Human-Machine Interaction as a Complex Socio-Linguistic Practice

Netaya Lotze

Abstract *This paper presents a socio-linguistic model for Human-Machine Interaction (HMI), examining the interplay of technological affordances, user cognitive awareness, and language strategies. The model features three continua: technological affordances, users' cognitive awareness, and language strategies. The first dimension evaluates the anthropomorphism degree of the system, including linguistic anthropomorphism and therefore tries to integrate Ruijten's et al. (2014/2019) Rasch-scale of human perception of anthropomorphic designs. The second dimension explores users' cognitive awareness, ranging from pre-conscious alignment to conscious strategies. The third dimension depicts a continuum of user language, from pre-conscious alignment (Gandolfi et al. 2023) and linguistic routines and behaviors, transferred from HHC (CASA: Reeves and Nass 1996; MASA: Lombard and Xu 2021) to various simplification strategies as robot-directed speech (RDS), simplified registers (SR) (Fischer 2011) and computer talk (CT) (Zoeppritz 1985). The paper argues from a diachronic perspective that HMI language evolution is influenced not only by anthropomorphic technology and user awareness but also by language variation and change, and societal factors. Therefore, the results of numerous studies of my own research group conducted between 2000 and the present (with a particular focus on Lotze 2016) will be summarized and interpreted in light of the model, and vice versa.*

1. Introduction: The AAS-Model of HMI as a Complex Socio-Linguistic Practice

Users seem to interact with AI either as they would with a human conversation partner or in a simplified form specifically designed for operating a machine. The fact that empirical studies on HMI (Human-Machine Interaction) continue to yield contradictory results regarding alignment, politeness markers, and routines similar to those found in HHC (Human-Human Communication) on the one hand, and simplifications, imperatives, and isolated keywords as bot-directed language on the other (Fischer 2006), has led to the emergence of two competing research paradigms. Each of these paradigms only captures and explains partial aspects of the phenomenon: a) the CASA/MASA approach ('All media are social actors'), which assumes that users always attribute social characteristics to the system (see Reeves and Nass 1996; Lombart and Xu 2021), and the 'Simplified Registers' approach, which focuses primarily on simplifications, repair and related phenomena (Fischer 2006; see also 'Computer Talk' (CT) Zoeppritz 1985). Both approaches can only analyze minor portions of the whole complexity of the empirical phenomena.

Therefore, in Chapter 3 of this article, I propose my AAS-model of HMI as a complex socio-linguistic practice that can claim broader validity. In my opinion, 'user awareness' seems to be the most relevant cognitive key concept for this purpose, which is often missed by other approaches. Accordingly, my model is structured into three dimensions: the degree of anthropomorphism of the technology, the cognitive awareness of the users, and the user language. This creates a decision space in which users can position themselves on three continua, always keeping the current interaction situation with the AI and its cultural and pragmatic implications in mind.

The model also aims to capture the dynamic nature of HMI language development, emphasizing metaphorical language use after Krause and Hitzenberger (1992) as it undergoes diachronic transformations with technological shifts – always with a little delay (c.f. Schmitz 2015 "stilistisches Trägheitsgesetz" ("stylistic inertia")). External influences on user language, such as dialog design, data foundations, and linguistic models, are supplemented by sociolinguistic factors like language change, cultural dynamics, and societal shifts. The model provides at this point in time a first glimpse into the intricate web of language evolution in HMI, offering a balanced perspective on both technological and socio-linguistic dimensions in human-machine interaction.

The most significant contribution of the model lies in its ability to a) not only integrate but also partially explain traditionally contradictory tendencies in user behavior (e.g. anthropomorphization vs. simplification), and b) maintain connections to the historical academic discourses on CASA/MASA and "computer talk" (CT) as a "simplified register" (SR), while going far beyond those approaches by conceptualizing HMI as inherently dynamic, negotiable, and subject to socio-cultural change. Consequently, the model does not rely on static categories but instead opens a decision space composed of continua in which the user can position themselves. The approach breaks with structuralist or positivist theories, yet it does not seek to entirely discard those aspects that were useful for our better understanding of HMI. Rather, it aims to demonstrate that HMI is more complex and diverse than traditional theories have recognized.

In the following two chapters, the research landscape of traditional approaches in HMI research (CASA/MASA vs. SR) will be presented, followed by an overview of the empirical studies conducted by my research group over the past twenty years. The focus will be on aspects that extend beyond CASA/MASA and SR to illustrate why my new model is necessary.

2. The CASA/MASA Approach as One of the Earliest Reference Points for Interpreting Linguistic User Behavior

Despite from a philosophical and sociological perspective contemporary artificial intelligences, even in dialogue systems employing GPTs, still lacking the characteristics of a social actor, as they neither possess self-reflective consciousness, emotions, empathy, nor exercise a free will capable of autonomously setting and being held accountable for their own goals, linguistic analyses of user language since the late 1990s suggest that individuals exhibit a tendency to transfer linguistic concepts from human-human communication (HHC) onto human-machine interaction (HMI) (cf. Reeves and Nass 1996; Nass and Moon 2000; Nass and Brave 2005; Reeves et al. 2020). Thus, cognitively, they engage in a certain form of anthropomorphization of systems on a conceptual level, triggered by the natural language dialogue serving as depiction of a human interlocutor (Clark and Fischer 2023). Accordingly, the anthropomorphic design of the system on a technological level, triggers an anthropomorphization on the level of user cognition and a conceptualization as a social actor. Reeves and Nass (1996) analyze this user behavior in their early studies as “mindless behavior”, interpreting it as a preconscious transfer of concepts, schemas, and action routines from HHC (Reeves and Nass 1996; Nass and Moon 2000; Nass and Brave 2005; Reeves et al. 2020). Reeves and Nass (1996) argue that individuals apply a social model when confronted with a complex entity whose mechanisms they do not immediately comprehend. Linguistically, this phenomenon manifests, for instance, in the transfer of certain ritualized protocols (Sacks et al. 1992), frames and scripts (Fillmore 1976), or levels of politeness (Brown and Levinson 1987) from HHC to HCI.

They consolidate this research stance into their "Computers are Social Actors (CASA) Paradigm", elucidated and systematized in "The Media Equation" Reeves and Nass 1996). Subsequently, they and other research groups discovered numerous cross-cultural pieces of evidence for social effects of dialogue systems that can be interpreted through the CASA framework (Nass and Moon 2000; Nass and Brave 2005; Reeves et al. 2020). However, the precise role of preconscious attribution of social attributes to the system in the context of HMI remains a subject of ongoing controversial discourse (see Lotze 2016 for a deeper exploration; Dippold 2023).

In more recent times, the CASA approach has been expanded and refined into MASA ("Media are Social Actors", Lombard and Xu 2021), which can be applied to various contemporary media and it incorporates the degree of anthropomorphism associated with these media. Ruijten et al. (2014) proposed a Rasch-like anthropomorphism scale for their psychology of AI perception, systemizing objects in general (and specifically robots, agents, and assistance systems), with varying effects on user reception. Since 2019, they have tested and confirmed this idea in different scenarios with diverse participants, yielding replicable results. Ruijten's et al. approach could show, that the technological level of anthropomorphic design and the user's perception of it as more or less social are closely intertwined in a systematic way. And only because of that, we can combine the otherwise separate levels in the Rasch-like anthropomorphism scale, they suggest. Lombard and Xu (2021) adopt this scale of degrees of anthropomorphism from psychology and integrate it into CASA. Unlike Nass and Reeves' early approach, MASA considers the degree of anthropomorphism in AI design, which significantly influences individuals, especially when the representation is more humanoid.

In my opinion, CASA is an exceptionally fruitful idea and model that can explain a significant number of user utterances across various contexts. Nevertheless, having scrutinized hundreds of real interactions from anonymous users with customer support bots in the field, I contend that, on the flip side, there still exists a stable corpus of user expressions over different applications and decades that unfortunately eludes explanation through the CASA/MASA paradigm. Consequently, I have come to the conclusion that CASA (and to a lesser extent, MASA) constitutes a position within the research community that adeptly captures only one singular driving force behind user behavior towards AIs – specifically, the transfer of behaviors from Human-Human

Communication (HHC), with all its implications for dialogue (as mentioned earlier: routines and protocols, frames and scripts, linguistic politeness). However, it exhibits a blind spot for all aspects of user behavior that deviate from HHC and are currently evolving: simplifications in the form of audience-specific "Simplified Registers" (Fischer 2006; 2011), such as syntactically simplified commands or questions observed in RequestandResponse systems like Amazon Alexa (Greilich, in preparation, see below), or isolated keywords, popular among Digital Natives, for instance in Social Bots used in customer service (Lotze and Ohrndorf, in preparation, see below), leading quickly to the dialogue goal, especially in written media.

In my opinion, it is particularly crucial to emphasize that aspects of linguistic economy (including Ronneberger-Sibold 1980; Köhler 2005) play a crucial role here, as they have become relevant in the context of digitization in real-time written communication among people (regarding IR chats and SMS: Siever 2011; concerning messenger apps: König 2019). Some of the simplifications observed in the field can be explained depending on the technological affordances of the respective language system, while others appear to represent emerging socio-linguistic practices for interacting with AI, evolving as variants currently just in the process of formation.

With this article, my intention is to strongly advocate for the notion that Human-Machine Interaction (HMI), at least at present (future systems might become even more human-like), constitutes a heterogeneous form of interaction. It incorporates preconscious to ritualized transfers from Human-Human Communication (HHC) but also exhibits numerous new stylistic parameters addressing the utilitarian nature of the application. These include simplifications, the absence of politeness, increased use of vulgarisms, and considerations of its representational character (cf. Lotze 2016), or its performative nature (the staging of interaction with AI as a philosophical game, discussing the AI with others during an ongoing dialogue, etc.). Clark and Fischer (2023) similarly underscore that dialogue systems and robots are always "depictions" of humans, and modern users (in contrast to those of Weizenbaum's ELIZA in the 1960s) are indeed conscious of this representational character. They provide numerous example dialogues that effectively illustrate how individuals intermittently engage more or less in this role-play. Fischer's (2006) user types – a "Player" who embraces the portrayal of an anthropomorphic conversational partner and a "Non-Player" who conceptualizes the application more as a tool –, in my opinion, are valuable key concepts for systematizing the heterogeneity of user strategies.

In this article, I aim to present HMI as a multi-dimensional socio-linguistic practice, considering not only varying degrees of anthropomorphism in system design but also increasing levels of user awareness. Our studies, particularly those focused on interactive alignment, have demonstrated that users exhibit more or less preconscious or routinized behavior in different dialogue sequences and phases. Importantly, users maintain such behavior only as long as the sequence proceeds without disruptions (cf. Lotze 2016; Krummheuer 2010). Thus, preconscious transfers from HHC depend on the user type according to Fischer (2006), the degree of linguistic anthropomorphism in the interface (Ruijten et al. 2014; 2019), the dialogue phase, and disruptions in the dialogue (Lotze 2016). Only a model that additionally incorporates the user's levels of awareness can adequately address the heterogeneity of HMI, as opposed to models that consider individual aspects in isolation. HMI, therefore, must be conceptualized in three dimensions: a) as user language (in variation and evolution), b) as user awareness (on a continuum from preconscious to conscious/strategic), and c) in relation to the degree of anthropomorphism in system design (both visually and linguistically). This approach creates a three-dimensional decision space, wherein users position themselves with each contribution to the conversation. Simultaneously, this framework serves as a model for the linguistic AI research community to better interpret linguistic user behavior in HMI.

In the article, my model of HMI will be introduced as a complex and heterogeneous socio-linguistic practice, grounded in theory (see Chapter 3) and motivated by the results of my research group (empirical evidence, see Chapter 2). It is imperative to firmly connect our research in both empirical evidence and theory to the existing and current research landscape.

In Chapter 2, I will present relevant studies conducted by my research group on linguistic user behavior, discussing those aspects (Alignment, Acceptance, Simplification (AAS)) that have been incorporated into the model:

a) Lotze (2016): Corpus study on rule- and plan-based chatbots.
b) Lotze and Ohrndorf (in preparation): Corpus study on Socialbots in customer service.
c) Greilich (in preparation): Psycho-linguistic experiment on Amazon Alexa.
d) Lotze and Aydin (in preparation): Qualitative-explorative study on ChatGPT following an ethnomethodology.

Subsequently, in Chapter 1, the research horizon will be outlined. In Subsection 1.2, the scientific-historical foundations of the discourse on "Computer-Talk" (CT, Zoeppritz 1985) and "Simplified Registers" (Fischer 2006; 2011) will be presented to better understand the relevance of simplifications for HMI. Considering the historical background of both terms, new and more nuanced conceptualizations will be explored. Section 1.3 will then focus on the heterogeneity of HMI (following Lotze 2016) in detail. Subsets of HMI will be delineated, described, and categorized.

Chapter 2, as mentioned earlier, follows as the empirical section, with Subsection 2.2 focusing on our current studies on ChatGPT concerning simplifications, addressing and discussing relevant aspects.

In Chapter 3, the model of HMI as a complex socio-linguistic practice is presented. It will be discussed within the context of a diachronic perspective on communication in the age of digitization. This chapter aims to provide a comprehensive understanding of HMI, drawing on the theoretical foundations and empirical findings outlined in the preceding chapters.

2.1 The Academic Discourse on "Simplified Registers" as a Counterpoint to CASA/MASA?

Fischer's (2011) framework of "Simplified Registers" emerges as a crucial starting point for analyzing strategic simplifications by users. When faced with a robot or agent, individuals engage in strategic actions, consciously simplifying their language. Fischer's benchmarks for HMI include other highly simplified registers such as child-directed or animal-directed language, along with intercultural communication. In these scenarios, speakers intentionally simplify their communication, tailoring it appropriately to the respective audience. While there is a certain level of intuition involved when interacting with AI, given users' prior experiences with other 'Simplified Registers,' the process, in my opinion, primarily constitutes a strategic and conscious decision rather than a preconscious behavioral mechanism. Therefore, it is crucial to conceptually distinguish between preconscious behavior and conscious action in the ensuing discussion. These states of consciousness should not be perceived as a dichotomy but rather as poles within a continuum of degrees of awareness (see Chapter 3).

2.2 Historical foundations of the academic discourse on "Computer-Talk" (Zoeppritz 1985)

Fischer's conceptualization of "Simplified Registers" emerged within the context of the much older academic discourse on "Computer Talk", instigated by Magdalena Zoeppritz in 1985 based on initial experiments with users of early rule-based systems. Zoeppritz observed "several instances of deviant or odd formulations that looked as if they were intended to be particularly suitable to use with a computer as the partner of communication" (Zoeppritz 1985, 1). She explained these linguistic acts by proposing that users had a concept of the system's functioning in mind, tailoring their utterances accordingly, with a focus on the system's tool-like nature. To describe this phenomenon, she introduced the term "Computer Talk" (CT), drawing parallels to "Baby-Talk" or "Foreigner-Talk."[1]

Krause and Hitzenberger (1992) found numerous instances in their early German-language DICOS experiments with users of an early system for grade recording with a speech interface that supported Zoeppritz's (1985) concept of "Computer Talk." They observed simplifications of syntactic constructions, an increasing number of overspecifications, a growing amount of formal coding, a decreasing number of frame elements in the dialogue, a diminishing number of politeness phrases, a declining number of partner-oriented dialogue signals, and a reduced use of particles as markers for the speaker's personal disposition toward the spoken content. Krause (in Krause and Hitzenberger 1992) interpreted these as "metaphorical language use," wherein the actual metaphor lies in users tailoring their language use to the concept they have of the internal processes of the early language processing system.

Example 1: Krause and Hitzenberger (1992, 159–60)

User: Welche Deutschnote in Quarta hat wie viele Schüler?
[What is the German grade distribution in the fourth grade?]

User: Wieviele Schüler repetieren 1 Klasse?
[How many students repeat 1 grade?]

---

1 However, both terms are now problematic, as "Talk" inherently carries a derogatory, paternalistic connotation. In L1 and L2 acquisition research, these terms have been discarded in favor of "child-directed language" and "intercultural communication" (as mentioned above).

System: nicht verstanden
[not understood]
User: Wieviele Schüler repetieren 2 Klassen?
[How many students repeat 2 grades?]
System: 25
User: Wieviele Schüler repetieren 1 Klassen?
[How many students repeat 1 grades?]
System: 99

In this early phase of AI history, this cognitive concept is directed towards formal expressions (in programming language). However, an actual understanding of the system's architecture and programming is only partially present and varies significantly among users. Nevertheless, in Krause's early experiments, users tend to align more with a tool metaphor (AI as a tool) rather than an assistant metaphor (AI as an anthropomorphic conversational partner).

Krause and Hitzenberger (1992) characterize "Computer-Talk" based on their DICOS experiments as a structural register. Fischer, drawing on research data from the Verbmobil project, expands upon this assumption and conceptualizes "Computer Talk" more broadly as a "functional variety" (Fischer 2006) and later as a "Simplified Register" and "Robot-Directed Speech" (Fischer 2011, 261). Similar to Womser-Hacker's earlier observations on a structural level in Krause and Hitzenberger (1992) Fischer (2006) notes that HMI, in comparison to HHC, is distinguished by either an increase or decrease in lexical variety, syntactic complexity, and politeness markers. Her significant contribution lies in shifting the interpretative perspective from empirically structural features of CT to functional parameters and concepts of user cognition. "By looking at the peculiarities observable as strategies, we stop thinking of CT as a particular product and turn instead to the process in which it is created – a negotiation process" (Fischer 2006, 78). Before Fischer's 2006 analysis, early studies in HMI had a far too broad focus and a structuralist bias.

I am fully aware, that the methodological implications of Krause and Hitzenberger (1992) and Fischer (2006) are not neatly compatible with my praxeological attempt, but in order to create a model, that can include a broader range of empirical linguistic parameters I choose a more open approach.

HMI was then and remains a highly asymmetric interaction situation in which humans and machines process dialogicity quite differently, utilizing rather distinct resources. When we compare HMI and HHC, the asymmetry is immediately apparent and manifests empirically across all linguistic levels (refer to lexicon, syntax, semantics – particularly disruptions in dialogue coherence and grounding attempts – and linguistic politeness: Lotze 2016; Lotze and Ohrndorf, in preparation, concerning phonetics/phonology in Amazon Alexa: Greilich, in preparation). In all of our rather diverse studies we can observe, that users seem to transfer only the basic principles of communication from HHC (preconscious alignment, adjacency principle, frame sequences, concepts of registers, concepts of repairs, grounding, and framing), as long as the assistant metaphor is successful. When the dialogue design is geared towards it (e.g., in the case of social bots through textuality and multimodality in the form of clickable areas or in the case of Alexa through Voice User Interface (VUI) and RequestandResponse architecture) or when disruptions occur during the ongoing dialogue, users across all system types increasingly resort to simplifications or other markers of Computer-Talk (CT), such as vulgarisms, abrupt terminations of the conversation, etc.

2.2.1 User types according to Fischer (2006)

Another notable contribution by Fischer is the introduction of two user types: Players and Non-Players. What is particularly valuable about this distinction is that it involves open categories based on functional criteria. Characteristic of the Player type is treating the system as if it were a human interlocutor. The Player engages in the metaphorical game, addressing the system with personal pronouns like "du" [you without social distance] or "Sie" [you with social distance], offering greetings, and/or providing information about their own well-being when prompted by the system. On the other hand, the Non-Player type views the bot as a tool and utilizes it accordingly. They do not greet the bot, nor use personal pronouns to address it, and avoid politeness indicators. While the Non-Player demonstrates fewer transfers from HHC that can be interpreted through CASA/MASA, there are more instances of a "Simplified Register." Therefore, it is crucial to consider both approaches together.

Both types are defined by the conversational strategies they employ based on their assumptions about the AI. Consequently, their utterances become somewhat predictable. Fischer suggests that one can infer the user's category based on their behavior in the opening sequence. If the user responds to the system's greeting, they are a Player; if they ignore it, they are a Non-Player. Lotze (2016) was able to replicate this fundamental distinction between Players and Non-Players, but states that user types are more complicated and not always dichotomous.

2.2.2 The heterogeneity of HMI (Lotze 2016)

How should we define HMI then? As the attribution of social characteristics to the system, accompanied by the transfer of linguistic behavior from HHC (CASA/MASA)? Or as, in any case, partially a strategic user decision for a simplified register in the sense of bot-directed speech according to Fischer (2011)? Does the interpretation of conflicting linguistic evidence in different studies lead to a dilemma?

In my dissertation (Lotze 2016, 346–47), I argue that this perceived dilemma can be easily resolved. HMI is, after all, a genuinely heterogeneous form of interaction that varies depending on system architecture, application context, user type, and awareness level. Therefore, we need a model that accommodates the entire variability of HMI by considering all relevant parameters and not focusing solely on individual aspects. The following variables must be taken into account when interpreting HMI data, as they all have a significant impact on HMI and contribute to the variation in user language. Accordingly, the asymmetry of HMI is not a monolithic feature but manifests in very different factors and variables that are all interconnected and have an important influence on the users' language behavior and stratetgies.

Levels of asymmetry (Lotze 2016, 346):

a) External Factors

- The scenario of the application domain determines the interaction situation.

b) System Variables

- Persona
- Robot, avatar, or interface design
- Hardware
- Input channel
- Dialog design
- System architecture

c) User Variables

- Technical expertise
- User type
- Assumptions about the system
- Dialog goals
- Pre-conscious priming
- Conscious action strategies

The heterogeneity of the HMI can only be addressed by a multi-dimensional conceptualization, taking into account that the HMI is influenced by numerous factors, and these factors are highly asymmetrical on the part of both the system and the users.

Table 1: Dimensions of the dialogically inherent heterogeneity of HMI (Lotze 2016, 348)

System Architecture	**User Guidance: guided – free – hybrid**
Dialog Design	· Textuality – Orality · Social Distance – Proximity · Different Phases of Dialogue (Introduction – Middle – Farewell) · Handling of Disruptions: Incoherences, Quasi-coherences, Default Responses, or Follow-up Questions
User	· User Type · Conscious Strategic – Preconscious or Routinized · CT („Computer Talk") – HHC (Human-Human Communication)

The factors I listed in 2016 remain relevant for current systems with Natural Language Processing/Understanding (NLP/U) and Machine Learning (ML), as well as GPTs. In the HMI, individuals who are self-aware and wish to freely choose and negotiate their dialogue goals still encounter machines that still exhibit significant challenges in these aspects. As mentioned above, this asymmetry manifests across all linguistic levels. Design decisions regarding system architecture and dialogue design are, of course, impactful for interaction. However, the type of user and whether they consciously act or preconsciously react are equally consequential for dialogue. Therefore, we must consider the following levels of asymmetry in the HMI with the following effects on dialogue:

Levels of asymmetry of HMI and their effects

User – System: Humans and machines fundamentally differ in terms of "world knowledge" (Habermas 1993), emotions, (first-) language acquisition, and self-reflective consciousness. This has significant implications for dialogue semantics and coherence.

$System_A$ – $System_B$: Systems differ significantly from one another. Different technical approaches are currently used for various applications, and their functionalities should not be generalized. System architectures, dialogue designs, and the mediality of interfaces (oral, literal, embodiment) vary. This affects the chosen simplification strategies of users. The technological affordances of the system, in general, are just as relevant as its degree of anthropomorphism.

$SystemA_{error\text{-}free}$ – $SystemA_{error\text{-}prone}$: Errors are a particularly relevant factor contributing to the heterogeneity of the HMI because users must reconsider their dialogue strategy in such situations. One possible consequence is that users transfer repair strategies from the HHC to the HMI (for grounding, see Fischer 2006). Since these often fail, an extreme outcome may involve a user type switch, where a cooperative, flexible, polite player transforms into a non-player who vulgarly insults the system and abruptly ends the dialogue without a farewell. The reverse principle we currently observe in users of ChatGPT who initially attempt to operate the system with isolated keywords but then switch to more elaborate prompts when they realize that the system is capable of generating longer sequences of disruption-free dialogues (see below).

$User_A$ – $User_B$: User types (Player / Non-Player) according to Fischer (2006; 2011) have implications for the level of conscious cognitive reflection and, consequently, the chosen linguistic register (see above).

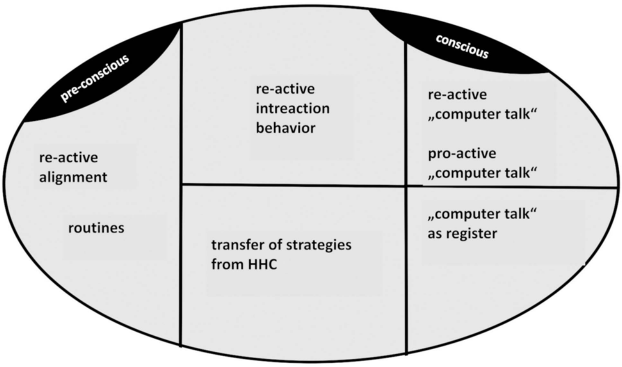
$User_{t1}$ – $User_{t2}$: Users fundamentally change their strategy when it fails (see errors and disruptions). This can occur in specific sequences without an immediate full user-type switch (extreme case). This also has implications for the level of conscious cognitive reflection, consequently affecting the chosen register, and makes the contributions of the same user in different sequences heterogeneous.

Context$_A$ – Context$_B$: Application contexts can vary extremely, impacting the attribution of social proximity or distance to the system, which linguistically reflects in politeness levels, etc.

Time$_{t1}$ – Time$_{t2}$: Mediatization manifests in diachronic variation and change affecting both systems (technological history from rule(+plan)-based systems to Big-Data approaches with NLP/U and ML, as well as GPTs) and users (Digital Non-Natives, Digital Natives, GenZ), who develop different strategies/styles/registers to interact with respective system types. Thus, in our diachronic corpora for the past 20 years, we can observe, in my opinion, how Krause's metaphor in user reception has shifted from "code" to "natural language 'enter'-key" (confirming a predefined dialog script) and "isolated keywords" (as a concept from Google search).

Within this theoretical framework that considers all relevant variables of HMI, the heterogeneity of HMI manifests empirically as user language as follows:

Figure 1: Composition of HMI (Lotze 2016, 359)



Preconscious behavior: In the preconscious realm, we find numerous instances of *re*-active alignment as a lower-level priming effect; i.e., humans adapt to the system – phonetically, syntactically, and lexically. We cannot speak of *inter*-active alignment (Pickering and Garrod 2004) here because it is not built up interactively or collaboratively. With Lotze (2016), I am still not referring to interactive alignment in the strict sense (Latin: *inter-agere*) but rather to the user's reactive alignment to the system. Also, routines transferred from HHC, such as turn construction and allocation, politeness levels, greeting sequences, etc., can be substantiated through our studies. These two aspects can only be interpreted within the CASA/MASA paradigm. However, the interaction has both preconscious and conscious components, and the better the illusion of a natural dialogue is maintained, the more "mindless behavior" (cf. Reeves and Nass 1996; Nass and Moon 2000; Nass and Brave 2005) is exhibited by the users. In contrast, during disruptions, the artificial dialogue situation must be reflected upon, and conscious strategic behavior is the logical response (cf. Fischer 2006).

Transitional behavior / strategies: Not only preconscious aspects of the HMI can be interpreted within the CASA/MASA framework, but also some of the conscious proactive action strategies involve transfers from HHC. On the functional level, for instance, all attempts by users to establish common ground or create dialog coherence (grounding, repair) can be interpreted as the users anthropomorphizing the system. Even though users, in most cases, theoretically know that systems cannot draw upon the same world knowledge as they do, they sometimes intuitively strive to promote common ground and a logically coherent dialog progression. However, this does not apply to a large portion of users. These reactive consumers of the HMI allow themselves to be guided by the system and do not attempt to address its logical-semantic deficiencies. This results in a reactive interaction that cannot be interpreted as a transfer from HHC but also does not align with CT in the narrower sense. Nevertheless, we frequently observe this passive behavior in our empirical studies, especially among digital natives of the player type who passively let the bot guide them through the application without a specific dialog goal. These two functional user attitudes can be interpreted as a transitional zone between CASA/MASA and CT. Therefore, empirical evidence suggests a continuum between preconscious behavior and strategic CT.

Conscious / strategic decisions: The scope of CT does not encompass the entire HMI, as it is heterogeneous and sometimes exhibits longer sequences of human-like dialogue, especially in contemporary applications. What can be termed as CT must be negatively defined as the subset of HMI where preconscious mechanisms (preconscious alignment, routines) or transferred strategies from the HHC (grounding, framing) do not apply. This subset can be functionally further subdivided into a) a reactive CT, directly triggered by the technological affordances of the system's architecture and dialog design (s.b.) and b) a classic, proactive CT derived from users' assumptions about the system (Krause's metaphorical language use). For the reactive form of CT we found some interesting tendencies in user behavior (triggered by the technological affordances):

- rule-based system – user behavior: isolated keywords
- plan-based systems – user behavior: passive reception attitude and "natural language 'enter'-key" ("ok", "continue", "back")
- request and response systems – user behavior: isolated imperative sentences

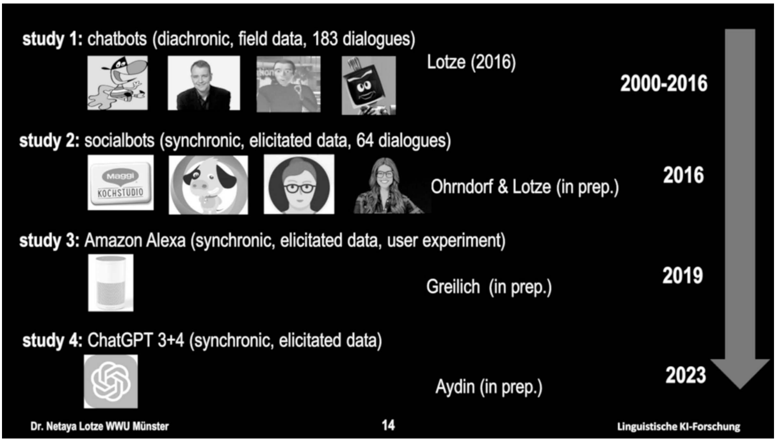
Both types of CT have functional and structural dimensions and undergo developmental processes. Function and structure do not always develop in tandem. For example, lexical and syntactic simplification today serves different functions than it did in the 1990s (programming language as a metaphor vs. keyword-based Google search as a metaphor). Overall, CT represents only an extreme case of linguistic user behavior that can be perceived as an outer pole within a continuum of user language.

3. How Do Users Linguistically Interact With AI in Our Empirical Studies? Alignment, Acceptance and Simplification (AAS)

Demonstrating the heterogeneous nature of HMI as a form of interaction, we could empirically substantiate our findings using various methods for diverse user groups with field and experimental data since the year 2000. In this article, I aim to provide an overview and, as a conclusion, present my model for HMI as a complex socio-linguistic practice. As mentioned earlier, I can only present the most relevant aspects of every study.

Study 1 is my dissertation, where I worked on a micro-diachronic level, analyzing user language in rule-based and plan-based, media-written chatbots from 2000–2016. These systems were all used in the help-desk sector and were more or less advanced for that early stage of technological development. It represented a first description of HMI using a mixed-method approach with qualitative (conversation analysis) and quantitative methods (corpus linguistics). The data consisted of system log files from various application scenarios with real users, providing high ecological validity. Human-to-human chats with help-desk character served as a parallel corpus (library information, Chat-Korpus Beißwenger and Storrer 2004). Statistical analysis of corpus data included relative frequencies, distance-frequency analyses, and inferential statistics.

Figure 2: Our studies over the past 20 years



In 2016, Study 2 applied the mixed-methods approach (qualitative and quantitative) to Socialbots on Facebook Messenger, conducting a synchronous analysis of customer support bots in that context.

Study 3 and 4 constitute projects undertaken by my research group. Study 3 is a psycho-linguistic, hypothesis-testing experiment focusing on user strategies in oral interaction with Amazon Alexa (in collaborative tasks). Study 4 is a purely qualitative first description of written interaction with ChatGPT in elicited dialogues with the AI in two collaborative tasks (travel planning and essay writing) following an ethnomethodological approach.

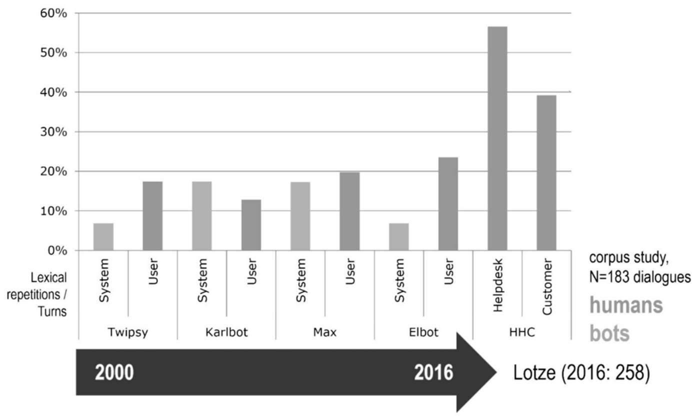
Even though our studies address different types of systems (rule-based, plan-based, VUI for RequestandResponse, GPT) and different modalities (written/oral), we observe similarities in the user language. In a simplified view, across older systems, social media systems, oral VUI, and the innovative GPT, we identify three fundamental tendencies in user language: preconscious (reactive) alignment; reactive adaptation strategies to the affordances of system architecture and dialog design, as well as simplifications in the sense of a "Simplified Register" (in extreme cases, even as CT according to Zoeppritz 1985). Alignment, acceptance, and simplification can be abbreviated to the acronym AAS, representing the main aspects of a heterogeneous HMI.

3.1 "Alignment" as a Preconscious Phenomenon

Interactive alignment in HHC, characterized by the tendency to adapt one's language use to that of the interlocutor (Hartsuiker et al. 2000; Pickering and Garrod 2004), serves as a good example of preconscious behavior, given reaction times in the microsecond range. Perception and reception are so closely linked in HHC that a form just perceived remains cognitively active when people begin to produce their own contribution. Thus, it is more likely to reproduce what has just been perceived.

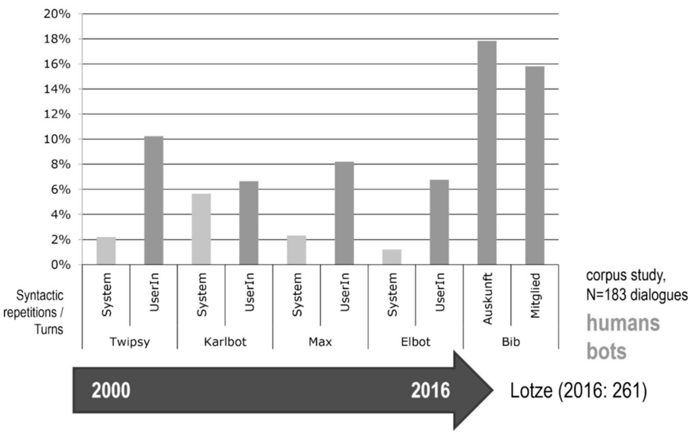
Reactive alignment of the user to the bot: Alignment in HMI has been demonstrated in various studies across all linguistic levels (Branigan et al. 2000; Branigan and Pearson 2006; Huiyang and Min 2022; Heyselaar 2017; Raveh et al. 2019; Lotze 2016; Fischer 2006; Linnemann and Jucks 2018).

Figure 3: Lexical alignment (Users of rule-based and plan-based systems) (Lotze 2016, 258)



In our corpora, it plays a less prominent role for users of our older systems compared to the HHC reference corpus (Lotze 2016, 254–55). Nevertheless, it appears consistently in every dialogue (approximately as frequently as in HHC) and has been identified on the syntactic and lexical levels. Nevertheless, these adoptions can be seen as indicating a transfer of the basic concept of interaction per se from human-human communication, especially among users of the more recent systems. The better the system works, the more the user goes along with the illusion. What is evident in my corpus studies on lexical alignment in HMI, which manifests as user repetitions of the systems lexis, is, that humans adapt less to the language of the system than to a human (on average 50 percent less for lexis and syntax) (c.f. Lotze 2016; 2018; 2019).

Figure 4: Syntactic alignment (Users of rule-based and plan-based systems) (Lotze 2016, 261)



However, we have to distinguish for lexical persistencies for each individual instance whether it is preconscious adaptation, socially motivated strategic adaptation, or a simplification strategy. In the latter case, a word form, that the system itself has already used, is selected for the user's own turn, because the user assumes, that this keyword is stored in the system's database. Thus, lexical alignment of the user can be interpreted either as a pre-conscious mechanism or as a strategic adaptation with different motivations (preconscious alignment as attribution of social proximity vs. simplification for the machine). Of course, as researchers, we can only speculate about the actual intentions behind the users' alignment to the system. However, the HMI research community agrees that users' intentions can vary and may also change over the course of a single dialogue, even for the same user.

Syntactic alignment is less frequent but shows the same trend and the interpretation as pre-conscious alignment is obvious here. With syntactic alignment, it becomes much clearer that in these instances users are not trying to find the right keyword. Instead, they are not aware of the adoption of the entire syntactic structure on a conceptual level. In the following example the user adopts the syntactic form of the predicative clause from the bot, even though there is a change of topic in the example, and the lexis is not adopted.

Example 2: Lotze (2019, 314)

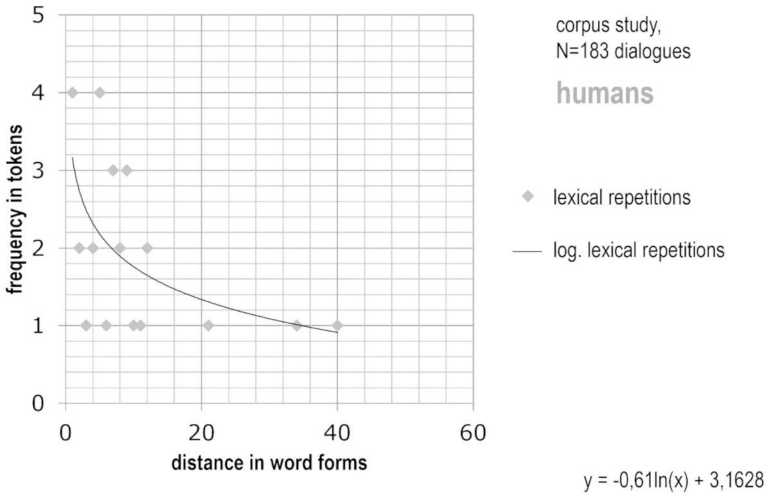
Max: **Das** [Nominativ] **ist** [Kopulaverb] **deine Meinung** [Nominativ].
*[**This is your opinion** [predicative sentence with "to be"]]*

User: **Stefan** [Nominativ] **wird** [Kopulaverb] **Informatiker** [Nominativ]
[Stefan will be a computer scientist? [predicative sentence with "to be"]]

In this interaction with the Max system at Bielefeld University, which was being tested by its developers at the time, Max concludes a prior dispute with the statement 'That's your opinion,' effectively suggesting to agree to disagree. The user then changes the topic and addresses the career ambitions of one of Max's developers (Stefan Kopp) with the remark, 'Stefan will become a computer scientist.' Despite the abrupt topic shift, the user syntactically aligns with the preceding system turn in form of a predicative clause.

Does human memory in HMI differ? Overall, the cognitive processing of dialogue by users in HMI is not fundamentally different from Computer-Mediated Communication (CMC) (more nuanced: Lotze 2016). For example, the rate of decay of the primes repeated by the user follows the "forgetting curve" of Ebbinghaus (1985); i.e., for users in HMI, a linguistic structure produced by the system becomes gradually less relevant, and repetitions by the user become rarer, as is the case in human-human communication as well. These research results support the idea that alignment in HMI is also preconscious behavior.

Figure 5: Rate of decay of primes (Users of rule-based and plan-based systems) (Lotze 2016, 279)

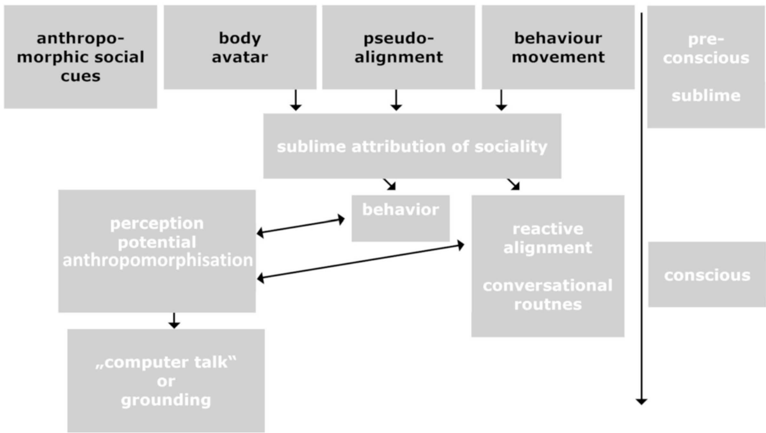


Thus, a model of interactive alignment in human-machine interaction has to be understood as a continuum of degrees of consciousness between automatic or sublime/subtle and conscious or strategic behavior. For this purpose, I added the new component "alignment" to the schematic representation of the social effects of artificial agents according to Krämer (2008, 202).

Reactive alignment of users is thus a proven phenomenon and, depending on the level of awareness and motivation, can be interpreted as: (a) an automatism of human dialog behavior (b) a linguistic indicator for attributing social characteristics to the system (c) a conscious simplification strategy (e.g. search for the right keyword) in the sense of a "computer talk" (CT) according to Zoeppritz (1985). Against the background of the discussion about dangers of interaction with AIs it then must be classified as (a) an unavoidable cognitive process, (b) as problematic because the system can only disappoint the expectations of the users and (c) as media literacy of users, who are familiar with the system. The example of "alignment" shows clearly, how multilayered the psycho-linguistic interpretation is.

The model can also be applied to other modalities. Greilich (in preparation) found even more alignment in the oral HMI among users of Amazon Alexa (in lexicon, syntax, and prosody). She conducted an experimental study as a collaborative task with Amazon Alexa with focus on referential expressions and topic continuity, both often manifesting in repetition. And she found not only more alignment in oral HMI, but also over longer sequences of persistencies in triplets or quadruplets.

Figure 6: Model alignment in HMI (Lotze 2016, 390)



Example 3: Greilich (in preparation)

User: Alexa, was sind die Öffnungszeiten?
[Alexa, what are the office hours?]

Alexa: Die Gemäldegalerie Alte Meister in Dresden öffnet in 4 Minuten?
[The galery Alte Meister in Dresden is going to open in 4 minutes?]

User: Alexa, wie lange ist sie am Montag geöffnet?
[How long is it open on Mondays?]

Alexa: Sie haben am Dienstag von 10 Uhr bis 18 Uhr geöffnet.
[They open on Thursdays from 10 a.m. to 6 p.m.]

In one of the scenarios of the collaborative tasks of the experiment the test users had to ask Alexa for information on Dresden from a tourist perspective.

With reference to the discussion about CASA/MASA vs. “Simplified Registers,” we can establish the following aspects based on our empirical findings:

a) Reactive alignment can be interpreted as a transfer from HHC to CASA/MASA.
b) The phenomenon is so stable in HMI that it occurs even in medial-written interaction with old, extremely error-prone plan- and rule-based systems and is cognitively processed regularly (forgetting curve according to Ebbinghaus 1985).
c) The phenomenon depends on modality and intensifies in orality (probably due to the anthropomorphic voice and shortened reaction time).
d) Strategic alignment as a search for the appropriate keyword by users must be interpreted as a simplification strategy in the sense of a "Simplified Register" (Fischer 2011).

In addition, all ritualized aspects of interaction such as turn construction and allocation, politeness levels, and ritualized greeting sequences can be interpreted as "mindless behavior." We find numerous examples of all these investigative parameters in our studies (further explored in Lotze 2022).

3.2 "Acceptance" as a Transitional Phenomenon

At the transition between preconscious behavior and strategic action[2], we find a) highly frequent passive reactions to the affordances and restrictions of system architecture and dialog design[3] and b) transfers of proactive strategies from HHC (e.g., grounding as a repair strategy). In the former case, it is an *affordance-bound*, passive receptive stance of users that guides them through mainly plan-based applications in the quickest way without disruptions. In the latter case, it is an *affordance-unbound* user reaction, which is indeed a conscious repair, but often not a conscious decision of the users since the older systems in our corpora lack any world knowledge (cf. Habermas 1993). Such user strategies can still be conceptually interpreted as transferred concepts from HHC according to CASA/MASA on the cognitive level, but functionally, they represent a conscious repair strategy.

2 The processes identified by Pickering and Garrod are preconscious and thus automatic, meaning they occur prior to more complex processes of conscious interpretation ($t < 600$ms, see Pickering & Garrod 2004).

3 rule-based systems – user behavior: isolated keywords; plan-based systems – user behavior: passive receptive stance ("ok", "continue", "go back"); RequestandResponse systems – user behavior: isolated imperative sentences

User language that lies in this borderline area and we can only interpret it with a model that considers such a transition zone. Furthermore, we need to differentiate between affordance-bound acceptance and affordance-unbound acceptance. A "passive receptive stance" of users is found especially in plan-based systems that guide people step by step through decision tree-based dialog scripts, requiring a lot of confirmation in the dialogue. In extreme cases, users deviate from their original dialogue goal and let the AI passively guide them through the application. This example of an interaction with the Max system comes from the period when it was deployed as a virtual museum guide at the Heinz Nixdorf Forum in Paderborn.

Example 4: Max corpus 501–526
Max: Should I show you the next exhibit?
User: no, go back
Max: The next exhibit is the AI exhibition. Should I explain that?
User: ok

Krause's metaphorical language use (Krause and Hitzenberger 1992) has thus undergone a shift in its pragmatic function – away from the metaphor of actively operating a machine to a) cooperation with the system in processing specific tasks (cf. RDS, Fischer 2011) and b) a passive receptive stance towards conversational technology (Lotze 2016, 358). The latter manifests as reactive user behavior closely tied to the bot's instructions. An extreme example of this is the absence of interventions in case of disruptions.

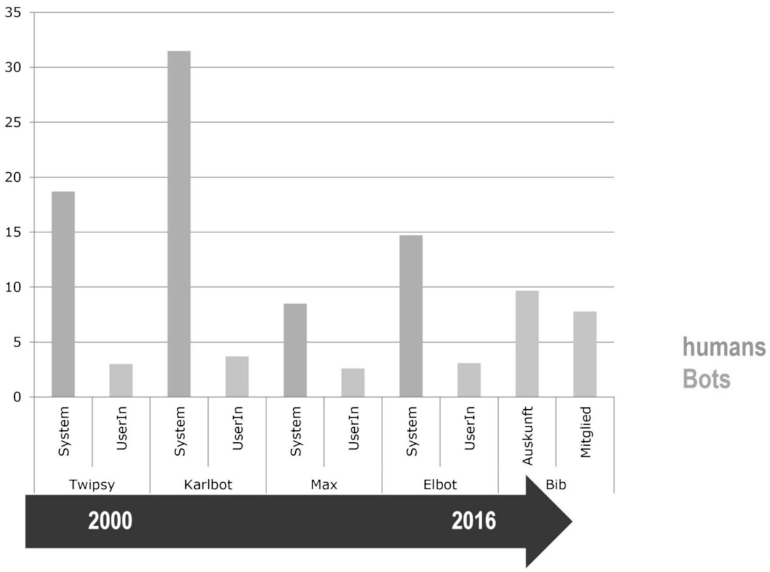
3.3 Simplification as an Affordance-Bound and Affordance-Unbound User Style

Simplifications by users are numerous across all examined systems on different linguistic levels, regardless of modality. Comparing the turn lengths of users and systems in the older rule- and plan-based chatbots, users consistently formulate extremely short turns, regardless of the length of the bot's turns. One could argue that this is due to the helpdesk scenario[4] with short questions and detailed responses, which represents the context of all corpora

4 All corpora examined in the 2016 study were log files of interactions with various chatbots in help-desk scenarios, specifically in customer support. The parallel corpus for human-human communication was the chat corpus by Beißwenger and Storrer (2004).

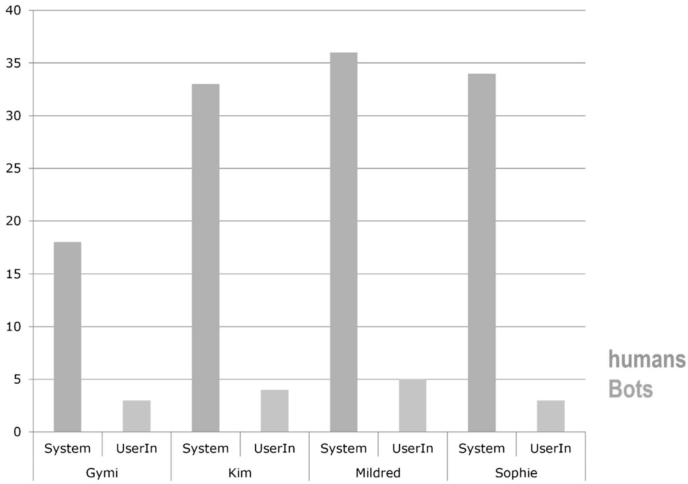
in this study. However, looking at the comparison corpus of CMC (computer-mediated communication) related to HHC, which was selected as a parallel corpus precisely because it is also a helpdesk, the quantification of the corpus study clearly shows that people adapt to each other regarding turn length, and this effect is stable even in the written medium.

Figure 7: Length of turns in chatbots (diachronic) (Lotze 2016, 234)



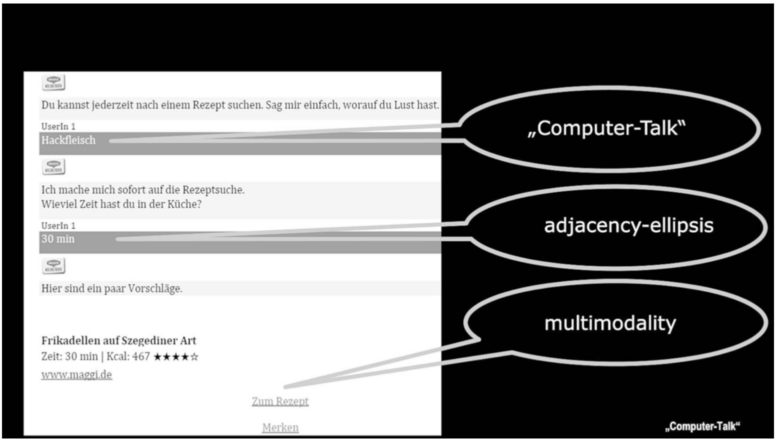
We could replicate this result in 2016 for a more recent type of intent-based social bots on Facebook Messenger and were able to reproduce my micro-diachronic study with corpus data from 2000–2016. Even with these significantly improved systems, turn lengths vary greatly, and users tend to become silent. This leads to various additional structural and functional simplifications (for further details, refer to Lotze 2016).

Figure 8: Length of turns in socialbots (synchronic) (Lotze and Ohrndorf, in preparation)



If the interfaces are additionally designed to be multi-modal with clickable areas, buttons or images, the effect is further enhanced.

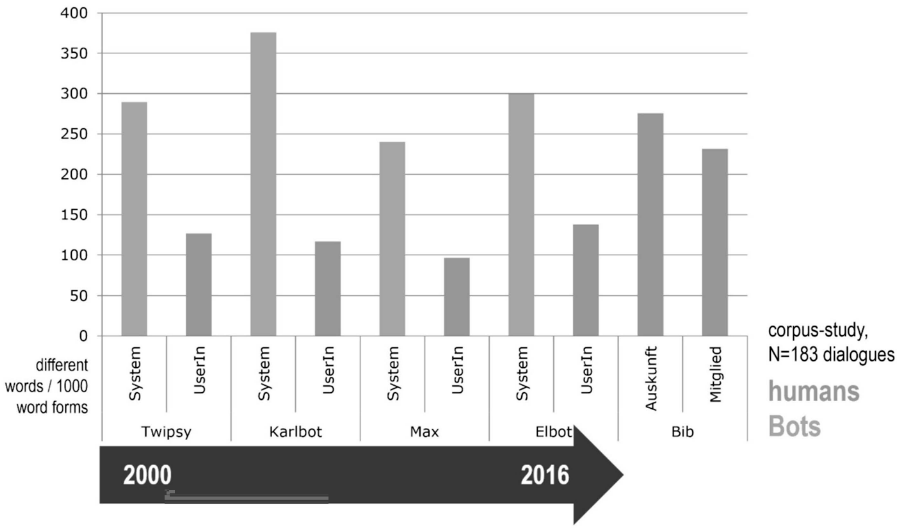
Figure 9: Isolated keywords in users of socialbots (Lotze and Ohrndorf, in preparation)



The example illustrates aspects of structural and functional computer talk at the interface between the desktop metaphor and the assistant metaphor, as well as the metaphor of Google search (for Krauses notion of metaphor see cap. 1.2.1, for further details s. Natale and Cooke 2021). The interfaces of the examined socialbots were primarily operated by their digital-native users using isolated keywords and adjacency ellipses related to the bot's previous turn. The tool metaphor dominates in this generation of users.

Not only does the turn length decrease, but lexical variability also decreases compared to computer-mediated communication (CMC). The Type-Token Ratio of users is significantly lower than in the parallel corpus, and that of bots is significantly higher, further emphasizing the asymmetry. In the older information bots, lexicon and syntax were primarily oriented toward written texts in a brochure, not the medial-written, quasi-synchronous dialogicity of CMC. This explains the richness of different lemmata, especially in the two oldest systems.

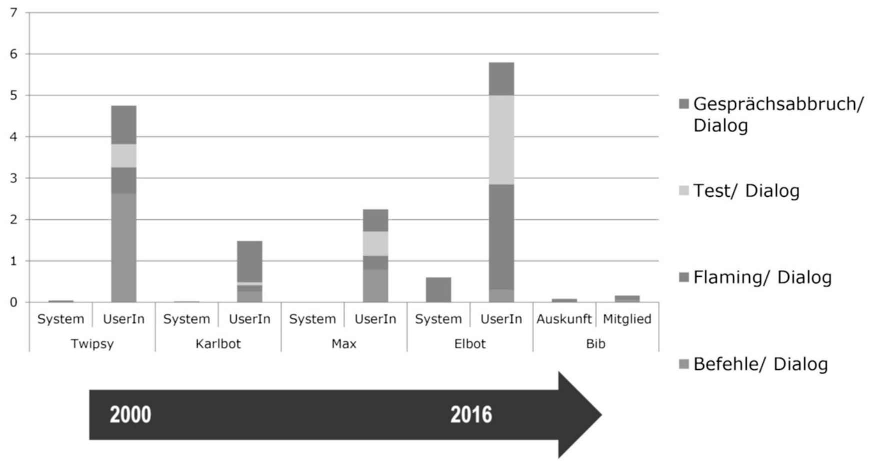
Figure 10: Simplification of lexical variation (Lotze 2016, 322)



Concerning syntax, users of rule- and plan-based systems utilize ellipses approximately 60–70 per cent of the time, not all of which are adjacency ellipses; some are also isolated keywords, simple imperatives, and confirmation signals. The remaining 20 per cent consist of simple sentences, often with a copula or main verb in the imperative form. Only 10 per cent form complex sentence structures with subordinate clauses. In comparison to the CMC reference corpus, ellipses constitute 40–50 per cent, predominantly being adjacency ellipses. In terms of syntax, humans strongly adapt to each other, while the frequencies for users and bots differ significantly, primarily due to the conceptual nature of the bot turns (Lotze 2016, 327).

Politeness in language remains a contentious research field in HMI, as studies yield different results depending on the application context, user type, and sophistication of the AI (see, e.g. Clark and Fischer 2023). Indicators of actual CT, following Zoeppritz (1985) as the extreme pole of a simplified user language, can only be interpreted through expressions that would be completely dispreferred in human interaction: isolated imperatives, vulgarisms (flaming), abrupt conversation interruptions, and playful testing of system functions by asking the bot personal, emotional or particularly complex questions. Instances of these forms of expressions are found in users of rule- and plan-based systems between 1.5 to 6 times per dialogue.

Figure 11: Linguistic (im)politeness (Lotze 2016, 338)



This result is particularly interesting as it demonstrates that impoliteness occurs more frequently in studies when the investigation corpora consist of unaltered field data (log files) with high ecological validity (as in Lotze 2016). Therefore, it can be inferred that users communicate more impolitely and directly with bots in real-world scenarios than in experimental settings. In her elicited data on Amazon Alexa, Greilich (in preparation) identifies more simple imperatives without politeness markers, but no vulgar language. This suggests that imperatives and isolated keywords depend on the affordances of the RequestandResponse architecture in orality. The absence of vulgarisms can most likely be analyzed as an experimenter effect.

Example 5: Imperative as affordance-bound simplification (Amazon Alexa, Greilich, in preparation)

User 1: Nenne mir bekannte Verfilmungen von Thomas Mann (Participant 1, Question 2, Attempt 2)
[Name famous film adaptations of Thomas Mann]

User 2: Erzähle mir etwas über die Werke von Thomas Mann (Participant 4, Question 2, Attempt 2)
[Tell me something about the works of Thomas Mann]

User 3: Finde mir bitte das Alter von Leonardo DiCaprio raus (Participant 7, Question 1, Attempt 1)
[Please find out the age of Leonardo DiCaprio for me]

More challenging to interpret are user utterances that are not triggered by the affordances of the dialog system.

Example 6: Isolated keywords as affordance-unbound simplification (Amazon Alexa, Greilich, in preparation)

User 4: Thomas Mann Nobelpreis (Participant 9, Question 2, Attempt 3)
[Thomas Mann Nobel Prize]

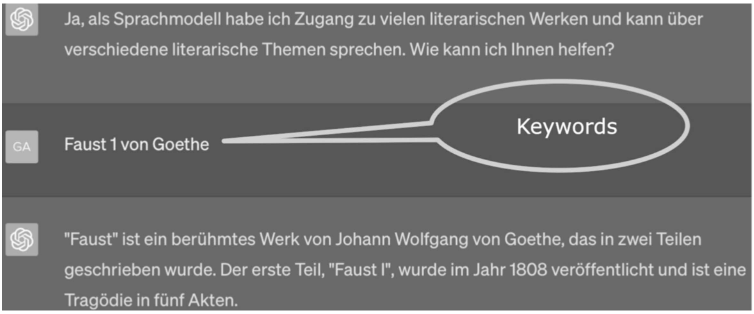
User 5: 5 Tickets Gemäldegalerie Alte Meister (Participant 9, Question 3, Attempt 6)
[5 tickets art gallery Alte Meister]

These affordance-unbound simplifications can indeed be interpreted as "metaphoric language use" according to Krause and Hitzenberger (1992), as users here experiment with the new metaphor of Google search in the form of isolated keywords. This provides deeper insights into users' assumptions about the system's functions. Alexa, as a voice-based assistant, has been optimized with a RequestandResponse architecture for oral commands for home automation and has been trained with empirical speech data. This suggests operation in complete sentences with finite verbs (especially in the imperative). However, some users transfer the cognitive concept of keyword-based bots and Google search from text to oral interaction with Alexa. For keyword-based searches, a socio-linguistic practice seems to have already developed, as indicated by the example of social bots (Fig. 9).

Even in generative transformers based on large language models (LLMs) that perform better when longer contexts are made explicit, we find affordance-unbound simplifications and evidence of reactive user behavior.

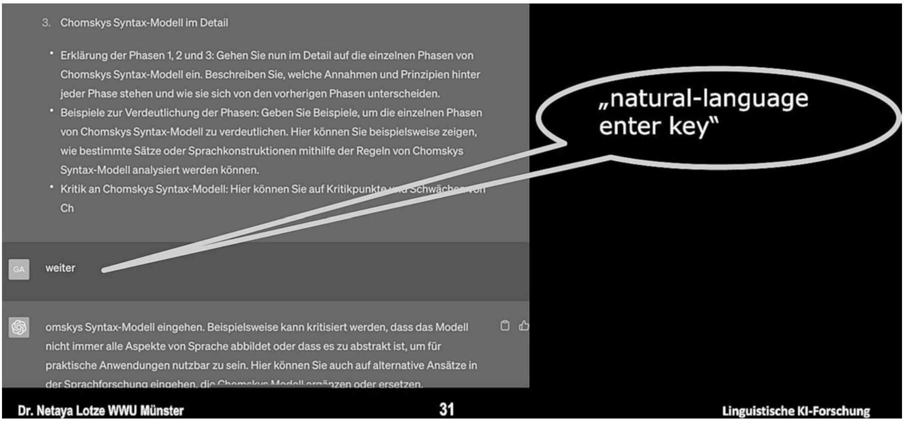
Example 7: Collaborative travel planning (1–3) and essay task with ChatGPT (Lotze and Aydin, in preparation)

Figure 12: Isolated keywords in ChatGPT



In the most extreme manifestation of this form of acceptance and passive user behavior (as in the example above), the user merely confirms the suggestions provided by the bot with "yes", "ok", or in the case of this example "next", which can be analyzed as an equivalent of the 'enter' key in natural language. This reactive strategy has evolved in the past, especially among users of plan-based systems, who playfully and exploratively let the system guide them through the application in this way. Now, they transfer this concept to interact with ChatGPT, thereby rendering the practice of "natural language 'enter'-key" no longer strictly interpretable as affordance-bound.

Figure 13: Reactive behavior in users of ChatGPT

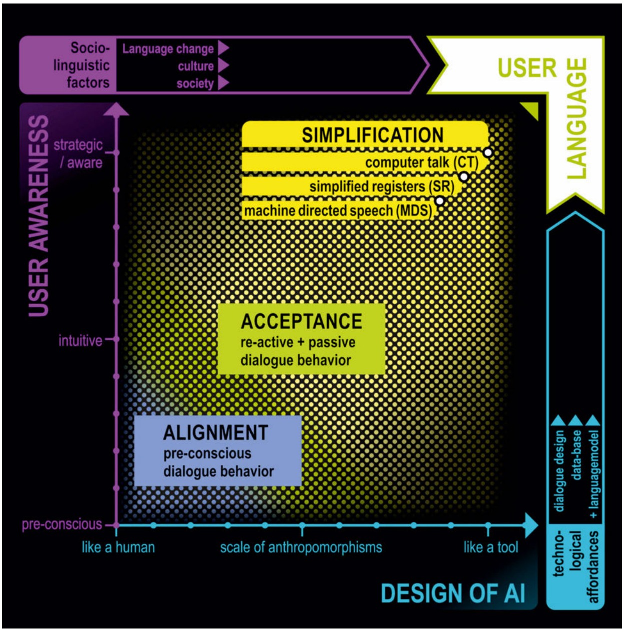


When technology changes, user strategies evidently do not shift immediately but with a time delay (cf. the "stylistic inertia law" (*stilistisches Trägheitsgesetz*, Schmitz 2015, 25–26; cf. Hauser 1958)). Therefore, if the simplification strategy does not align with the affordances of the (new) technology and cannot be understood as a transfer from the HHC (e.g., child-directed speech, etc.), Krause and Hitzenberger's (1992) idea of metaphorical language use for this small subset of linguistic simplifications within the HMI, in my opinion, remains relevant. Conceptual metaphors seem to undergo a diachronic change, which follows the technological revolutions with a delay. Metaphor is a concept that can manifest linguistically and structurally in various ways, and it must be considered partially independently of the affordances of technology and mediality (see above).

For a (still extremely young) diachronic research perspective on HMI, this means that user concepts also undergo changes over time. This becomes apparent whenever user assumptions about how the technology works lag behind: code as a metaphor for operating the first natural language interfaces, isolated keywords, and the "natural language 'enter'-key" (or dialog scripts) as metaphors for operating learning-capable and pre-trained systems based on large language models (LLMs). The time-delayed adaptation of the metaphor is interesting for linguistic discussions because in transition phases, one can observe that the affordances of technology do not directly trigger linguistic behavior but are always mediated by cognitive concepts, which are only discarded when they are no longer efficient. The conceptual level and the medial level are not always congruent in HMI (for HHC: see Koch and Oesterreicher 1994; for CMC: see Dürscheid 2003).

4. A Model for HMI as a Complex Socio-Linguistic Practice

Figure 14: AAS-Model of HMI as a complex socio-linguistic practice



In the following, a model for HMI as a complex socio-linguistic practice will be introduced. To do this, I will first present its three continua: a) a technological continuum, b) a human-cognitive continuum, and c) a human-linguistic continuum, with the latter being partially dependent on the first two. However, a strict dependency of user language on the degree of system anthropomorphism and individual cognitive awareness cannot be postulated, as language, in general, is subject to numerous social, cultural, and historically grown factors and undergoes language-specific changes. Moreover, the model is by no means deterministic but assumes spontaneous, flexible, and adaptable users. The multi-dimensionality of the model provides users with the ongoing opportunity to linguistically position themselves depending on the degree of system anthropomorphism. This occurs partly at a preconscious level and partly consciously and strategically. If preconscious behavior or a conscious strategy fails, the same user can reposition themselves in the interaction diagram. Concrete linguistic structures, manifested as lexical, syntactic, or phonological forms, can be analyzed by the linguistic community using the model.

Dimension 1: Technological affordances and anthropomorphic design

The technological dimension is graded in degrees of the anthropomorphization of the system (cf. Ruijten et al. 2014; 2019). This dimension is adopted from MASA (Lombard and Xu 2022). It refers not only to the degree of anthropomorphism of the system as an interface or robot doll but also includes the cognitive reception by its users. As reception effects coincide with the degree of anthropomorphism, Ruijten et al. (2014; 2019) argue that these parameters can be combined as one parameter for human reception of more or less anthropomorphic systems.

I would like to expand the visual, movement-based etc. anthropomorphism and its reception by the degree of linguistic anthropomorphism, which appears more relevant from a linguistic perspective. The gradual variation here is crucial, indicating to what extent a natural language dialogue succeeds in being coherent and cohesive over longer sequences or, for example, only at individual adjacency pairs (cf. old rule-based bots, and partially RequestandResponse systems). Therefore, the anthropomorphism of pragmatics in AI per se is particularly important for interpreting our data of interface-based AI (oral and written). Additional parameters in our studies include the voice or name of the AI, as in the case of Alexa. Regarding robotics, we cannot make any statements based on our own studies and rely primarily on Clark and Fischer (2023), Fischer (2011), Habscheid et al. (2018), Lenz et al. (2019), and can build upon MASA (Lombard and Xu 2022).

Dimension 2: Cognitive awareness levels of the user

Human consciousness can be understood as a temporally staggered phenomenon, ranging from “pre-conscious” to “conscious,” as cognitive availability hierarchies organize processing in the brain chronologically. For the cognitive processing of HMI by users, alignment as a lower-level priming presents a key phenomenon (see above). The processes considered by Pickering and Garrod (2004; with Gandolfi 2023) are pre-conscious and automatic, i.e., they occur temporally before more complex processes of conscious interpretation (t < 600ms, cf. Pickering and Garrod 2004). Lower-level priming alone can be understood as the driving force of the interaction in this area. The interactive alignment model (Pickering and Garrod 2004) does not provide information about factors related to conscious interpretation. Its mechanisms must precede considerations of the social goal orientation or intentionality of utterances both temporally and logically. Lower-level priming thus constitutes the starting point of the second dimension.

Especially assigned to consciousness are the activated memory, focal attention, and controlled (non-automatic) processes of information processing (cf. Wirtz 2021). Reflected, thoughtful user strategies that intentionally pursue their own agenda or individual dialogue goal accordingly represent the endpoint of the consciousness continuum.

Fischer takes initial steps in this direction by defining CT as functional (2006) and as a Simplified Register (2011). She assumes conscious user strategies that control dialogue behavior depending on assumptions about the bot and the user type. She emphasizes the tool character of user language. This contrasts with preconscious cognitive alignment as the cause of preconscious user behavior. Depending on the HMI application, strategies are developed more or less consciously (Lotze 2016, 334–336). Greetings and farewells follow transferred protocols from HHC, while repair strategies for disruptions are mostly consciously chosen. Therefore, a continuum between "awareness" strategies and "mindless behavior" (stereotypes, assumptions, cf. among others Reeves and Nass 1996 alignment, among others Pickering and Garrod 2004) should be assumed (Lotze 2016, 334–35).

Dimension 3: User language as a continuum of AAS (Alignment, Acceptance, Simplification)

The third dimension represents a continuum of degrees of simplification in user language – from pre-conscious alignment through passive and reactive behaviors to different simplification strategies (from RDS to CT). The starting point of this dimension is a user language that exactly matches the HHC and should, therefore, be interpreted radically according to CASA/MASA. This language is for us purely a hypothetical placeholder in the model, for which we (yet!) have no evidence. Innovative systems of the future may one day fill this space (or may not).

Alignment can be understood primarily as "mindless behavior" and manifests itself in persistences (repetitions), which I therefore include in the first section of the language continuum (see Chapter 2.1.1). Partial transfers from the HHC are naturally present in our studies, and we must consider them directly after alignment on the scale: turn construction and allocation, linguistic anthropomorphisms (e.g., through "you"/"they" pronouns, linguistic politeness, ritualized greetings, semantic-thematic anthropomorphisms (e.g., personal questions), etc.).

Acceptance phenomena and reactive behaviors, which are affordance-dependent, constitute the transitional area (see Chapter 2.1.2), such as "natural-language 'enter'-key") as a user reaction to plan-based systems.

Then come simplifications ("Machine Directed Speech" (MDS), "Simplified Registers" (SR)), initially those directly triggered by the affordances of the respective technology, and then those that are independent of them (see Chapter 2.1.3). The outer extreme pole in the continuum represents affordance-independent simplifications that occur across technologies (isolated keywords, abrupt terminations, and vulgar language as tests or after disruptions, etc.). These come closest to CT according to Zoeppritz (1985) and metaphorical language use (Krause and Hitzenberger 1992) and seem to have emerged as a new digital practice per se. Here, the tool character of the application alone seems to be the cognitive guiding concept.

External influencing factors:

In my view, user language evolves in dialogue not simply between the reception of an anthropomorphic technology and the level of consciousness in its cognitive processing but might also actually be modeled as an independent dimension. We need to consider it as an only partially dependent variable. Language follows its own principles, which manifest in the formation of style and register over extended periods. New technologies with new affordances give rise to new linguistic practices that should not be solely interpreted as technology deterministic. Besides the technological realm, there exist a social and language inherent realm. Language variation and change always occur in the interplay between explicitness and simplifications. Grammar and lexicon of each individual language also play a role. Lexicalization and grammaticalization and language- and culture-specific parameters for variation and change must be considered in a modern model for HMI. Otherwise, one cannot explain technology-transcendent new socio-linguistic practices (e.g., isolated keywords as a user strategy in all bots, from old, rule-based systems to Voice User Interfaces and ChatGPT). Therefore, the model also takes into account socio-linguistic factors such as language variation and change, culture, and society alongside technological affordances like dialogue design, data basis, and language model. This makes the multi-dimensional AAS-model compatible with more abstract, humanities-oriented discourses on AI.

References

Beißwenger, Michael, and Angelika Storrer. 2004. '*Dortmunder Chat-Korpus.' Chat-basierte Bibliotheksauskunft.* http://www.chatkorpus.tu-dortmund.de/files/releasehtml/index.html.

Branigan, Holly P., Michael J. Pickering, and Andrew A. Cleland. 2000. "Syntactic co-ordination in dialogue." *Cognition* 75 (2): 13–25. https://doi.org/10.1016/S0010-0277(99)00081-5.

Branigan, Holly P., and Jamie Pearson. 2006. "Alignment in human-computer interaction."In *Proceedings of the Workshop on How people talk to computers, robots, and other artificial communication partners*, edited by Kerstin Fischer, 140–56. Hanse-Wissenschaftskolleg Institute for Advanced Study.

Brown, Penelope, and Stephan C. Levinson. 1987. *Politeness. Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Clark, Herbert. H., and Kerstin Fischer. 2023. "Social robots as depictions of social agents." *Behavioral and Brain Sciences* 46: 1–33. https://doi.org/10.1017/S0140525X22000668.

Dippold, Doris. 2023. "'Can I have the scan on Tuesday?' User repair in interaction with a task-oriented chatbot and the question of communication skills for AI." *Journal of pragmatics* 204: 21–32.

Dürscheid, Christa. 2003. "Medienkommunikation im Kontinuum von Mündlichkeit und Schriftlichkeit: theoretische und empirische Probleme." *Zeitschrift für angewandte Linguistik* 38: 37–56.

Ebbinghaus, Herman. 1985. *Über das Gedächtnis. Untersuchungen zur experimentellen Psychologie*. Leipzig: Duncker and Humblot.

Fillmore, Charles J. 1976. "Frame semantics and the nature of Language." In *Annals of the New York Academy of Sciences. Origins and Evolution of Language and Speech*, 20–32.

Fischer, Kerstin. 2006. *What computer talk is and isn't. Human-Computer Conversation as Intercultural Communication*. Saarbrücken: AQ-Verlag.

Fischer, Kerstin. 2011. "How people talk with robots: Designing dialog to reduce user uncertainty." *Ai Magazine* 32 (4): 31–38.

Gandolfi, Greta, Michael J. Pickering, and Simon Garrod. 2023. "Mechanisms of alignment: shared control, social cognition and metacognition." *Philosophical Transactions of the Royal Society B* 378 (1870).

Greilich, Anna. In preparation. *Information structure in human-computer interaction: The linguisticanalysis of multi-turn information search with Amazon Alexa*. Dissertation thesis.

Habermas, Jürgen. 1993. *Theorie des kommunikativen Handelns*. Frankfurt am Main: Suhrkamp.

Habscheid, Stephan, Christine Hrncal, Jens Lüssem, Rainer Wieching, Felix Carros, and Volker Wulf. 2018. "Robotics and Emotion." In *EuropeNow* 19. *A Special Feature on Anxiety Culture*, edited by Nicole Shea and Emmanuel Kattan. https://www.europenowjournal.org/2018/07/01/robotics-and-emotion.

Hartsuiker, Robert J., and Casper Westenberg. 2000. "Word order priming in written and spoken sentence production." *Cognition* 75 (2): 27–39.

Hauser, Arnold. 1958. *The Philosophy of Art History. Routledge Revivals*. London: Routledge.

Heyselaar, Evelien. 2017. *Influences on the magnitude of syntactic priming*. Nijmegen: Radboud University Nijmegen.

Huiyang, Shen, and Wang Min. 2022. "Improving Interaction Experience through Lexical Convergence: The Prosocial Effect of Lexical Alignment in Human-Human and Human-Computer Interactions." *International Journal of Human-Computer Interaction* 38 (1): 28–41. https://doi.org/10.1080/10447318.2021.1921367.

Koch, Peter, and Wulf Oesterreicher. 1994. "Schrift und Schriftlichkeit." In *Schrift und Schriftlichkeit: Ein interdisziplinäres Handbuch internationaler Forschung*, edited by Herbert Günther and Otto Ludwig, 587–603.

Köhler, Reinhard. 2005. "Synergetic Linguistics." In *Quantitative Linguistik – Quantitative Linguistics. Ein internationales Handbuch*, edited by Reinhard Köhler, G. Altmann, and R. G. Piotrowski, 760–74. Berlin and New York: De Gruyter.

König, Katharina. 2019. "Sequential patterns in SMS and WhatsApp dialogues: Practices for coordinating actions and managing topics." *Discourse and Communication* 13 (6): 612–29.

Krämer, Nicole. 2008. *Soziale Wirkungen virtueller Helfer: Gestaltung und Evaluation von Mensch-Computer-Interaktion*. Stuttgart: Kohlhammer.

Krummheuer, Antonia L. 2010. *Interaktion mit virtuellen Agenten. Zur Aneignung eines ungewohnten Artefakts*. Stuttgart: Lucius & Lucius.

Krause, Jürgen, and Ludwig Hitzenberger. 1992. *Computer Talk*. Hildesheim, Zürich, and New York: Olms.

Lenz, Gaby, Jens Lüssem, Hannes Eilers, and Hannah Wachter. 2019. "Soziale Robotik in der Altenpflege." *Soziale Arbeit* 68 (11): 402–09.

Linnemann, Gesa Alena, and Regina Jucks. 2018. "'Can I Trust the Spoken Dialogue System Because It Uses the Same Words as I Do?' – Influence of Lexically Aligned Spoken Dialogue Systems on Trustworthiness and User Satisfaction." *Interacting with Computers* 30 (3): 173–86. https://doi.org/10.1093/iwc/iwy005.

Lombard, Matthew, and Kun Xu. 2021. "Social responses to media technologies in the 21st century: The media are social actors paradigm." *Human-Machine Communication* 2: 29–55.

Lotze, Netaya. 2016. *Chatbots: Eine linguistische Analyse*. Frankfurt: Peter Lang.

Lotze, Netaya. 2018. "Zur sprachlichen Interaktion mit Chatbots – Eine linguistische Perspektive." In *Talk with the Bots – Gesprächsroboter und Social Bots im Diskurs*, edited by Theo Hug and Günther Pallaver, 9–50. Innsbruck: University Press.

Lotze, Netaya. 2019. "Psycholinguistik der KI-Forschung – Beeinflussen Künstliche Intelligenzen unsere Art zu kommunizieren?" *Psychologische Aspekte von Automation und Robotik. Psychologie in Österreich* 4: 310–16.

Lotze, Netaya. 2022. "Zur Adressierung des Unbelebten – Grenzen von pragmatischer Konzeption." In *Mensch – Tier – Maschine: Sprachliche Praktiken an und jenseits der Außengrenze des Humanen*, edited by Miriam Schmidt-Jüngst, 305–327. Bielefeld: transcript, 305–27.

Lotze, Netaya, and Gabriel Aydin. In preparation. *Chat-GPT: Kohärente Interaktion?* Concept paper.

Lotze, Netaya, and Laura Ohrndorf. In preparation. *Socialbots. Eine linguistische Analyse*. Lang.

Nass, Clifford, and Scott Brave. 2005. *Wired for speech: how voice activates and advances the human-computer relationship*. MIT Press.

Nass, Clifford, and Young Me Moon. 2000. "Machines and mindlessness: Social responses to computers." *Journal of Social Issues* 56 (1): 81–103. http://stanford.edu/~nass/vita.html.

Natale, Simone, and Henry Cooke. 2021. "Browsing with Alexa: Interrogating the impact of voice assistants as web interfaces." *Media, Culture & Society* 43 (6): 1000–1016. https://doi.org/10.1177/0163443720983295.

Pearson, Jamie, Holly M. Branigan, Martin J. Pickering, J. Hu, and Clifford Nass. 2006. "Adaptive Language Behavior in HCI: How Expectations and Beliefs About a System Affect Users' Word Choice." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1177–80. https://doi.org/10.1145/1124772.1124948.

Pickering, Martin J. and Simon Garrod. 2004. "Towards a mechanistic psychology of dialogue." *Behavioural and Brain Sciences* 27: 169–225. https://doi.org/10.1017/S0140525X04000056.

Raveh, Eran, Ingo Siegert, Ingmar Steiner, Iona Gessinger, and Bernd Möbius. 2019. "Three's a Crowd? Effects of a Second Human on Vocal Accommodation with a Voice Assistant." In *Proceedings of the 20th Annual Conference of the International Speech Communication Association*, 4005–4009. https://doi.org/10.21437/Interspeech.2019-1825.

Reeves, Byron, and Clifford Nass. 1996. *The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places*. Cambridge University Press.

Reeves, Byron, Jeff Hancock, and S. X. Liu. 2020. "Social robots are like real people: First impressions, attributes, and stereotyping of social robots." *Technology, Mind, and Behaviour* 1 (1): 1–37. https://doi.org/10.1037/tmb0000018.

Ronneberger-Sibold, Elke. 1980. *Sprachverwendung – Sprachsystem. Ökonomie und Wandel.* Tübingen: Niemeyer. https://doi.org/10.1515/9783111346076.

Ruijten, Peter A., Diane H. L. Bouten, Dana C. J. Roushop, Jaap Ham, and Cees J. H. Midden. 2014. "Introducing a rasch-type anthropomorphism scale." In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 280–81. https://doi.org/10.1145/2559636.2559825.

Ruijten, Peter A., Antal Haans, Jaap Ham, and Cees J. Midden. 2019. "Perceived human-likeness of social robots: testing the Rasch model as a method for measuring anthropomorphism." *International Journal of Social Robotics* 11, 477–94. 10.1007/s12369-019-00516-z.

Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1992. *Lectures on Conversation, Volumes I and II.* Blackwell: Oxford.

Schmitz, Ulrich. 2015. *Einführung in die Medienlinguistik*. Darmstadt: WBG.

Siever, Torsten. 2011. *Texte i. d. Enge. Sprachökonomische Reduktion in stark raumbegrenzten Textsorten*. Frankfurt am Main: Peter Lang.

Wirtz, Markus A. 2021. *Dorsch – Lexikon der Psychologie*. Göttingen: Hogrefe.

Zoeppritz, Magdalena. 1985. *Computer talk? Technical Report TN 85.05*. Heidelberg: IBM Heidelberg Scientific Center.