a robot have been suggested to indicate that a user treats a robot as social [19]. However, it is not always humans who greet first, but often robots are (intuitively or deliberately) programmed to greet as a first action. You might now think of robots that have natural speech generation engines like Pepper, Nao or Furhat. However, greeting patterns are relevant for all kinds of robots. For instance, the non-speech sounds that toy robot Cozmo produces can be treated as greeting actions. In the following excerpt (Figure 2), Cozmo is just "waking up" and then produces a sequence of sounds ("wa?" - "oh" - and on leaving its charger "dadu?").

01	Cozmo	((waking up)) wa?	
02	Wife	er hat was gesagt	
		he said something	
03	Cozmo	oh	
04	Wife	oh	
05	Cozmo	((leaving charger)) dadu?	GREETING
06		(0.8)	
07	Husband	serv+us	RETURN
		hi (in dialect)	GREETING
08		(1.6)	
09	Husband	hat der servus gesacht?	
		did he say hi?	
10	Husband	((gazes@wife))	
11	Husband	((gazes@researcher))	

Figure 2. Greetings at the beginning of the encounter between a German couple and a Cozmo robot E18-12-30 [01:45-02:05]. Translation in italics. Transcription symbols: ? rising intonation, (0.8) silence in seconds.

One of the participants, a German adult is responding to the robot's "dadu?" sound (l. 05) with "Servus" (l. 07), a greeting that is used in the Bavarian part of Germany and in Austria. As the robot is not continuing but stays idle, he asks "did he say Servus?" (l. 09), looking at his wife (l. 10) and subsequently the researcher (l. 11). He is thereby questioning his initial interpretation since the robot does not respond to his friendly greeting.

Greetings and mutual recognition are important elements in interaction with a simple robot like Cozmo. A particularly engaging element is for instance Cozmo's ability to learn faces and then "greet" people by saying their name, whenever Cozmo recognizes their face. In my data, people often respond to that by saying "hi" to the robot. This illustrates that a simple sound (with a relevant prosodic contour, as in the transcript) or just a name can also be treated as a greeting.

Similarly, greetings have been explored in non-anthropomorphic robotic objects that interact through small movements (see e.g. the Greeting Machine [20]). HRI work already draws on social science when building theories about social interaction. However, these theories are often abstract and do not provide much guidance on how to design for local, real-world contexts. In designing specific robot actions, designers and programmers then draw on their own (often tacit) knowledge about human interaction.

Becoming more aware of our own preconceptions and engaging more systematically with our knowledge about interaction is crucial for designing better, more socially adaptive robots. Conversation analytic research offers both specific examples (through detailed transcripts of interaction) and theories that are grounded in these real-world observations. It can thereby be drawn upon in building more coherent intermediate-level concepts. For instance, conversation analytic transcripts of greeting sequences illustrate what humans say and do when greeting each other and what robot actions they are willing to accept as a greeting. Identifying adjacency pairs can help in recognising that a greeting by the robot should be followed by a human return greeting. This can help in evaluating whether humans treat the robot's nonverbal actions as a greeting. In that sense, interactional patterns are more than just an observation and have a degree of predictive power, (i.e., that in human-human interaction, greetings should be responded to by a return greeting).

4. USING CA TO ENGAGE WITH INTERMEDIATE-LEVEL KNOWLEDGE

As I hope to have demonstrated, we can turn to conversation analysis when exploring what should be considered intermediate-level knowledge about human-robot interaction. Conversation analytic research covers a wide range of interactional patterns, such as how humans formulate instructions (see e.g. [21], [22]) or how we move in traffic (see e.g. [23], [24]). The communities' wiki www.emcawiki.net can serve as a helpful starting point when searching for existing work on interactional patterns.

However, conversation analysis has more to offer than useful literature for robot interaction design. Taking a CA perspective at several stages of the design process can be a way to actively engage with our intermediate-level knowledge about interaction patterns. Videotaping and transcribing a short interaction between humans or with a robot (prototype) can be a way to validate our intuition and to explicitly formulate intermediate-level knowledge concepts that others can draw upon.

For instance, one might initially only consider verbal greetings when thinking of ways to start an interaction, or one might not be fully aware of the fact that greetings usually come in pairs. The transcription process encourages us to challenge such preconceived ideas about interaction. A transcript forces us to first look at what is observable in the interaction before starting to relate the observations to what we already know. For instance, before saying "the human user says hello, which is a sign of treating the robot as social" in Excerpt 2, conversation analysis encourages us to ask "why that now?" [25], i.e. Why is the husband saying "Servus" at this particular point in the unfolding interaction? If we scrutinize the video by noting actions down in a transcript, we discover that the robot is leaving its charger and plays a friendly "dadu?" sound. The husband apparently hears this as a greeting, as he responds to the robot's sound by a greeting himself, i.e. a return greeting. Without carefully looking at the interaction, we might miss the fact that (from the participant's perspective) Cozmo was the one who actually greeted first, through a sound.

Since conversation analysis works with a small number of interactions or even single cases, it can be applied repeatedly during an iterative design process. Thereby, CA can facilitate the gradual transformation of tacit knowledge into more explicit formats that can be discussed with others and subsequently reported, for instance through showing transcripts in a final paper. Being unique in its grounding in single cases and simultaneous generalization to more general patterns, CA is particularly suitable for scientifically engaging with intermediate-level knowledge.

5. CONCLUSION

This paper has argued that designerly knowledge about interactional patterns should be considered a form of intermediate-level knowledge that is drawn upon in the design process. Critically discussing and possibly rectifying tacit knowledge that we bring to the design process will benefit the HRI community and improve human-robot interaction in the long run. To facilitate critical engagement with tacit knowledge, I introduced multimodal conversation analysis, a research field that studies the detailed organization of human action. We can draw on CA methods and literature to engage with our intuitive understanding of interaction. So far, CA has often taken an evaluative perspective in design research, analyzing interaction with a final product. However, it can and should also be used to continuously engage with and reflect upon designerly knowledge throughout the design process. The script-like transcripts used in CA encourage playful exploration of our own experiences while simultaneously enabling cross-case comparison and discovery of general patterns. Supporting systematic engagement with interactional patterns, CA can stimulate critical discussion of what often remains tacit knowledge in HRI.

Acknowledgment —This work is funded by the Swedish Research Council, project no. 2016-00827. I would like to thank my supervisors Prof. Leelo Keevallik and Prof. Mathias Broth as well as the anonymous reviewer for their helpful comments on this paper.

References

[1] H. Sacks, "Notes on methodology," in Structures of Social Action: Studies in Conversation Analysis, J. Heritage and J. M. Atkinson, Eds. Cambridge, UK: Cambridge University Press, 1984, pp. 21–27.

[2] K. Höök and J. Löwgren, "Strong concepts: Intermediate-level knowledge in interaction design research," ACM Trans. Comput.- Hum. Interact., vol. 19, no. 3, October 2012.

[3] L. A. Suchman, Plans and Situated Actions: The Problem of Human- Machine Communication. Cambridge: Cambridge University Press, 1987.

[4] L. Keevallik, "The interdependence of bodily demonstrations and clausal syntax," Research on Language and Social Interaction, vol. 46, no. 1, pp. 1–21, January 2013.

[5] S. Wiggins, "Talking with your mouth full: Gustatory mmms and the embodiment of pleasure," Research on Language and Social Interaction, vol. 35, no. 3, pp. 311–336, July 2002.

[6] J. Streeck, C. Goodwin and C. D. LeBaron, Ed., Embodied Interaction: Language and Body in the Material World. New York: Cambridge University Press, 2011.

[7] C. Goodwin, "Action and embodiment within situated human interaction," Journal of Pragmatics, vol. 32, no. 10, pp. 1489–1522, 2000.

[8] L. Mondada, "Contemporary issues in conversation analysis: Embodiment and materiality, multimodality and multisensoriality in social interaction," Journal of Pragmatics, vol. 145, pp. 47–62, May 2019.

[9] P. Dourish, Where the Action is: The Foundations of Embodied Interaction. Cambridge: MIT Press, 2001.

[10] K. Höök, Designing with the body: Somaesthetic interaction design. Cambridge: MIT Press, 2018.

[11] A. Woodruff, M. H. Szymanski, R. E. Grinter, and P. M. Aoki, "Practical strategies for integrating a conversation analyst in an iterative design process," in Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, ser. DIS

’02. New York: Association for Computing Machinery, 2002, pp. 255–264.

[12] D. McMillan, B. Brown, I. Kawaguchi, R. Jaber, J. Solsona Belenguer, and H. Kuzuoka, “Designing with gaze: Tama – a gaze activated smart-speaker,” Proceedings of the ACM on Human- Computer Interaction, vol. 3, no. CSCW, November 2019.

[13] H. R. M. Pelikan, M. Broth, and L. Keevallik, “Are you sad, Cozmo?”: How humans make sense of a home robot’s emotion displays,” in Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, ser. HRI ’20. New York, USA: Association for Computing Machinery, 2020, pp. 461–470.

[14] R. Gehle, K. Pitsch, T. Dankert, and S. Wrede, “How to open an interaction between robot and museum visitor? Strategies to establish a focused encounter in HRI,” in Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, ser. HRI ’17. New York, USA: Association for Computing Machinery, 2017, pp. 187–195.

[15] M. Porcheron, J. E. Fischer, S. Reeves, and S. Sharples, “Voice interfaces in everyday life,” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ser. CHI ’18. New York, USA: ACM, 2018, pp. 640:1-640:12.

[16] E. A. Schegloff, “Sequencing in conversational openings,” American Anthropologist, vol. 70, no. 6, pp. 1075–1095, August 1968.

[17] E. A. Schegloff, Sequence Organization in Interaction. Cambridge: Cambridge University Press, 2007.

[18] B. Heenan, S. Greenberg, S. Aghel-Manesh, and E. Sharlin, “Designing social greetings in human robot interaction,” in Proceedings of the 2014 Conference on Designing Interactive Systems, ser. DIS ’14. New York, USA: Association for Computing Machinery, 2014, pp. 855–864.

[19] M. K. Lee, S. Kiesler, and J. Forlizzi, “Receptionist or information kiosk: How do people talk with a robot?” in Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work, ser. CSCW ’10. New York, USA: Association for Computing Machinery, 2010, pp. 31–40.

[20] L. Anderson-Bashan, B. Megidish, H. Erel, I. Wald, G. Hoffman, O. Zuckerman, and A. Grishko, “The greeting machine: An abstract robotic object for opening encounters,” in 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), August 2018, pp. 595–602.

[21] O. Lindwall and A. Ekström, “Instruction-in-interaction: The teaching and learning of a manual skill,” Human Studies, vol. 35, no. 1, pp. 27–49, 2012.

[22] O. Lindwall, G. Lymer and C. Greiffenhagen, “The sequential analysis of instruction,” in The Handbook of Classroom Discourse and Interaction, N. Markee, Ed. John Wiley & Sons, Ltd, 2015, ch. 9, pp. 142–157.

[23] E. De Stefani, M. Broth, and A. Deppermann, “On the road: Communicating traffic,” Language & Communication, vol. 65, pp. 1–6, 2019.

[24] P. Haddington and M. Rauniomaa, “Interaction between road users: Offering space in traffic,” Space and Culture, vol. 17, no. 2, pp. 176–190, 2014.

[25] H. Sacks and E. A. Schegloff, “Opening up closings,” Semiotica, vol. 8, no. 4, pp. 289–327, 1973.