

Portland State University

**PDXScholar**

---

Applied Linguistics Faculty Publications and  
Presentations

Applied Linguistics

---

7-2022

# Common Ground, Cooperation, and Recipient Design in Human-computer Interactions

Judit Dombi  
*University of Pécs*

Tetyana Sydorenko  
*Portland State University, tsydorenko@pdx.edu*

Veronika Timpe-Laughlin  
*Educational Testing Service*

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/ling\\_fac](https://pdxscholar.library.pdx.edu/ling_fac)



Part of the [Applied Linguistics Commons](#)

**Let us know how access to this document benefits you.**

---

## Citation Details

Dombi, J., Sydorenko, T., & Timpe-Laughlin, V. (2022). Common ground, cooperation, and recipient design in human-computer interactions. *Journal of Pragmatics*, 193, 4-20.

This Article is brought to you for free and open access. It has been accepted for inclusion in Applied Linguistics Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).



# Common ground, cooperation, and recipient design in human-computer interactions



Judit Dombi <sup>a, \*</sup>, Tetyana Sydorenko <sup>b</sup>, Veronika Timpe-Laughlin <sup>c</sup>

<sup>a</sup> Department of English Linguistics, University of Pécs, Hungary

<sup>b</sup> Department of Applied Linguistics, Portland State University, USA

<sup>c</sup> Center for Language Education and Assessment Research (CLEAR), Educational Testing Service, USA

## ARTICLE INFO

### Article history:

Received 10 October 2021

Received in revised form 26 February 2022

Accepted 3 March 2022

Available online 28 March 2022

### Keywords:

Socio-cognitive approach

Human-computer interaction

Common ground

Cooperation

Recipient design

## ABSTRACT

In recent years, the number of human-machine interactions has increased considerably. Additionally, we have evidence of linguistic differences between human-machine interactions and human-human conversations (e.g., Timpe-Laughlin et al., 2022). Therefore, it is reasonable to revisit theoretical frameworks that conceptualize interactional language use and investigate to what extent they still apply to technology-mediated interactions. As a first attempt at exploring whether pragmatics theories apply to human-machine interaction, we examined how well Kecskés's (2013) socio-cognitive approach (SCA) focusing on asymmetric interactions (e.g., between interlocutors of different language backgrounds) applies to the asymmetry of human-machine interactions.

Using examples from experimental data, we present the nature of common ground between human and machine (spoken dialogue system) interlocutors, focusing on the construction of and reliance on the emergent side of common ground that is informed by the actual situational experience. Like Kecskés, we argue that both egocentrism and cooperation play a role in human-machine interaction. While the former is manifested in approaching the machine interlocutor as if it was human, the latter appears to play a role in common ground seeking and building as well as in recipient design. We demonstrate that Kecskés's SCA is a fitting framework for analyzing human-machine communication contexts.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The use of spoken dialogue systems (SDS), which rely on natural language processing techniques and automated speech recognition, has been on the rise in recent years due to advancements in artificial intelligence (AI) technology. Since SDS allow for dialogue branching and can thus simulate the feel of a relatively naturalistic conversation (e.g., Timpe-Laughlin et al., 2020), the number of interactions taking place between humans and machines (i.e., computers, automated agents, robots) has increased considerably across a range of communicative purposes and contexts (e.g., Lin et al., 2014; Traum et al., 2008). These include leadership skills trainings (Hays et al., 2012), doctor-patient interactions (Kenny et al., 2007; Stevens et al., 2006), improving mental health (e.g., Ramachandran et al., 2020), and language learning (Johnson and Wang, 2010; Ockey and Chukharev-Hudilainen, 2021; Timpe-Laughlin and Dombi, 2020; Timpe-Laughlin et al., 2020).

\* Corresponding author. Ifjuság Str 6, Pécs, 7624, Hungary.

E-mail address: [dombi.judit@pte.hu](mailto:dombi.judit@pte.hu) (J. Dombi).

Given this rapid increase in human-machine interactions and differences in the linguistic quality of the interactions (e.g., Timpe-Laughlin et al., 2022), theoretical models and frameworks that conceptualize interactional language use should be revisited to see to what extent they can account for language use in technology-mediated communication. As one of the first attempts, Searle (1984) explored human-machine interactions in his famous Chinese room thought experiment and he argued that in interactional situations a computer is not involved in real meaning making. Mey (2010) also argued that AI, unlike human interlocutors, is not able to relate to situated interactivity. While earlier work on the pragmatics of human-computer interaction has been theoretical, a few recent studies have drawn upon experimental data (e.g., Chai et al., 2014; Fischer, 2016a, 2016b; Georgila et al., 2010; Saygin and Cicekli, 2002; Wolters et al., 2009). Chai et al. (2014), for instance, investigated situatedness and recipient design in human-robot interactions, whereas Georgila et al. (2010) examined patterns in interactions between humans and an automated agent (SDS) (see also Wolters et al., 2009). Saygin and Cicekli (2002) studied computers' violations of Grice's cooperative principle and conversational maxims in the context of the Turing Test and found that it is through the violation of the maxims that computers reveal their identity – however, results also suggested that the effects violations of maxims create are varying. As a result, Saygin and Cicekli (2002:227) note that “studying human-computer communication may require some modifications of existing frameworks in pragmatics because of certain characteristics of these conversational environments”. Thus, in our paper, we revisit pragmatic theories and discuss prominent theoretical aspects against the backdrop of human-machine interaction. Throughout, we underpin our points using experimental data that feature interactions between a human and a fully-automated SDS agent.

### 1.1. Motivation for this paper

Most current pragmatic theoretical frameworks are built on communication between speakers of the same L1 (for an exception see Kecskés, 2013, 2019). However, more than 1 billion students globally are learning English (Beare, 2016) in order to be able to communicate with others in an increasingly multilingual world, and most interactions in English take place between speakers who have different L1s (Jenkins, 2007). Thus, it is reasonable to assume that SDS technologies are and will be used in English by speakers of diverse L1s. A secondary motivation for our study is an observation we made in a previous study in which we demonstrated that oral performances collected via face-to-face and SDS task formats are comparable in terms of linguistic indices (e.g., use of vocabulary, syntactic complexity) (Timpe-Laughlin et al., 2022). However, in analyzing whether such data are comparable in terms of pragmatic phenomena, some interesting findings emerged. For example, SDS conversations with an automated agent turned out to be more transactional with fewer but longer turns in contrast to a larger number of short turns and a considerable amount of backchanneling in face-to-face interactions. In other words, our data provide valuable insights into how situatedness affects interlocutor output in an environment where the conversation partner is a computer. Building on these initial observations, we set out to examine how well Kecskés's (2013, 2019) socio-cognitive approach focusing on common ground, cooperation (vs. egocentrism), and recipient design applies (or not) to technology-mediated interaction. Given limited prior research in this specific area, our paper is necessarily exploratory.

## 2. The applicability of the socio-cognitive approach to human-machine interaction

Two seemingly contradictory lines of pragmatics theory have dominated the field in the past decades: the more individualistic line is intention-based and it has been labelled Cognitive Philosophical Pragmatics or Anglo-American Pragmatics (Haugh, 2008). The more context-based Sociocultural-Interactional line is also called European-Continental Pragmatics (Haugh, 2008). One of the main differences between the two traditions is related to the speaker's intention in the interpretation of utterances. Intention, as described in Grice's theory of meaning (1957), has a central role in Cognitive-Philosophical Pragmatics: most theories propose that communication takes place when speakers express their intentions and hearers try to recognize these intentions.

This view postulates that among the main purposes of pragmatic theory is to describe (1) how hearers infer intended speaker meaning and (2) how participants in the conversation know if correct inferences were made. Pragmatics theories that derive from the Gricean framework (e.g., Horn, 2005; Levinson, 2006) tend to emphasize intention and cooperation as the driving forces underlying communication. A divergent line in theorizing that has been labelled Sociocultural-Interactional Pragmatics recognizes intentions as somewhat less straightforward. Versheuren (1999), for instance, pointed out that communication is not always dependent on intentions. Mey (2001) emphasized the importance of considering socio-cultural context in linguistic analysis, highlighting the determining function of society and context in communication, accounting for its situatedness. To summarize, cognitive-philosophical pragmatics investigates speaker meaning by focusing mainly on linguistic constraints on language use. By contrast, socio-cultural interactional pragmatics articulates that research into social and cultural constraints on language use should also be included in pragmatics research.

Connecting these seemingly antagonistic lines of research, Kecskés (2013) proposed a new framework incorporating elements of both traditions. He argued that humans are both individualistic and social beings which is also reflected in their communication. Kecskés emphasized that intention cannot solely be conceived as an *a priori* construct characterizing speaker utterance. Additionally, he highlights that emergent intention also needs to be considered, especially in second language (L2) interactions (Kecskés, 2019:114). His Socio-Cognitive Approach (SCA, Kecskés, 2008; Kecskés and Zhang, 2009) emphasizes that language production and comprehension involve experience and knowledge of prior contexts (traditionally labelled *common ground*), as well as experience and knowledge emerging in the actual situation, co-constructed by interlocutors.

Although SCA was developed with only human interlocutors in mind, we believe that it provides a framework with explanatory potential to understand human interactional behavior in human-machine communication contexts.

We propose various arguments for the applicability of the SCA to human-machine interaction. First, contrary to most pragmatics theories, SCA was developed to explain intercultural encounters, in which language production and comprehension are not strongly affected by conventions and norms of communality. Our key argument is that this is the case in human-machine interaction. Second, the theory revisits the notion of *common ground*, which is taken for granted by traditional pragmatics theories, and it makes a distinction between a *priori* and *emergent* sides of common ground. It steps outside the framework of hearer-centered pragmatics and highlights the role of the speaker's prior experience which affects the speaker's linguistic output (Kecskés, 2010). Also, SCA combines the intention-based pragmatic view of cooperation with the cognitive view of egocentrism to incorporate emerging features of communication (Kecskés 2010, 2012, 2013). Consequently, speaker meaning is not “impoverished”, but speaker utterances involve “speaker intention, personal attitude, and privatized actual contextual elements” (Kecskés, 2013:55). We argue that this combination applies to human-machine interactions: prior experiences surface in initially more egocentric language use, yet actual situational experiences shape communication and fuel conscious cooperation manifested in recipient design often realized via repairs and adjustments. To showcase these arguments, we will utilize sample interactions between L2 learners of English and a fully-automated SDS agent. These interactions were collected by means of the “Request Boss” task.

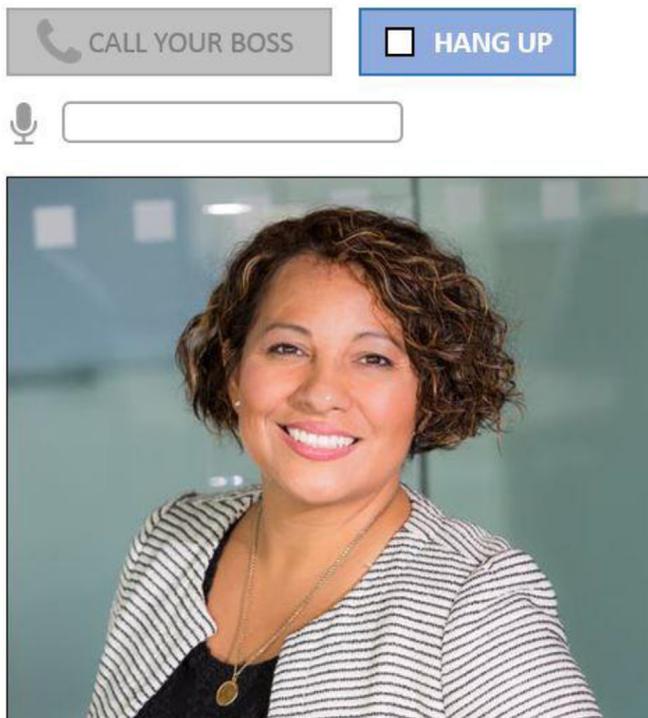
### 2.1. The Request Boss task, the underlying SDS system, and the data

The human-machine interaction data we use to illustrate our points were collected by means of an SDS-based, multi-turn conversation task (see Fig. 1). The task instructions directed the participants to make two requests to their supervisor: to schedule a meeting and to review slides.

You are about to call your boss, Lisa Green. Your goals are to:

1. Get her to agree to have a meeting with you and
2. Ask her to review your presentation slides before the meeting.

Your schedule is free for the rest of the week, so any time proposed by Lisa will work for you.



Copyright © 2020 Educational Testing Service. [www.ets.org](http://www.ets.org).

Fig. 1. Request boss SDS task.

As described in detail in Timpe-Laughlin et al. (2017), the task was created for deployment in HALEF, an open-source, modular, cloud-based framework for designing and operationalizing SDS-based speaking activities (Ramanarayanan et al., 2017). A dialogue structure, shown in Fig. 2, was implemented into HALEF using an OpenVXML design tool. As outlined in Evanini et al. (2017), users access the application through a web browser. Once the user provides a spoken response, such as “Can we have a meeting?”, the utterance is converted to text via an automatic speech recognition (ASR) system. The output of the speech recognizer is then processed by a language understanding component that aims to identify regular expressions based on a list of predefined regular expressions for the given dialogue state. Based on this automatic detection of the regular expression such as “meeting” or “meet”, the dialogue manager then determines the spoken dialogue system’s response. Finally, a language generation component selects the appropriate response from a list of pre-specified system responses. If an utterance such as “Can we have a meeting” is processed correctly, the system would respond, “Yeah, sure. I’m available on Friday at 12. Does that work for you?” Similarly, the system would react to the request for the review of the presentation slides by means of identifying predefined regular expressions such as “presentation”, “slides” etc. regardless of when in the conversation the user makes that request. The cycle repeats for the next user utterance. All user audio is processed in real time on a server using voice activity detection in order to determine the end of a user’s response. Given the small number of possible system responses, instead of using text-to-speech synthesizer, we audio-recorded the system responses (i.e., Lisa Green’s utterances) using a female human voice actor to increase the perceived naturalness of the responses.

Turn	Interlocutor	Utterance
1	Lisa Green	Hello?
	Participant	[greeting]
2	Lisa Green	Hi, how’s it going? What can I do for you?
	Participant	[request for meeting]
3	Lisa Green	Yeah, sure. I’m available on Friday at noon. Does that work for you?
	Participant	[positive response]
4	Lisa Green	Was there anything else you needed?
	Participant	[request to review slides]
5	Lisa Green	Sure, no problem. Send them over.
	Participant	[expression of gratitude]

Fig. 2. Baseline dialogue structure of the SDS request task (adopted from Timpe-Laughlin and Dombi, 2020:232).

In the following, we revisit core theoretical concepts that have played a key role in pragmatics in general, and in SCA in particular: common ground, cooperation (vs. egocentrism), recipient design (vs. salience effect). Then, we discuss how they apply against the backdrop of human-machine interaction. Throughout the discussions we illustrate our points by presenting utterances (authentic transcribed excerpts) from conversations between humans (university-level learners of English) and a fully-automated SDS agent. These utterances were collected by means of the SDS task outlined above. In presenting our argument, we focus on discursive processes and attempt to be reconstructive as we recreate the process through which interlocutors have achieved their communicative goals. In line with Kecskés’s suggestion, we follow a holistic approach and go beyond the utterance level to analyze the dialogic sequence or discourse segment (2017a:89). Thus, we show how the asymmetric nature of human-machine interaction impacts communication and triggers a very conscious approach to what is being said, and how. In the last section, we present implications for further examinations of human-machine interaction.

### 3. The role of cooperation, common ground, and recipient design in human-machine interaction

#### 3.1. Cooperation

Grice tied cooperation to speaker–hearer’s rationality: even if the linguistic system is ubiquitously ambiguous and has a remarkable potential to fail, people tend to use it satisfactorily. This happens, as Grice argues, because we assume that our partner is a cooperative agent (Grice, 1975). By contrast, the view of egocentrism (e.g., Barr and Keysar, 2005; Giora, 2003; Keysar, 2008) challenges the idea that cooperation is automatic and systematic. Drawing on findings from cognitive psychology, Keysar (2008) argues that our own perspective, knowledge, and beliefs have priority over everything else we know about the others’ perspective, knowledge, and beliefs. Additionally, he emphasizes that taking the perspective of the other requires considerable attention and effort (2008:279) and it does not happen as automatically as traditional cooperation-based pragmatics assumed.

Trying to incorporate elements of both cooperation-based pragmatics and findings from cognitive psychology, Kecskés claims that “cooperation and egocentrism are present in the communication process all the time in a varying degree, and the interplay of conscious recipient design (result of cooperation) and subconscious salience (driving force behind egocentrism) is what shapes speaker production” (2017b:244). SCA was developed for human interlocutors who do not share an L1 (and with that, common ground). We believe that similar tendencies can be observed in human-machine interactions, where an obvious

communication asymmetry is present in the sense that interlocutors have neither equal resources and abilities, nor a shared system of knowledge and experiences to function as common ground for them to rely on.

Although the integrated notion of conscious cooperation and subconscious salience has great potential to explain interactional behavior in human-machine communication, a difference that needs to be highlighted is that in Kecskés's approach both participants negotiate meaning and co-construct common ground in a shared and cooperative effort. In a human-machine interaction, however, there is a shift towards the human interlocutor as they have an increased role in finding and establishing the common ground and cooperating in the course of the interaction.

In cases when only one of the interlocutors is a human, the automatic and subconscious cooperative behavior described by Grice (1975) may face a challenge. As Mey (2010) highlights, pragmatics will not operate unless the communicative partners work together. He notes that "if cooperation is offered and ratified, the interactant becomes more than just a sounding board. In interacting, there are no exclusive, isolated subjects and objects: in 'interaction', the action is co-created by the concerted efforts of the interactants" (2010:2887). In human-machine interactions there are also, necessarily, concentrated efforts from both sides: the machine was programmed to communicate with the human, so it evidently has all necessary efforts for cooperation that the particular technology can afford.

However, it is reasonable to assume that the human, interacting with an automated agent or a robot, needs to find more conscious strategies for cooperation than when interacting with other humans because interactions with machines are less natural. A very similar tendency was demonstrated in diverse asymmetric communicative situations: interlocutors who do not share a common first language and sociocultural norms of interaction tend to make conscious cooperative efforts to avoid breakdowns in communication. For example, they use repetition and fronting, engage in self-repairs and avoid semantically less transparent formulaic expressions (e.g., Björkmann, 2011, 2014; Cogo, 2009; Cogo and Dewey, 2012; Dombi, 2020; Kecskés, 2007). Humans may use diverse strategies in a similar tendency to consciously cooperate when interacting with a machine.

Participants in our study were eager to cooperate - this is abundantly evident in our data. However, similarly to Kecskés (2013:150) we also believe that in situations affected by both prior and actual situational contexts, we need to reconsider and further analyze what interactional cooperativeness means. It may be more than the mere subconscious and automatic cooperative behavior described by Grice. Based on our data, we propose that cooperation may play a role in seeking common ground and formulating the utterances to fit the interactional partner (i.e., recipient design). The next two sections aim to explore how this was realized in our data.

### 3.2. Common ground

#### 3.2.1. Core common ground: assumed shared knowledge

As a starting point, common ground (Clark, 1996, 2009) needs to be considered. Common ground denotes the assumed, shared, collective knowledge of interlocutors (Clark, 2009): when speakers form an utterance, they have an assumption of the hearers' ability to correctly interpret the intended message (Allan, 1986; Colston, 2008; Stalneker, 1978). Pragmatic theories are ambiguous in their claims about the degree to which speakers use this assumed mutual knowledge to guide their utterances. Intention and cooperation-based pragmatics theories tend to take this knowledge for granted. For example, accounts like Relevance Theory (Gibbs and Tendahl, 2006; Carston and Powell, 2006; Sperber and Wilson, 1986), or the view of language as joint activity (Bangerter et al., 2004; Bangerter and Clark, 2003; Clark et al., 1983), hold that successful language exchange is anchored in the consideration of shared knowledge. However, recent works urge to rethink the central role of common ground in language production and comprehension (Colston, 2005).

#### 3.2.2. Common ground in human-machine interaction

Even though technologies like Siri, Alexa, and Google assistant are becoming ubiquitous, many people still lack extensive experience in communicating with machines via spoken language and they cannot resort to linguistic conventions in this context (Gaspard and Madison, 2021; Mou and Xu, 2017; Fischer, 2016b). In such cases, communication means interacting with an interlocutor about whom they do not have much knowledge, and with whom they have to get acquainted in the course of the interaction – thus, only an "emerging," online constructed knowledge can be relied on (Kecskés, 2013, 2019). In cases of human-machine interaction, common ground is limited as interlocutors have "significantly mismatched capabilities in perceiving the shared environment" (Chai et al., 2014), they have discrepant knowledge of previous experiences, which would be key in establishing common ground (Kecskés, 2019). Colston advocates that the reliance on common ground may be liable to vary depending on the context of the interaction. Consequently, comprehension should not be viewed as a fixed phenomenon, but rather as a "loose continuum that ranges from something minimal to something richly elaborated upon with inferences" (2008:173).

In an SDS environment, speaker and hearer cannot be equal participants in the communicative process, as the automated agent (just like the agent inside Searle's famous Chinese Room) is not involved in real meaning making. When humans interact with machines, interlocutors cannot rely on common ground in the traditional sense: the automated agent does have some knowledge or common ground programmed into its memory, but this knowledge is limited. In Example 1 it is clear that S and H do not share a common ground: the speaker uses the word *appointment* which is not a word preprogrammed into the system to proceed to the next dialogue stage. Any competent speaker of English would most likely have inferred speaker intention from the utterance italicized in Example 1.

### Example 1: Participant ID29, Human-Machine Interaction

SDS: Hello?  
 ID29: Hi, hello. Uh.  
 SDS: Hi, how's it going? What can I do for you?  
 ID29: Uh, so Lisa, uh, could, uh, *I would like to, like make an appointment with you. Are you available for that?*  
 SDS: I am sorry I didn't get that. What can I do for you?  
 ID29: So uh, um, *I would like to, uh, appointment with you.*  
 SDS: I'm sorry, I'm having trouble understanding you. I have to go, but please feel free to call back later.  
 ID29: Okay.

To contrast, in human–human interaction, core common ground helps the communication process. In Example 2, the participant does not say the word *meet/ing* at all, yet the human interlocutor uptakes the request for a meeting to review the slides, most probably because both participants in the conversation have shared knowledge they can rely on (i.e., knowledge of presentation slides required from employees, of occasions in which people provide feedback on one another's work etc.). This existing knowledge of prior contexts helps them move forward in the communication process.

### Example 2: Participant ID22, Human-Human Interaction

ID22: Uh, hi, boss.  
 Human: Hi.  
 ID22: How's it going?  
 Human: Uh, it's good. Uh, ho- what can I do for you?  
 ID22: Uh, actually, I need you to help me with to review some some work with me.  
 Human: Okay.  
 ID22: Yeah.  
 Human: Uh, what kind of work?  
 ID22: Uh, just uh presentation slides.  
 Human: Okay.  
 ID22: Yeah, because...  
 Human: So you want me...  
 ID22: Yeah, because I I need your feedback. Because your feedback is always uh always helpful.  
 Human: Oh, thank you.  
 ID22: Yeah.  
 Human: Uh, sure. I can review your slides.  
 ID22: Um.  
 Human: Just send them over.  
 ID22: Yeah, but, uh *I I think uh I think we have to review them together. It doesn't matter the day. Today or tomorrow uh. As you find your yourself free, just let me know.*  
 Human: Uh, well, I am uh available to meet on Friday at twelve o'clock. Does that sound good?  
 ID22: Uh, yeah. The it's it's really good for me.

Moreover, Example 2 comprises a hint (italicized), a non-conventionally indirect request (Blum-Kulka et al., 1989). In hints speakers' utterance meaning and the sentence meaning come apart in various ways (Searle, 1975:168). Hints are rare in SDS interactions (cf. Timpe-Laughlin et al., 2022; Timpe-Laughlin and Dombi, 2020), most probably because the human interlocutors are aware that the availability of common ground as well as the possible recognition of semantically opaque utterances are limited. Holtgraves (2002) also highlights that the likelihood of using and understanding indirect speech acts and figurative language increases if the common ground among participants is high, whereas literal interpretations tend to be preferred when common ground is low.

#### 3.2.3. The emergent side of common ground

Example 1 demonstrates that not only are S and H lacking core common ground, the S (human) also fails at establishing emergent common ground.

Clark (1996) differentiated between communal common ground (based on community co-membership) and personal common ground (based on personal acquaintance or joint experience) which is likely to comprise the interactively established referential terms (Clark and Wilkes-Gibbs, 1986). This distinction is very similar to the one proposed by Kecskés (2013): core common ground, the relatively static, generalized common knowledge and beliefs that belong to a community as a result of prior interactions and experiences, and emergent common ground, the dynamic, particularized knowledge created in the course of the interaction in a particular communicative context (Kecskés, 2013:160). As Kecskés (2013) emphasized, participants not sharing a common ground may create an emergent common ground in the course of the interaction, as part of cooperation, in order to manage the interaction. In a human-machine interactional context, the human interlocutor's task in finding this emergent common ground is more demanding. Humans are, theoretically, able to make deductions about how a system works and they may modify their utterances based on this emergent common ground, as Example 3 illustrates. The system does not advance to the next stage because it does not understand the words *questions* and *discussion*; therefore, the participant in the subsequent utterance uses *meet* instead.

### Example 3: Participant ID30, Human-Machine Interaction

SDS: Hello?  
 ID30: Good afternoon, Lisa. This is (name) and uh I have some questions discussio-, I want to discuss some questions with you. And uh–  
 SDS: Hi, how's it going? What can I do for you?  
 ID30: Uh I wonder, can I meet you at tomorrow nine a.m.? And uh I want to talk about uh somethings and a discussions to you.  
 SDS: Yeah, sure. I'm available on Friday at 12. Does that work for you?

Example 3 supports the argument that common ground comprises both a priori and post factum elements (Kecskés, 2013), and information emerging as the dialogue unfolds (Brennan et al., 2010) may shape utterances that occur later. Instead of careful a priori planning, messages may be altered and adapted to addressees when adjustments are needed, as part of a deductive process (Kecskés, 2017c), which is a part of cooperation. Research in cognitive psychology found that the initial planning of utterances often ignores common ground; speakers have been found to monitor their utterances more closely if they perceive misunderstandings related to common ground (Barr, 2004; Barr and Keysar, 2005; Horton and Keysar, 1996; Keysar, 2007). Similarly, in human-machine interaction, a number of studies report that participants adjust their utterances in line with what they progressively discover about the interactional capabilities of machines over the course of the interaction. For instance, they may adapt their word selection, length of turns and even prosody (Avgustis et al., 2021; Branigan et al., 2010; Fischer and Saunders, 2012; Pelikan and Broth, 2016).

Kecskés's idea of emergent common ground is also partly in line with the process view of common ground (e.g., Clark, 1996; Schegloff, 1997; Stalneker, 2002). It claims that common ground is constructed and reconstructed in the process of communication, whereas the SCA holds that both core common ground and emergent common ground play a role in the actual conversational situation, albeit to different degrees. In Example 3 above, Participant 30 realized in the course of the interaction that they need to change their language to be understood by the machine. Their attempt to use other lexical items (*meet, talk*) in the second turn can be understood as an example of building the emergent common ground. As humans have limited ideas on how different SDS actually work, building of the emergent common ground is a discovery process for the human interlocutor (Avgustis et al., 2021; Pelikan and Broth, 2016). In an SDS environment, like the one our data were collected in, only the human participant can have emergent common ground, as the system is not intelligent, in the sense that it does not rely on input in broadening its store of possible answers/cues. However, in different human-machine interactional contexts, including the ones relying on machine learning, it is possible that both interactants rely on emerging as well as core common ground. Thus, core common ground may be conceived of as relying on “memories of schemata, frames, scenarios and scripts” (Allan, 2013:18), whereas emergent common ground derives from “blending individuals' prior personal experiences with perceptions of the actual situational context” (Kecskés, 2013:164) in a cooperative effort to reach understanding of the message.

#### 3.2.4. Common ground and social affiliations

Enfield (2008:235) argues that the management of common ground is tied to managing personal relationships within social networks. In our data, we have ample evidence that in human–human interactions, as compared to human-SDS interactions, interlocutors are more social: a quantitative analysis of the same dataset (Timpe-Laughlin et al., 2022) revealed that in human–human interactions there are more supportive moves to make up for the imposition realized by the request. Also, human–human interactions often contained four to six closing turns, whereas in SDS interactions there were never more than three closing turns.

Participants' social language use can be illustrated via the following data. In Example 2, Participant 22 used a compliment, a positive politeness strategy (Brown and Levinson, 1987), “*your feedback is always uh always helpful,*” and the human interlocutor noticed and acknowledged it. Such deviation from the “task” would not have been acknowledged by our SDS, as the system would have failed to uptake the social element because it was not programmed into it.

Our data also shows that initially even in human-SDS interactions, there are identifiable social elements (italicized in Examples 4 and 5), which either persist throughout the interaction, or are reduced, as participants become more aware of the lack of common ground. In Example 4 the participant keeps being social despite repeated breakdowns in communication, whereas Example 5 shows interactional behavior that accommodated to the situational context and changed to a more factual tone.

### Example 4: Participant ID22, Human-Machine Interaction

SDS: Hello?  
 ID22: Hi, Lisa, how are you?  
 SDS: Hi, how's it going? What can I do for you?  
 ID22: Actually, *can you do a favor for me?* I need to review some things with you because I need to speak with someone about the work. (pause) Yeah, actually, I, I have uh...  
 SDS: I am sorry I didn't get that. What can I do for you?  
 ID22: Uh, in fact uh I need to to speak with you to to review the slides about about the next meeting. *So I need, I really, really need your help in this situation.* And by the time I have uh I don't have any problem any problem to meet any day in uh the next week. So any time good for you it's good for me too.  
 SDS: Sure, no problem. I'll review the slides as soon as you send them to me, and I'm available on Friday at 12 for a meeting. Does that work for you?

ID22: Sure, it works for me. *Thank you for giving me this chance and I know you are busy, but thank you so much.*  
 SDS: I am sorry I didn't get that. What else can I do for you?  
 ID22: Uh no, *thank you so much for your time. I really appreciate that.*  
 SDS: Okay, see you Friday, then.  
 ID22: See you. *Thank you. Have a good rest of the day.*

In Example 4 the same participant who complimented the human interlocutor (in Example 2) used a very social voice with the machine throughout the whole conversation, even after repeatedly failing to achieve the desired illocutionary effect. By contrast, Participant 12 in Example 5 does take notice of the system's inability to uptake social language (*Hope you have a good day*) and simplifies the repair, in accordance with his understanding of the emerging common ground.

#### Example 5: Participant ID12, Human-Machine Interaction

[ ... ]  
 SDS: Sure, no problem. I'll review the slides as soon as you send them to me, and I'm available on Friday at 12 for a meeting. Does that work for you?  
 ID12: Yes. I really appreciate that.  
 SDS: Was there anything else you needed?  
 ID12: *Uh, that's pretty much of it. I really appreciate that. Hope you have a good day.*  
 SDS: I am sorry I didn't get that. What can I do for you?  
 ID12: *No, that's it. Thank you so much.*  
 SDS: Okay, see you Friday, then.  
 ID12: See you, bye.

In Example 4, Participant 22 relied on various supportive moves across the different turns (e.g., *can you do a favor for me, I know you are busy, but thank you so much*) and used lexical intensifiers (e.g., *really, really need your help*). Additionally, Participant 22 repeatedly emphasized that a favour was needed. It seems that it is unclear for the participant how well, and to what extent the machine can handle the use of socially loaded expressions – this, however, may unfold in the course of common ground building, as shown in Example 5, where Participant 12 adjusted their language use to match the context.

Whether or not humans proceed in the same manner even if their social language use is not reciprocated by a machine may depend on numerous reasons: 1) if they conceived the task as a tool (e.g. a machine assisting with navigation) or as a simulation of a real-life interlocutor (e.g., interacting with a boss in a polite manner) (Traum, 2017); 2) how much prior experience or what preconceptions they have about human-machine talk (Fischer, 2016b; Kecskés, 2017b); or 3) their interactional style (Jain et al., 2018; Wolters et al., 2009). However, no matter which of these possibilities are at work, there seems to be evidence from human-machine interactions supporting the claim put forward by Colston that “common ground may best be viewed as an adjustable component of conversations, both in terms of how much it is needed for a task at hand, as well as how much different speakers wish to rely on it for a given conversational exchange” (2008:178). Also, there is evidence for a conscious cooperative behavior that is manifested in adapting to emerging ideas on how the system works, as well as incorporating of prior experiences with talking to machines.

In sum, common ground concerns what is assumed to be shared between speaker and addressee, either as a result of prior experiences or experiences emerging in the specific context of interaction. Thus, it needs to involve speakers' hypotheses about their partners (Fischer, 2016a:31). How and to what extent the communication partner is being taken into consideration by a speaker is addressed by numerous theories, including the audience design model (Bell, 1984), communication accommodation theory (e.g., Coupland et al., 1988), and partner modeling (e.g., Horton and Gerrig, 2005). These theories emphasize another important construct: recipient or audience design.

### 3.3. Recipient design and salience

Speakers design utterances in ways that their interlocutors can easily compute what they meant on the basis of the utterance itself and their common ground (Allan, 2013; Clark et al., 1983). A kind of linguistic accommodation, recipient design, is described as “the multitude of respects in which the talk by a party in a conversation is constructed or designed in ways which display an orientation and sensitivity to the particular other(s) who are the co-participants” (Sacks et al., 1974:727). Research is not congruent on whether recipient design is manifest from initial steps of communication (e.g., Allan, 2013; Clark, 1996), or whether it is only used to repair failed attempts or to make adjustments, if needed (e.g., Horton and Keysar 1996; Keysar et al., 1998). Kecskés argues that these processes may not be viewed as mutually exclusive: in SCA, recipient design and salience both shape language production (Kecskés, 2013:58). In this view, recipient design is the result of cooperation whereas salience drives egocentric language use. As Kecskés (2017c:151) puts it, “recipient design usually requires an inductive process that is carefully planned, while salience effect generally appears in the form of a deductive process that may contain repairs and adjustments.” Salience in linguistics refers to the accessibility of entities in speakers' memory, and how this affects language production and comprehension (Giora, 2003; Kecskés, 2013). It is argued that initially, speakers tend to rely on their own experience of communication (influenced by the many prior contexts they have experienced) and they automatically and unconsciously reach out for the most salient (that is, familiar or frequent) meaning (Giora, 2003) when they design an utterance (Kecskés, 2013). As prior experiences are not necessarily shared by the interlocutors, this individualistic approach is highly egocentric (Keysar, 2008). In the next section, we provide examples of conscious, carefully planned

recipient design in our data, as well as dialogues which are affected by salience and comprise subsequent repairs and adjustments.

### 3.3.1. Conscious recipient design from the beginning of the conversation: cooperation

As Fischer emphasizes, the central role of recipient design in formulating an utterance may become most apparent if aspects of communication “usually taken for granted cannot be relied on” (2016a:7). The less common ground participants share, the more effort can be detected in formulating utterances for the communication partner (Fischer, 2016b; Kecskés, 2013, 2017b, 2017c, 2019). Research has shown that not only linguistic decisions, like lexical or syntactic choices (e.g., Fox, 2008), and interactional processes, such as repairs (e.g., Jefferson, 1974), but also more global, discourse-level choices (e.g., storytelling initiations: Schegloff, 1980; argumentation strategy: Schmitt and Deppermann, 2009) may be tailored to addressees. In our data we could trace conscious recipient design in terms of lexical choices, syntactic choices, interactional style, and discourse-level choices.

In the interactional situations examined in this paper, awareness of the absence of common ground is thought to lead to a very conscious approach to what is being said. Some participants (e.g., Participant 19 in Example 6) displayed a conscious cooperative effort when talking to the machine from the beginning of the conversation. Social language was kept to a minimum, there were clear and short sentences, and avoidance of turn-internal self-repairs and repetition. When contrasting Participant 19’s talk to a human with his talk to a machine, it is clear that short, simple sentences are not prevalent in human–human interaction.

#### Example 6: Participant ID19, Human–Human vs. Human–Machine Interactions

Human–Human	Human–Machine
Human: Hi.	SDS: Hello?
ID19: Hi, uh, hi, [Name]. Uh, uh, can, can, do you have any time this week to meet up with me?	ID19: Hello Lisa. May I meet with you?
Human: Uh, yeah, I have time this week. Um, Friday at twelve is best for me. How does that sound?	SDS: Yeah, sure. I’m available on Friday at 12. Does that work for you?
ID19: Friday at twelve? Oh, yeah, that, that sounds great. I, I’m available then too, yeah. <i>Can we, so can we meet up that day, I need to meet you?</i>	ID19: Yeah, great. That’s available for me.
Human: Uh, sure, we can meet then. Uh, was there anything else you needed?	SDS: Was there anything else you needed?
ID19: Uh, yeah, uh, <i>I need you to review a presentation slide before the meeting also. Can-, could you do that?</i>	ID19: Yeah, uh, I would like to also review my, uh, my presentation slide before the meeting.
Human: Sure, send them over. Um, I’ll review them once I receive them.	SDS: Sure, no problem. Send them over.
ID19: Oh, great. Okay. Uh, I will send it immediately after, after this to your email.	ID19: All right, good. I’ll send them over, over the email. Thank you very much.
Human: Uh, great. Uh, was there anything else you needed?	SDS: Okay, see you Friday, then.
ID19: Uh, no that’s all. Thank you very much for your time.	ID19: See you on Friday. Have a good day.
Human: All right. See you Friday at twelve then.	
ID19: See you there.	
Human: Bye.	
ID19: Bye.	

Our data is in line with prior research that examined speech directed at artificial communication partners and found differences between human–human and human-machine interactions (e.g., Amalberti et al., 1993; Herberg et al., 2008; Kanda et al., 2008), particularly as regards the social aspect of the interaction. Presumably, Participant 19 has experience in conversing with machines (e.g., Google Assistant, Siri, etc.) and uses this *a priori* knowledge to make the conversation with a machine smooth.

Research reports that more experienced users have more realistic ideas about personal assistants as opposed to less experienced users who base their expectations on their familiarity with human–human interactions (Cho et al., 2019; Luger & Sellen; 2016). Differences in spoken human–human and human-machine interactions have been found at suprasegmental level, that is, changes in loudness, intonation, and rhythm (Siegert and Krüger, 2018) as well as at lexical and syntactic levels, when users simplify their utterances and focus on functional keywords (Luger and Sellen, 2016). Luger and Sellen (2016) also observe that more experienced users are more likely to understand the causes of miscommunication and are more likely to continue their attempts, while less experienced users tend to give up earlier. Similarly, Mavrina et al. (2022) observe that as individuals gain more experience communicating with intelligent assistants, their communicative strategies change.

Monitoring existing experience and applying it to the actual situational context can be understood as very conscious cooperation. Participant 19, for instance, selects between conversational styles relative to interlocutor. In the human–human interaction, they apply two different request strategies in both turns containing the two requests (italicized in Example 6), probably in an attempt to emphasize the request. Also, the word “needs” is used two times in the human–human interaction but not used at all with the machine. Participant 19 may have made that lexical choice in an attempt to attend to the hearers’ positive face, thus evoking sympathy.

Similarly, in Example 7 Participant 05 treated the human interlocutor differently from the beginning of the conversation: they started with a greeting, proceeded with a preparator (Blum-Kulka et al., 1989) “can I ask you question” in turn 2, and only arrived at asking the request for the meeting in turn 3. By contrast, a very bold, unmodified direct request for the meeting was issued to the machine in the very first turn.

Example 7: Participant ID05, Human-Human vs. Human-Machine Interactions

Human-Human	Human-Machine
ID05: Hi, [Name], [Name].	
Human: Hi, how's it going? Uh what can I do for you?	SDS: Hi, how's it going? What can I do for you?
ID05: Uh can I ask (pause), can I ask you ques-, question?	ID05: Hi, Lisa. I want to go to meet you.
Human: Sure, um, what do you need help in?	SDS: Yeah, sure. I'm available on Friday at 12. Does that work for you?
ID05: I want to meet with you (pause) and m-, please, my presentation check for you.	ID05: Yes, it's okay.
	SDS: Was there anything else you needed?
	ID05: Uh I need to meet. I want to review my presentation slides.

These moves are all believed to be part of the speaker's strategy to establish rapport with the human interlocutor, which is not really needed in human-machine interaction, as manifested in the first turn containing a simple greeting and a straightforward, direct request for a meeting. Initiating the request early on in human-machine interaction may be conceived as a discourse-level choice - social turns may be thought of as unimportant in human-machine interaction, and crucially important in human-human interaction (see also Example 8, Participant 13 below). This is in line with Fischer's (2016a:27) idea that social and communicative goals are among the features suspected to influence the amount of adjustments to the recipient.

We also found examples where syntactically more complex sentences were uttered in interaction with the human than with the machine. In Example 8, Participant 13 used a rather complex structure (the reason why I came here today is ...) with the human interlocutor, whereas this complex preparator is missing in the interaction with the machine. Instead, the request for the meeting is made in the first turn, using a syntactically transparent request (I like to, uh, meet you, uh, about the business project.)

Example 8, Participant ID13, Human-Human vs. Human-Machine Interactions

Human-Human	Human-Machine
Human: Hi.	SDS: Hello?
ID13: Hi, [Name].	ID13: Hi. My name is [name]. Uh, I like to, uh, meet you, uh, about the business project. Is it okay for you?
Human: Hello.	SDS: Yeah, sure. I'm available on Friday at 12. Does that work for you?
ID13: Um uh how- how's it going?	ID13: Uh, Saturday at twelve, right? Yeah, I'm available on Saturday ni-, uh, Saturday at twelve.
Human: Uh not bad. How about you?	
ID13: Yeah. I'm good, and yeah, so the reason why I came here today is uh to uh to get agree- agreement or to to has to to the meeting uh with this. So um yeah, so what what what time are you available?	

Such differences are thought to support Fox's ideas that “the syntactic structure of a given utterance is constructed for the particular action under production, for this particular recipient at this particular moment in the interaction” (2008:264). Attendance to the system's hypothesized linguistic limitations was also observed by Fischer (2016a:183) who found that humans had preconceptions about whether the robot they were talking to understood particular words or syntactic structures.

3.3.2. Initial salience with subsequent adjustments: egocentrism and cooperation at work

Conscious recipient design, however, is not always at play. Kecskés (2013) argues that speakers are driven not only by the intent that the hearer understands their utterances as they were meant (a cooperative, social component of behavior), but also by salience (a more individualistic component of behavior affecting production in a subconscious way), which can be understood as an example of egocentrism. Fox (2008) highlights that while sensitivity to the recipient may play an important role in utterance design, some of the shape of an utterance may reflect the speaker's own needs. This is an idea that corroborates Kecskés's (2013) view of the co-presence of egocentric and cooperative elements in utterance design. In Example 9, Participant 07 used the non-standard expression *vacant day* (which the system likely did not understand) together with the verb *meet*. As Fox puts it, selection of a word or phrase may “arise from an internal need (an inability to find a sought-for word); speakers may choose particular noun phrases as grammatical subject, based on their own associations or alignments with the referents of those nouns, or cognitive priming, or other speaker-internal processes” (2008:255). As the utterance also included the verb *meet*, there were no repair initiations on behalf of the system, and no breakdown in communication. Yet, the use of *vacant day* perfectly exemplifies an egocentric bias.

### Example 9: Participant ID07, Human-Machine Interaction

SDS: Hello?  
 ID07: Hello? Can I ask your vacant day to meet with you?  
 SDS: Yeah, sure. I'm available on Friday at 12. Does that work for you?  
 ID07: Yes. Uh and I want to, you to check my presentation slide before the meeting.  
 SDS: Sure, no problem. Send them over.  
 ID07: Thank you.  
 SDS: Okay, see you Friday, then.  
 ID07: (silent)

In Kecskés's view, “egocentrism (....) involves a subconscious processing bias based on individual salience and prior experience. Interlocutors just cannot help saying something (...) from their own perspective or expressing something that is on their mind subconsciously. Of course, if what is said or comprehended does not match the actual situational context, repair and adjustment may happen right away” (2013:34).

After miscommunication, many participants in our study displayed more careful attendance to the recipient's needs, which was frequently manifested through lexical modifications. This happened ly due to the way the SDS was designed, that is, to uptake certain expressions and only move on in the dialogue sequence once having elicited them. According to Clark, common ground influences lexical choice, and word selection depends on judgements of what is shared (1998:72). Participants in the SDS task could not have clear ideas on which exact words the system would uptake, but the actual situational context, in this case, the task description (Fig. 1?), may have influenced their choices. In Example 10, Participant 36 used the word *appointment*, however, upon repair initiation, they repeated the request but also added a different request (I want to meet with you) that included the word *meet/ing* which was included in the task description, probably as part of a more conscious conversational cooperation (see also: Participant 36 in Example 10, Participant 30 in Example 3).

### Example 10: Participant ID36, Human-Machine Interaction

SDS: Hi, how's it going? What can I do for you?  
 ID36: Oh, I want to, um, I have some problem with my presentation, so could I make appointment with you someday?  
 SDS: I am sorry I didn't get that. What can I do for you?  
 ID36: Uh, I have some questions with my presen presentation so could I make appointment with you? I want to meet with you.  
 SDS: Sure, no problem. I'll review the slides as soon as you send them to me, and I'm available on Friday at 12 for a meeting. Does that work for you?

In the context of human-machine interactions the extent to which humans take into account that they talk to an automated interlocutor, and how this affects their utterances has attracted increasing attention (e.g., Groom et al., 2009; Lee and Nass, 2010; Nass and Moon, 2000). All of the above examples in this section illustrate that interlocutors' initial lexical choices tend to be affected by prior experiences and privileged meanings, and thus speakers are initially egocentric in the interaction (see also findings of Nass and Moon, 2000; Mou and Xu, 2017). The “mindless transfer” hypothesis (Nass and Moon, 2000) holds that speakers mindlessly employ interactional behavior typical among humans even if they know their interactional partner is a computer, as evidenced in automatic application of social rules or in the breadth and depth of social responses issued to a computer (Groom et al., 2009; Lee and Nass, 2010; Nass and Moon, 2000). There is also research indicating that users initially approach computer interactions in the same way as they would interact with a human (Pelikan and Broth, 2016; Martinovsky and Traum, 2003; Reeves and Nass, 1996; Schneider, 2020), but they adjust their utterances to the context as the interaction unfolds. However, as we have shown in the first part of the Recipient Design section, some speakers make a conscious effort from the very beginning of the interaction to tailor their utterances to their interlocutor, and in the context of human-machine interaction, some studies showed that human–human and human–machine interactions are different (e.g., Amalberti et al., 1993; Herberg et al., 2008; Kanda et al., 2008). The reasons for such individual differences (i.e., people treating automated agents as machines or as humans) are not yet well understood and should be examined in future studies (see our concluding section).

The next example (11) illustrates how prior experiences affect utterance formation, and also exemplifies the effort on Participant 22's part to overcome the breakdown, an example of pragmatic adjustments.

### Example 11: Participant ID22, Human-Machine Interaction

SDS: Hello?  
 ID22: Hi, Lisa, how are you?  
 SDS: Hi, how's it going? What can I do for you?  
 ID22: *Actually, can you do a favor for me? I need to review some things with you because I need to speak with someone about the work. (pause) Yeah, actually, I, I have uh...*  
 SDS: I am sorry I didn't get that. What can I do for you?  
 ID22: Uh, in fact uh I need to to speak with you to to review the slides about about the next meeting. So I need, I really, really need your help in this situation. And by the time I have uh I don't have any problem any problem to meet any day in uh the next week. So any time good for you it's good for me too.  
 SDS: Sure, no problem. I'll review the slides as soon as you send them to me, and I'm available on Friday at 12 for a meeting. Does that work for you?  
 ID22: Sure, it works for me. Thank you for giving me this chance and I know you are busy, but thank you so much.

The first, indefinite attempt to request for meeting/slide was not taken up by the SDS, as it lacked the keywords. In this turn Participant 22 is rather vague and talks in general (*some things ... someone*) most probably in an attempt to make the request appear less of an imposition. Giora's graded salience hypothesis (2003) as well as Kecskes's dynamic model of meaning (2008) highlight the egocentric element in language production and emphasize that what the speaker says relies on privileged meanings (Giora, 2003) and on prior conversational experience (Kecskés, 2008), which surface in lexical choices. Although (relying on Giora, 2003) Kecskés links salience primarily to lexical choices, there seems to be evidence in our data that it may also affect other facets of utterance formation, for example pragmatic or syntactic choices (cf., Fox, 2008). Most probably, unlike Participant 19 in Example 6, participant 22 lacked extensive experience in human-machine talk, and instead relied on human-human conversational experience in which avoidance of face threats is a rational behavior. However, upon realizing that the system would not advance, Participant 22 changed the strategy and used clear, simplified language: *I need to speak with you to review the slides about about the next meeting*, which turned out to be effective.

In Example 12, Participant 39 made significant adjustments to the pragmatic organization of their utterances when talking to the machine. Initially, they tried to insert more pragmatic moves into one turn (Turn 1: greeting + request). However, upon realizing that the system only reacted to the social element, they, in turn, repeated the social part as well. This caused the system to initiate repair, which again made the participant adjust their strategies, neglected the social talk and used only one speech act in the next turn. This actually worked out well, the system proceeded and the participant quickly adapted to the less creative, more functional language required by the speech recognition system. It is especially in contrast with the same participant's H-H interaction (Example 12 on left) that we can see that talking to a human interlocutor, Participant 39 is less proactive as regards the request (appears in the 4th turn, preceded by three turns of greetings and social exchange), and uses strategies to establish rapport (any time that works for you, 2nd floor of the building).

#### Example 12: Participant ID39, Human-Human vs. Human-Machine Interactions

Human-Human	Human-Machine
Human: Hi.	SDS: Hello?
ID12: Hi excuse- uh excuse me.	ID12: Hello Lisa, this is (name). (pause) Uh how are you doing? And I'm wanting to ask you to um review my..-
Human: Hi. Yeah.	SDS: Hi, how's it going? What can I do for you?
ID12: Hello. How are you doing today, [name]?	ID12: Oh, I am doing good. How about you?
Human: Uh I'm good. Thanks for asking. How are you?	SDS: I am sorry I didn't get that. What can I do for you?
ID12: Yeah. I'm really good as well.	ID12: Oh, I'm here to ask uh if you could review my presentation slides before the meeting.
Human: Great.	SDS: Sure, no problem. I'll review the slides as soon as you send them to me, and I'm available on Friday at 12 for a meeting. Does that work for you?
ID12: Yeah. So I just done the report uh about the presentation slide that we just discussed last time, so uh I'm here to ask you if you have any time to uh meet with me and then ha- you can review my slides?	ID12: Oh, yes. That sounds great to me. I'll see you on Friday at twelve then.
Human: Okay. Sure.	SDS: Was there anything else you needed?
ID12: Yeah. Mm-hmm.	ID12: No, thank you so much.
Human: So you want me to review your slides and you want to meet later on, right?	
ID12: Yes. Okay.	
Human: Okay. Great.	
ID12: Uh so is any time that is work for you?	
Human: Uh yeah.	
ID12: Mm-hmm.	
Human: Uh actually, Friday at twelve works for me if that's okay with you?	
ID12: Oh, okay. Yeah. That sounds great. Yeah.	
Human: Great. Um, and then I can review your slides whenever you send them to me.	
ID12: Okay. Yeah. So, we can meet on the second floor of this building. Yeah.	
Human: Okay. Yeah. Sure.	
ID12: I'll see you on Friday at twelve.	
Human: Great. Was there anything else you needed?	
ID12: Um I think, yeah. We're fine with that and then we can go over it on Friday	
Human: Yeah, great. Sounds good.	
ID12: Yeah.	
Human: See you later.	
ID12: Thank you. Bye.	
Human: Bye.	

In terms of syntax, in most dialogues, participants seemed to start off with utterances that are similar in terms of syntactic complexity in both human-human and human-machine interactions, with a lot of adjustments in human-machine dialogues, which were due to misunderstandings. For instance, in Example 13, Participant 35 used an elaborate routine formula

“I’m wondering ...” to make a polite, conventionally indirect request. However, upon facing problems with ASR, Participant 35 removed that part of the request in their next attempt, which resulted in a syntactically less complex request in the repair turn.

Example 13: Participant ID35, Human-Human vs. Human-Machine Interactions

Human-Human		Human-Machine	
ID35:	Uh I'm wondering could I have a hmm make an appointment next week <%>?	SDS:	Hello?
Human:	Sure. Yeah. Yeah. Next week, I'm free Friday at twelve o'clock.	ID35:	Hello? Uh Yeah, my name's (name). Uh, I, I'm wondering if I can make appointment with you for next week. (pause) So could you review my-
ID35:	Okay.	SDS:	Hi, how's it going? What can I do for you?
Human:	Does that sound good?	ID35:	Oh. Uh, good. Uh, so could you, uh, review my presentation slides before the meeting?
ID35:	Yes.	SDS:	Sure, no problem. I'll review the slides as soon as you send them to me, and I'm available on Friday at 12 for a meeting. Does that work for you?
		ID35:	Uh, yes. So, uh, I will email my presentation slides soon. Thank you.

In sum, repairs were obvious examples of cooperation in our human-to-machine data: in case the system did not understand the utterances, participants tended to rephrase them. Repairs were either issued to overcome salience-related misunderstandings, or technology-related breakdowns (failure in ASR, for example). Either way, repairs evidence that the human interlocutor, informed by the particular situational context, relying on emergent common ground, reformulated the initial utterance that most probably had been affected by egocentric choices. Participants making adaptations to their utterances in order to be understood by the system are illustrative of cooperative behavior which surfaces in an “orientation and sensitivity to the particular other” (Sacks et al., 1974:727), that is, recipient design.

These examples support the many ways in which talk in a conversation is designed with an orientation or sensitivity toward the particular other co-present in the interaction. We argue that this planning evidences a conscious effort on behalf of the human participant to cooperate with the artificial agent in the course of the interaction, to make up for the limited common ground that is shared, and to co-construct the emergent common ground that is affected by actual situational experience.

As Fischer (2016a:2) observes, addressee orientation in interaction becomes more obvious when speakers talk to partners' “peculiar” in some way” - thus, in the case of human-machine interaction more conscious recipient design could be presupposed. The fact that many participants still approached the computer with language characteristic of human-human interactions, also highlights the paramount effect of salience. In short, even if technologies like Siri and Alexa are widespread, they still do not constitute the most salient form of communicative interaction for many humans.

4. Concluding remarks

In this paper we argued that Kecskés's socio-cognitive approach, a framework that was originally conceived for communication between human speakers not sharing an L1, is a suitable theoretical framework to analyze and account for human language use in human-machine interactions. It revisits the notion of common ground, which is crucial in human-machine interaction. Furthermore, instead of focusing on the hearer, as most current pragmatic theories do, the SCA urges us to also consider the speaker's prior experiences affecting production.

As has been demonstrated, the construction of emergent common ground in human-machine interactional contexts shows considerable similarities to the same phenomenon in intercultural communication contexts, for the very reason that interactants have little common ground to share. Thus, in this context, similarly to the arguments put forward by Kecskés, a highly context-dependent, emergent side of common ground is constructed, that is informed by actual situational experience.

Our experimental data also provide insights into how this particular context impacts speakers' utterances as regards cooperation, an idea central to most pragmatic theories. Relying on Kecskés's idea on the altered role of cooperation in intercultural interactions, we argue that in human-machine interaction cooperation also gains a new meaning, as conscious cooperative efforts surface in seeking emergent common ground and in recipient design.

We have also provided examples that this is not always a straightforward issue: some participants were more cooperative from the beginning of the conversations, whereas others exhibited a more egocentric behavior. This egocentrism manifested in approaching the computer in the same or in very similar way as they approached the human interlocutor. We explained this phenomenon with the help of the construct of linguistic salience (Giora, 2003) that also features in Kecskés's theory: in the SCA it is argued that both recipient design (cooperation) and salience (egocentrism) shape utterance production. This duality was also evidenced in our data: initially more egocentric speakers turned more cooperative and engaged in recipient design, which was manifested in adjustments and repairs as the dialogues unfolded.

The numerous examples that demonstrate the role of emergent common ground as well as the conscious efforts on behalf of speakers to go on with the conversation even upon facing difficulties show that pragmatic theories that take common ground for granted and cooperation automatic may not be suitable for explaining human language use in interaction with machines. Relying on the examples presented above we argue that it is the SCA that has the explanatory potential to account for this particular type of communication.

Certain limitations of our observations must be acknowledged. First, the fact that participants were L2 speakers with developing language skills in English (see Timpe-Laughlin et al., 2022) may have impacted their utterances. We cannot be

sure about the extent to which the utterances analyzed in this paper were influenced by the facts that participants used English as a lingua franca or that participants conversed with an automated agent. However, as the SCA was developed with non-L1 speakers in mind, we believe that our argumentation is valid in this context. Our points are also supported by studies identifying similar patterns in human-machine interactions involving native speakers (e.g., Fischer, 2016a; Mou and Xu, 2017; Nass and Moon, 2000). Moreover, as English is increasingly used as a lingua franca, and automated agents are increasingly used in diverse areas, including language learning and teaching, it is reasonable to investigate this specific type of context and how it impacts language production.

A potential limitation is that our SDS task featured a social component, i.e., the machine was the employer, a “social superior” to participants, which might also have affected their utterances.

As for the technical affordances, it needs to be highlighted that we used a system with a rigid turn-taking model instead of a system that provides incremental grounding via backchannels mid-utterance (cf. Visser et al., 2014) or that relies on machine learning (Khouzaimi et al., 2018). Such more advanced systems may render different interactional patterns - suspectedly more similar to those of human-human interaction. We believe this is a fruitful area for further research, especially as constraints on technology decrease.

Also, there are many paths for further research that can enhance our understanding of what exactly shapes speaker production in human-machine interactions. A potential line of research could gauge participants' experience with speaking to machines and relating that to their recorded performance or interaction with a computer and with another human to see if there are differences in the interaction approach or style. Second, triangulating performance with verbal reports could shed light on how participants feel about communicating with machines and how they approach the task. Longitudinal inquiries would make it possible to see if speakers' output changes as they gain experience in conversing with machines. Very likely, as it is the case with most facets of language use, individual variation may also play a role here. Thus, a study including the investigation of individual personality traits could provide information on the extent to which personality accounts for variation in talking to machines.

## Declaration of competing interest

The Authors have no competing interests do declare.

## References

- Allan, Keith, 1986. *Linguistic Meaning*. Routledge, New York.
- Allan, Keith, 2013. What is common ground? In: Capone, Alessandro, Lo Piparo, Franco, Carapezza, Marco (Eds.), *Perspectives on Linguistic Pragmatics*. Springer, New York, pp. 285–310. [https://doi.org/10.1007/978-3-319-01014-4\\_11](https://doi.org/10.1007/978-3-319-01014-4_11).
- Amalberti, René, Carbonell, Noëlle, Falzon, Pierre, 1993. User representations of computer systems in human-computer speech interaction. *Int. J. Man Mach. Stud.* 38, 547–566.
- Augustis, Iuliia, Shirokov, Aleksandr, Iivari, Netta, 2021. Please connect me to a specialist”: scrutinising ‘recipient design’ in interaction with an artificial conversational agent. *Lecture notes in computer science*. In: Ardito, Carmelo, Lanzilotti, Rosa, Malizia, Alessio, Petrie, Helen, Piccinno, Antonio, Desolda, Giuseppe, Inkpen, Kori (Eds.), *Human-Computer Interaction – INTERACT 2021*, vol. 12935. Springer, Cham. [https://doi.org/10.1007/978-3-030-85610-6\\_10](https://doi.org/10.1007/978-3-030-85610-6_10).
- Bangerter, Adrian, Clark Herbert, H., 2003. Navigating joint projects with dialogue. *Cognit. Sci.* 27, 195–225.
- Bangerter, Adrian, Clark, Herbert H., Katz, Anna R., 2004. Navigating joint projects in telephone conversations. *Discourse Process* 37 (1), 1–23. [https://doi.org/10.1207/s15326950dp3701\\_1](https://doi.org/10.1207/s15326950dp3701_1).
- Barr, Dale J., 2004. Establishing conventional communication systems: is common knowledge necessary? *Cognit. Sci.* 28 (6), 937–962.
- Barr, Dale J., Keysar, Boaz, 2005. Making sense of how we make sense: the paradox of egocentrism in language use. In: Colston, Herbert L., Katz, Albert N. (Eds.), *Figurative Language Comprehension: Social and Cultural Influences*. Lawrence Erlbaum, Mahwah, NJ.
- Bell, Allan, 1984. Language style as audience design. *Lang. Soc.* 13 (2), 154–204.
- Björkman, Beyza, 2011. Pragmatic strategies in English as an academic lingua franca: ways of achieving communicative effectiveness? *J. Pragmat.* 43 (4), 950–964. <https://doi.org/10.1016/j.pragma.2010.07.033>.
- Björkman, Beyza, 2014. An analysis of polyadic English as a lingua franca (ELF) speech: a communicative strategies framework. *J. Pragmat.* 66, 122–138. <https://doi.org/10.1016/j.pragma.2014.03.001>.
- Blum-Kulka, Shoshana, House, Juliane, Kasper, Gabriele, 1989. *Cross-cultural Pragmatics: Requests and Apologies*. Ablex, Norwood, NJ.
- Branigan, Holly P., Pickering, Martin J., Pearson, Jamie, McLean, Janet F., 2010. Linguistic alignment between people and computers. *J. Pragmat.* 42 (9), 2355–2368. <https://doi.org/10.1016/j.pragma.2009.12.012>.
- Brennan, Susan E., Galati, Alexia, Kuhlen, Anna K., 2010. Two minds, one dialog: coordinating speaking and understanding. *Psychol. Learn. Motiv.* 53, 301–344.
- Brown, Penelope, Levinson, Stephen C., 1987. *Politeness: Some Universals in Language Usage*. Cambridge University Press, Cambridge.
- Carston, Robyn, Powell, George, 2006. Relevance theory – new directions and developments. In: Lepore, Ernest, Barry, C., Smith (Eds.), *Oxford Handbook of Philosophy of Language*. Oxford University Press, Oxford, pp. 341–360.
- Chai, Joyce Y., She, Lanbo, Fang, Rui, Ottarson, Spencer, Littlely, Cody, Liu, Changsong, Hanson, Kenneth, 2014. Collaborative effort towards common ground in situated human-robot dialogue. In: 2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, pp. 33–40.
- Cho, Minji, Lee, Sang Su, Lee, Kun Pyo, 2019. Once a kind friend is now a thing: understanding how conversational agents at home are forgotten. In: *DIS 2019 – Proceedings of the 2019 ACM Designing Interactive Systems Conference*, pp. 1557–1569. <https://doi.org/10.1145/3322276.3322332>.
- Clark, Herbert H., 1996. *Using Language*. Cambridge University Press, Cambridge.
- Clark, Herbert H., 1998. Communal lexicons. In: Malmkjaer, Kirsten, Williams, John (Eds.), *Context in Language Learning and Language Understanding*. Cambridge University Press, Cambridge, pp. 63–87.
- Clark, Herbert H., 2009. Context and common ground. In: Mey, Jacob (Ed.), *Concise Encyclopedia of Pragmatics*. Elsevier, Oxford, pp. 116–119.
- Clark, Herbert H., Wilkes-Gibbs, Deanna, 1986. Referring as a collaborative process. *Cognition* 22, 1–39.

- Clark, Herbert H., Schreuder, Robert, Buttrick, Samuel, 1983. Common ground and the understanding of demonstrative reference. *J. Verb. Learn. Verb. Behav.* 22, 245–258.
- Cogo, Alessia, 2009. Accommodating difference in ELF conversations: a study of pragmatic strategies. In: Anna, Mauranen, Ranta, Elina (Eds.), *English as a Lingua Franca: Studies and Findings*. Cambridge Scholars, Newcastle upon Tyne, pp. 254–273.
- Cogo, Alessia, Dewey, Martin, 2012. *Analyzing English as a Lingua Franca: A Corpus-Driven Investigation*. Continuum, London.
- Colston, Herbert L., 2005. Social and cultural influences on figurative and indirect language. In: Colston, Herbert L., Katz, Albert N. (Eds.), *Figurative Language Comprehension: Social and Cultural Influences*. Lawrence Erlbaum, Mahwah, NJ, pp. 99–130.
- Colston, Herbert L., 2008. A new look at common ground: memory, egocentrism, and joint meaning. In: Kecskés, István, Jacob, L., Mey (Eds.), *Intention, Common Ground and the Egocentric Speaker-Hearer*. Mouton de Gruyter, New York, pp. 151–187.
- Coupland, Nikolas, Coupland, Justine, Giles, Howard, Henwood, Karen, 1988. Accommodating the elderly: invoking and extending a theory. *Lang. Soc.* 17, 1–41.
- Dombi, Judit, 2020. Email requests in an ELF academic setting - "Good luck to you, Doctor Judit...". *Appl. Pragmat.* 2 (1), 54–79. <https://doi.org/10.1075/ap.18017.dom>. John Benjamins.
- Enfield, Nicholas J., 2008. Common ground as a resource for social affiliation. In: Kecskés, István, Jacob, L., Mey (Eds.), *Intention, Common Ground and the Egocentric Speaker-Hearer*. Mouton de Gruyter, Berlin, pp. 223–254.
- Evanini, Keelan, Tsuprun, Eugene, Timpe-Laughlin, Veronika, Ramanarayanan, Vikram, Lange, Patrick, Suendermann-Oeft, David, 2017. Evaluating the impact of local context on CALL applications using spoken dialog systems. In: Colpaert, Jozef, Aerts, Ann, Kern, Rick, Kaiser, Mark (Eds.), *CALL in CONTEXT proceedings*, pp. 239–248. Berkeley, CA.
- Fischer, Kerstin, 2016a. Designing Speech for a Recipient: the Roles of Partner Modeling, Alignment and Feedback in So-Called 'simplified Registers'. John Benjamins, Amsterdam.
- Fischer, Kerstin, 2016b. The situatedness of pragmatic acts: explaining a lamp to a robot. In: Allan, Keith, Capone, Alessandro, Kecskés, István (Eds.), *Pragmatics and Theories of Language Use*. Springer, Cham, pp. 901–910.
- Fischer, Kerstin, Saunders, Joe, 2012. Getting acquainted with a developing robot. In: Salah, Albert A., Ruiz-del-Solar, Javier, Meriçli, Çetin, Oudeyer, Pierre-Yves (Eds.), *Human Behaviour Understanding*. Springer, Berlin, pp. 125–133.
- Fox, Barbara A., 2008. Dynamics of discourse. In: Antos, Gerd, Ventola, Eija V., Weber, Tilo (Eds.), *Handbook of Applied Linguistics*. de Gruyter, Berlin, pp. 255–284.
- Gaspard, Timothy, Madison, T. Phillip, 2021. Functions of imagined interactions with AI among Alexa-users. *Imagin., Cognit. Pers.* 41 (2), 137–157. <https://doi.org/10.1177/02762366211005361>.
- Georgila, Kallirroi, Wolters, Maria, Moore, Johanna D., Logie, Robert H., 2010. The MATCH corpus: a corpus of older and younger users' interactions with spoken dialogue systems. *Comput. Humanit.* 44 (3), 221–261.
- Gibbs, Raymond, Tendahl, Markus, 2006. Cognitive effort and effects in metaphor comprehension: Relevance theory and psycholinguistics. *Mind Lang.* 21, 379–403.
- Giora, Rachel, 2003. *On Our Mind: Salience, Context, and Figurative Language*. Oxford University Press, New York.
- Grice, Herbert Paul, 1957. Meaning. *Phil. Rev.* 66 (3), 377–388.
- Grice, Herbert Paul, 1975. Logic and conversation. In: Cole, Peter, Jerry, L., Morgan (Eds.), *Syntax and Semantics*, vol. 3. Academic Press, New York, pp. 41–58. Speech acts.
- Groom, Victoria, Nass, Clifford, Chen, Tina, Nielsen, Alexia, Scarborough, James K., Robles, Erica, 2009. Evaluating the effects of behavioral realism in embodied agents. *Int. J. Hum. Comput. Stud.* 67 (10), 842–849.
- Haug, Michael, 2008. The place of intention in the interactional achievement of implicature. In: Kecskés, István, Mey, Jacob (Eds.), *Intention, Common Ground and the Egocentric Speaker-Hearer*. Mouton de Gruyter, Berlin, pp. 45–85.
- Hays, Matthew Jensen, Campbell, Julia C., Trimmer, Matthew A., Poore, Joshua C., Webb, Andrea K., King, Teresa K., 2012. Can Role-Play with Virtual Humans Teach Interpersonal Skills? Interservice/Industry Training, Simulation and Education Conference (I/ITSEC), Orlando, FL, United States [Paper presentation].
- Herberg, Jonathan S., Saylor, Megan M., Ratanaswasd, Palis, Levin, Daniel T., Wilkes, D. Mitchell, 2008. Audience-contingent variation in action demonstrations for humans and computers. *Cognit. Sci.* 32 (6), 1003–1020.
- Holtgraves, Thomas M., 2002. *Language as Social Action: Social Psychology and Language Use*. Erlbaum, Mahwah, NJ.
- Horn, Laurence R., 2005. More issues in neo- and post-Gricean pragmatics: a response to Robyn Carston's response. *Intercult. Pragmat.* 3 (1), 81–95.
- Horton, William S., Gerrig, Richard J., 2005. The impact of memory demands on audience design during language production. *Cognition* 96, 127–142.
- Horton, William S., Keysar, Boaz, 1996. When do speakers take into account common ground? *Cognition* 59, 91–117.
- Kecskés, István, 2017a. From pragmatics to dialogue. In: Weigand, Edda (Ed.), *The Routledge Handbook of Language and Dialogue*. Routledge, New York, pp. 78–92.
- Jain, Alankar, Pecune, Florian, Matsuyama, Yoichi, Cassell, Justine, 2018. A user simulator architecture for socially-aware conversational agents. In: *Proceedings of the 18th ACM International Conference on Intelligent Virtual Agents*. Association for Computing Machinery, New York, pp. 133–140. <https://doi.org/10.1145/3267851.3267916>.
- Jefferson, Gail, 1974. Error correction as an interactional resource. *Lang. Soc.* 3, 181–199.
- Jenkins, Jennifer, 2007. *English as a Lingua Franca: Attitude and Identity*. Oxford University Press, Oxford.
- Johnson, W. Lewis, Wang, Ning, 2010. The role of politeness in interactive educational software. In: Hayes, Caroline C., Miller, Christopher A. (Eds.), *Human-computer Etiquette*. Taylor & Francis, New York, pp. 91–113. <https://doi.org/10.1201/b10420-8>.
- Kanda, Takayuki, Miyashita, Takahiro, Osada, Taku, Haikawa, Yuji, Ishiguro, Hiroshi, 2008. Analysis of humanoid appearances in human-robot interaction. *IEEE Trans. Robot.* 24, 725–735.
- Kecskés, István, 2007. Formulaic language in English lingua franca. Explorations in. In: Kecskés, István, Horn, Laurence R. (Eds.), *Pragmatics: Linguistic, Cognitive and Intercultural Aspects*. Mouton de Gruyter, Berlin/New York, pp. 191–219.
- Kecskés, István, 2008. Dueling contexts: a dynamic model of meaning. *J. Pragmat.* 40 (3), 385–406.
- Kecskés, István, 2010. The paradox of communication: a socio-cognitive approach. *Pragmatics and Society* 1 (1), 50–73.
- Kecskés, István, 2012. Is there anyone out there who really is interested in the speaker? *Lang. Dialogues* 2 (2), 283–297.
- Kecskés, István, 2013. *Intercultural Pragmatics*. Oxford University Press, Oxford.
- Kecskés, István, 2017b. Sequential structure of discourse segments shaped by the interplay of recipient design or salience. In: Blochowiak, Joanna, Griset, Cristina, Durrleman, Stephanie, Lenzlinger, Christopher (Eds.), *Formal Models in the Study of Language*. Springer, Cham, pp. 243–260. [https://doi.org/10.1007/978-3-319-48832-5\\_13](https://doi.org/10.1007/978-3-319-48832-5_13).
- Kecskés, István, 2017c. The interplay of recipient design and salience in shaping speaker's utterance. In: de Ponte, María, Korta, Kepa (Eds.), *Reference and Representation in Thought and Language*. Oxford University Press, pp. 144–162. <https://doi.org/10.1093/oso/9780198714217.003.0009>.
- Kecskés, István, 2019. *English as a Lingua Franca: the Pragmatic Perspective*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/9781316217832>.
- Kecskés, István, Zhang, Fenghui, 2009. Activating, seeking and creating common ground: a socio-cognitive approach. *Pragmat. Cognit.* 17 (2), 331–355.
- Beare, Kenneth, 2016. How many people learn English globally? About Education. Retrieved on September 29, 2021 from [http://esl.about.com/od/englishlearningresources/ff\\_eslmarket.html](http://esl.about.com/od/englishlearningresources/ff_eslmarket.html).

- Kenny, Patrick, Parsons, Thomas D., Gratch, Jonathan, Leuski, Anton, Rizzo, Albert A., 2007. Virtual patients for clinical therapist skills training. In: Pelachaud, Catherine, Martin, Jean-Claude, André, Elisabeth, Chollet, Gérard, Karpouzis, Kostas, Pelé, Danielle (Eds.), *Proceedings of 7th International Conference on Intelligent Virtual Agents*. Springer, New York, pp. 197–210.
- Keysar, Boaz, 2007. Communication and miscommunication: the role of egocentric processes. *Intercult. Pragmat.* 4, 71–84.
- Keysar, Boaz, 2008. Egocentric processes in communication and miscommunication. In: Kecskés, István, Jacob, L., Mey (Eds.), *Intention, Common Ground and the Egocentric Speaker-Hearer*. Mouton de Gruyter, Berlin, pp. 277–296.
- Keysar, Boaz, Barr, Dale, Horton, William, 1998. The egocentric basis of language use: insights from a processing approach. *Curr. Dir. Psychol. Sci.* 7, 46–50.
- Khouzaimi, Hatim, Laroche, Romain, Lefevre, Fabrice, 2018. A methodology for turn-taking capabilities enhancement in Spoken Dialogue Systems using Reinforcement Learning. *Comput. Speech Lang.* 47, 93–111.
- Lee, Jong-Eun R., Nass, Clifford I., 2010. Trust in computers: the computers-are-social-actors (CASA) paradigm and trustworthiness perception in human-computer communication. In: Latusek, Dominika, Gerbasi, Alexandra (Eds.), *Trust and Technology in a Ubiquitous Modern Environment: Theoretical and Methodological Perspectives*. IGI Global, pp. 1–15.
- Levinson, Stephen C., 2006. Cognition at the heart of human interaction. *Discourse Stud.* 8 (1), 85–93.
- Lin, Raz, Ya'akov (Kobi), Gal, Kraus, Sarit, Mazliah, Yaniv, 2014. Training with automated agents improves people's behavior in negotiation and coordination tasks. *Decis. Support Syst.* 60, 1–9.
- Luger, Eva, Sellen, Abigail, 2016. Like having a really bad PA\*: the gulf between user expectation and experience of conversational agents. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 5286–5297. <https://doi.org/10.1145/2858036.2858288>.
- Martinovsky, Bilyana, Traum, David, 2003. The Error Is the Clue: Breakdown in Human-Machine Interaction. ISCA Tutorial and Research Workshop, Château d'Oex, Vaud, Switzerland [Paper presentation].
- Mavrina, Lina, Szczulka, Jessica, Strathmann, Clara, Bohnenkamp, Lisa M., Krämer, Nicole, Kopp, Stefan, 2022. Alexa, you're really stupid\*: a longitudinal field study on communication breakdowns between family members and a voice assistant. *Front. Comput. Sci.* 4. <https://doi.org/10.3389/fcomp.2022.791704>.
- Mey, Jacob L., 2001. *Pragmatics: an Introduction*, 2nd ed. Blackwell, Oxford.
- Mey, Jacob L., 2010. Reference and the pragrameme. *J. Pragmat.* 42, 2882–2888.
- Mou, Yi, Xu, Kun, 2017. The media inequality: comparing the initial human-human and human-AI social interactions. *Comput. Hum. Behav.* 72, 432–440.
- Nass, Clifford, Moon, Youngme, 2000. Machines and mindlessness: social responses to computers. *J. Soc. Issues* 56 (1), 81–103. <https://doi.org/10.1111/0022-4537.00153>.
- Ockey, Gary J., Chukharev-Hudilainen, Evgeny, 2021. Human versus computer partner in the paired oral discussion test. *Appl. Ling.* <https://doi.org/10.1093/applin/amaa067>.
- Pelikan, Hannah R.M., Broth, Mathias, 2016. Why that Nao?: how humans adapt to a conventional humanoid robot in taking turns-at-talk. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/2858036.2858478>.
- Ramachandran, Maya, Sanaa, Suharwardy, Leonard, Stephanie A., Gunaseelan, Anita, Robinson, Athena, Darcy, Alison, Judy, Amy, 2020. 74: acceptability of postnatal mood management through a smartphone-based automated conversational agent. *Am. J. Obstet. Gynecol.* 222 (1). <https://doi.org/10.1016/j.ajog.2019.11.090>. S62.
- Ramanarayanan, Vikram, Suendermann-Oeft, David Lange, Patrick, Mundkowsky, Robert, Ivanov, Alexei, V., Yu, Zhou, Qian, Yao, Evanini, Keelan, 2017. Assembling the jigsaw: how multiple open standards are synergistically combined in the HALEF multimodal dialog system. In: Dahl, Deborah A. (Ed.), *Multimodal Interaction with W3C Standards: towards Natural User Interfaces to Everything*. Springer International Publishing, Switzerland, pp. 295–310. [https://doi.org/10.1007/978-3-319-42816-1\\_13](https://doi.org/10.1007/978-3-319-42816-1_13).
- Reeves, Byron, Nass, Clifford, 1996. *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*. CSLI Publications, CA.
- Sacks, Harvey, Schegloff, Emanuel A., Jefferson, Gail, 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50 (4), 696–735.
- Saygin, Ayse P., Cicekli, Ilyas, 2002. Pragmatics in human-computer conversations. *J. Pragmat.* 34 (3), 227–258.
- Schegloff, Emanuel A., 1980. Preliminaries to preliminaries: 'Can I ask you a question. *Socio. Inq.* 50 (3–4), 104–152.
- Schegloff, Emanuel A., 1997. Whose text, whose context? *Discourse Soc.* 8 (2), 165–187.
- Schmitt, Reinhold, Deppermann, Arnulf, 2009. Damit Sie mich verstehen\*: genese, Verfahren und Recipient Design einer narrativen Performance. In: Schneider, Jan Georg, Buss, Mareike, Frank, Liedtke, Habscheid, Stephan, Jautz, Sabine (Eds.), *Theatralität des sprachlichen Handelns: Eine Metaphorik zwischen Linguistik und Kulturwissenschaften*. Fink, München, pp. 79–107.
- Schneider, Florian, 2020. How users reciprocate to Alexa - the effects of interdependence. In: Stephanidis, Constantine, Antona, Margherita, Ntoa, Stavroula (Eds.), *HCI International 2020 - Late Breaking Posters*. Springer, Cham, pp. 376–383.
- Searle, John R., 1975. Indirect speech acts. In: Cole, Peter, Morgan, Jerry L. (Eds.), *Syntax and Semantics*, vol. 3. Academic Press, New York, pp. 59–82. *Speech acts*.
- Searle, John R., 1984. *Minds, Brains, and Science*. Harvard University Press, Cambridge, MA.
- Siegert, Ingo, Krüger, Julia, 2018. How do we speak with ALEXA - subjective and objective assessments of changes in speaking style between HC and HH conversations. *Kognitive Systeme* 1. <https://doi.org/10.17185/dupublico/48596>.
- Sperber, Dan, Wilson, Deirdre, 1986. *Relevance: Communication and Cognition*. Blackwell, Oxford.
- Stalnaker, Robert C., 1978. Assertion. In: Cole, Peter (Ed.), *Syntax and Semantics*, vol. 9. Academic Press, New York, pp. 315–332. *Pragmatics*.
- Stalnaker, Robert C., 2002. Common ground. *Ling. Philos.* 25, 701–721.
- Stevens, Amy, Hernandez, Jonathan, Johnsen, Kyle, Dickerson, Robert, Raji, Andrew, Harrison, Cyrus, DiPietro, Meredith, Allen, Bryan, Ferdig, Richard, Foti, Sebastian, Jackson, Jonathan, Shin, Min, Cendan, Juan, Watson, Robert, Duerson, Margaret, Lok, Benjamin, Cohen, Marc, Wagner, Peggy, Lind, D Scott, 2006. The use of virtual patients to teach medical students history taking and communication skills. *Am. J. Surg.* 191 (6), 806–811.
- Timpe-Laughlin, Veronika, Dombi, Judit, 2020. Exploring L2 learners' request behavior in a multi-turn conversation with a fully automated agent. *Intercult. Pragmat.* 17 (2), 221–257. <https://doi.org/10.1515/ip-2020-0010>. De Gruyter.
- Timpe-Laughlin, Veronika, Evanini, Keelan, Green, Ashley, Blood, Ian, Dombi, Judit, Ramanarayanan, Vikram, 2017. Designing interactive, automated dialogues for L2 pragmatics learning. In: *Proceedings of the 21st workshop on the semantics and pragmatics of dialogue*. SemDial, pp. 143–152. In: [http://www.saardial.unisaarland.de/wordpress/wp-content/uploads/SemDial2017SaarDial\\_proceedings.pdf](http://www.saardial.unisaarland.de/wordpress/wp-content/uploads/SemDial2017SaarDial_proceedings.pdf).
- Timpe-Laughlin, Veronika, Sydorenko, Tetyana, Daurio, Phoebe, 2020. Using spoken dialogue technology for L2 speaking practice: what do teachers think? *Comput. Assist. Lang. Learn.* 1–24. <https://doi.org/10.1080/09588221.2020.1774904>. Routledge.
- Timpe-Laughlin, Veronika, Sydorenko, Tetyana, Dombi, Judit, 2022. Human versus machine: investigating L2 learner output in face-to-face versus fully automated role-plays. *Comput. Assist. Lang. Learn.* <https://doi.org/10.1080/09588221.2022.2032184>.
- Traum, David, 2017. Computational approaches to dialogue. In: Edda, Weigand (Ed.), *The Routledge Handbook of Language and Dialogue*. Taylor & Francis, New York, pp. 143–161.
- Traum, David, Marsella, Stacy C., Gratch, Jonathan, Lee, Jina, Hartholt, Arno, 2008. Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In: *International Workshop on Intelligent Virtual Agents*. Springer, Heidelberg, pp. 117–130.
- Verschueren, Jef, 1999. *Understanding Pragmatics*. Edward Arnold, London.
- Visser, Thomas, Traum, David, DeVault, David, op den Akker, Rieks, 2014. A model for incremental grounding in spoken dialogue systems. *Journal on Multimodal User Interfaces* 8 (1), 61–73.
- Wolters, Maria, Georgila, Kallirroi, Moore, Johanna D., MacPherson, Sarah E., 2009. Being old doesn't mean acting old: how older users interact with spoken dialog systems. *ACM Transactions on Accessible Computing* 2 (1), 1–39. Article 2.

**Judit Dombi** (Ph.D., University of Pécs, Hungary) is an Associate Professor of Linguistics at the University of Pécs, Hungary, where she teaches undergraduate and graduate courses in linguistics. Her fields of interest include theoretical and applied aspects of intercultural communication and pragmatics in intercultural and ELF contexts. Her recent research has focused on intercultural communicative competence, the communicative functions of directness and indirectness in interaction, and communication asymmetries.

**Tetyana Sydorenko** (Ph.D., Michigan State University) is an Associate Professor of Applied Linguistics at Portland State University. Her research interests include L2 pragmatics, computer-assisted language learning, psycholinguistic processes in SLA, and assessment. She is currently focusing on the use of adaptive computer-simulated conversations in the teaching of L2 pragmatics.

**Veronika Timpe-Laughlin** (Dr. phil., TU Dortmund University, Germany) is a research scientist in the Center for Language Education and Assessment Research at Educational Testing Service. Her research interests include pragmatics, language assessment, task-based language teaching, and technology in L2 instruction and assessment. Veronika has recently published in *Intercultural Pragmatics* and *Language Testing* and is the co-author of the 2017 book "Second Language Educational Experiences for Adult Learners" (Routledge). Prior to joining ETS, Veronika taught in the English Department at TU Dortmund University, Germany.