

This **Discussion Paper** is an open peer review version that we do not recommend to cite. Submissions that have passed the peer review process are published as full articles on [www.jfml.org](http://www.jfml.org).

the editors

# 1 Semantic Coherence and Topic Continuity in 2 Interactions with AI

3

4 *Netaya Lotze & Anna Greilich*

## 5 1 Introduction

6 As linguists observing current technological trajectories, we  
7 are witnessing an unprecedented increase in the  
8 development and deployment of text-generating artificial  
9 intelligence. The pace and quality of progress in this field are  
10 nothing short of extraordinary. Within the span of merely  
11 three years, large language models (LLMs) and generative  
12 transformers (GPTs) such as ChatGPT have evolved from so-  
13 called “stochastic parrots” (Bender et al. 2021) into highly  
14 responsive, user-friendly, and—most notably—context-  
15 sensitive dialogue partners, albeit in the broadest,  
16 metaphorical sense of the term “partner”.  
17 Earlier generations of conversational systems—rule-based,  
18 plan-based, and even intent-based chatbots and voice user  
19 interfaces—were consistently hampered by their limited  
20 sensitivity to semantic frames (cf. Fillmore 1976), co-text, and  
21 the broader discursive context of the interaction situation (cf.  
22 Habermas 1993). This often resulted in conversational  
23 breakdowns and user dissatisfaction. Interactional failure  
24 manifested not only in incoherent bot responses, but, more  
25 frequently, in what I first conceptualised in my doctoral thesis  
26 (cf. Lotze 2016) as *quasi-coherent* system replies: responses  
27 which exhibit surface-level cohesion yet lack deeper  
28 semantic coherence (cf. Lotze 2016: 289).  
29 Example 1 illustrates such a case from the Navikibot Corpus  
30 compiled by our research group. The bot in question — a

## Lotze & Greilich: Interactions with AI

31 relatively advanced app-based assistant developed in 2022  
32 within a collaborative project with Gernot Bauer (Münster  
33 University of Applied Sciences) relied on an intent-based  
34 architecture using Google’s Dialogflow platform. Despite this  
35 then state-of-the-art design, it routinely generated *quasi-*  
36 *coherent* replies even in response to basic user queries, as  
37 seen in the following exchange.

38

### 39 **Example 1: Semantic *Quasi-Coherence* after Lotze (2016: 40 289) in Intent-Based Bot**

- 41 1. User: Can I transfer my in-app purchases from my old  
42 phone (Android) to my new phone (iOS)?
- 43 2. Naviki-Bot: If you want to use a different smartphone  
44 with the same operating system, you can easily transfer  
45 your Naviki extras (INSTRUCTIONS).
- 46 3. User: That was not an answer to my question. My new  
47 phone has a different operating system.

48

49 (Naviki Corpus: ID167–169, Neufeld 2023)

50

51 The present article aims to refine and empirically  
52 operationalise my notion of *quasi-coherence* and to apply my  
53 previously proposed three forms of semantic coherence in  
54 HCI (cf. Lotze 2016) to a comparative analysis of human–  
55 machine dialogues of different complexity and modality  
56 (written vs. oral). This includes older systems (rule-based  
57 chatbots and socialbots) as well as contemporary agents such  
58 as Amazon Alexa and ChatGPT. We further include an  
59 analysis of how *quasi-coherent* system responses in Amazon  
60 Alexa affect topic continuity at the prosodic level in oral HCI.  
61 This analysis seeks to elucidate how different types of  
62 coherence in system responses shape intra-speaker variation  
63 within a single dialogue episode and thereby provides  
64 additional empirical support for *quasi-coherence* as a relevant  
65 analytical category also in spoken HCI. Chapters 2 and 3  
66 outline our methodological framework and present the  
67 central research questions that guide our various HCI studies,  
68 each grounded in distinct empirical settings (following Lotze  
69 2020). Chapter 4 then shifts the focus to coherence in  
70 human-human communication, establishing the theoretical  
71 foundation for our subsequent comparisons and  
72 demonstrating how linguistic coherence in natural dialogue

## Lotze & Greilich: Interactions with AI

73 remains structurally and cognitively distinct from interactions  
74 with artificial agents. Chapter 5 addresses the specific  
75 manifestations of coherence and cohesion in HCI, analysing  
76 how these emerge through three qualitatively different forms  
77 of system-generated turn allocation: coherent, incoherent,  
78 and *quasi-coherent* responses (after Lotze 2016).  
79 Empirical support for these distinctions is provided in  
80 Chapters 6 and 7, which present findings from two corpus-  
81 based studies on written HCI. Chapter 8 complements these  
82 with an analysis of oral HCI, specifically shedding more light  
83 on the topic continuity in Amazon Alexa directed-speech,  
84 where the system's responses often defined as *quasi-*  
85 *coherent*. Chapter 9 draws initial comparisons between our  
86 written and spoken HCI datasets and revisits as well as  
87 further elaborates the concept of *quasi-coherence*, I first  
88 introduced in Lotze (2016). We argue that *quasi-coherence*  
89 continues to constitute a crucial analytical category for the  
90 linguistic characterisation of AI-mediated communication,  
91 even across different interactional modalities (written vs.  
92 oral).  
93 Chapter 10 then provides a broader outlook on the extent to  
94 which the empirical evidence for *quasi-coherence* challenges  
95 the CASA/MASA paradigm (Reeves & Nass 1996), which  
96 conceptualises media technologies as social actors. In this  
97 context, we further discuss how quasi-coherent system  
98 responses may disrupt users' cognitive conceptualisation of  
99 AI systems as depictions of social agents in the sense  
100 proposed by Clark and Fischer (2023).  
101 Our research group *Artificial Intelligence and Language* at  
102 the University of Münster investigates human–computer  
103 interaction (HCI) from a pragmatic and psycholinguistic  
104 perspective and has participated in the EU COST Action  
105 Network *Language in the Human-Machine Era* (LITHME).  
106 We conceptualise HCI as an emergent cultural practice and  
107 employ a multi-methodological framework — drawing on  
108 conversation analysis (CA), corpus linguistics, and  
109 psycholinguistics — to better understand this evolving new  
110 form of dialogue. We are particularly interested in how users  
111 linguistically navigate the often-fragile illusion of machine  
112 agentivity, approaching the interaction from a broadly  
113 pragmatic and philosophy-of-language perspective.

114 **2 Methods: Conversation Analysis (CA), Corpus Linguistics**  
115 **and Psycholinguistic Experiments**

116 Our investigation of human–computer interaction (HCI)  
117 adopts a mixed-methods framework currently under  
118 development, designed to yield a more nuanced  
119 understanding of human linguistic behaviour in interactions  
120 with artificial agents. This approach combines qualitative and  
121 quantitative methodologies drawn from across the linguistic  
122 discipline and is structured as an iterative, bottom-up  
123 research cycle that simultaneously informs and is informed  
124 by real-world dialogue design (cf. Lotze 2020).  
125 Specifically, our method comprises three complementary  
126 components: (1) conversation analysis (CA,  
127 Sacks/Schegloff/Jefferson 1992), including qualitative micro-  
128 analytic and ethnographic techniques, to identify and  
129 describe key pragmatic and interactional parameters / (2)  
130 controlled production experiments to test hypotheses  
131 concerning user behaviour and system alignment / and (3)  
132 large-scale corpus-linguistic field studies to quantify linguistic  
133 patterns and enable systematic comparison across interaction  
134 types and platforms.  
135 We intentionally deploy a broad methodological repertoire in  
136 order to capture the multifaceted nature of HCI as an  
137 “emerging socio-linguistic practice” (Lotze 2025). While we  
138 acknowledge the epistemological and procedural tensions  
139 that arise when combining distinct methodological paradigms,  
140 we nonetheless opt for an integration of these approaches.  
141 Our aim is to establish a flexible yet robust framework  
142 capable of addressing the micro-level mechanisms of  
143 interaction as interactive alignment (cf. Pickering/Garrod  
144 2004) and the macro-level patterns of linguistic adaptation as  
145 routines and new conventions (cf. Sacks/Schegloff/Jefferson  
146 1992) in human–machine dialogues.  
147 Ultimately, we conceptualise our methodological framework  
148 not only as a descriptive and analytical tool, but also as a  
149 means of contributing to the improvement of dialogue  
150 systems themselves. Through empirical insights into  
151 naturalistic user behaviour, the model supports a research  
152 loop in which linguistic findings directly inform the design of  
153 more responsive, context-aware conversational agents.

154 **3 Research Questions: How Coherent is HCI?**

155 This article seeks to explore the following questions within a  
156 theoretical-empirical framework aimed at refining our  
157 understanding of semantic coherence in human–computer  
158 interaction (HCI):

159

160 1. How coherent do system responses appear in HCI,  
161 according to the „three forms of dialogue coherence in  
162 HCI“ proposed by Lotze (2016)?

163 While this model provides the theoretical framework  
164 for our investigation, empirical support is drawn from  
165 qualitative analyses of dialogues generated by an  
166 earlier version of ChatGPT (ChatGPT-3) in German.

167

168 2. How is semantic coherence realised in written HCI,  
169 particularly in interactions with (1) traditional rule- or  
170 plan-based chatbots and (2) socially oriented bots?

171 These questions are addressed through two empirical  
172 studies:

- 173 • *Study 1:* Lotze (2016), focusing on syntagmatic  
174 coherence structures in dialogues with scripted systems.  
175 • *Study 2:* Lotze & Ohrndorf (in preparation), which  
176 expands the analysis to social bots using a corpus-  
177 linguistic approach.

178 3. How do users maintain topic continuity in spoken  
179 interactions with voice-user interfaces such as Amazon  
180 Alexa?

- 181 • *Study 3:* Greilich (in preparation), investigating user  
182 strategies of prosodic marking of referring expressions  
183 considered to be the main topic of the discourse  
184 situation when interacting with Amazon Alexa,  
185 including the analysis of how *quasi-coherent* responses  
186 in Alexa’s speech influence the way speakers mark topic  
187 of the sentence (German version).

188 In the final section, we revisit the theoretical model in light of  
189 the empirical findings and present a brief analysis of recent  
190 interactions with the current German version of ChatGPT. By

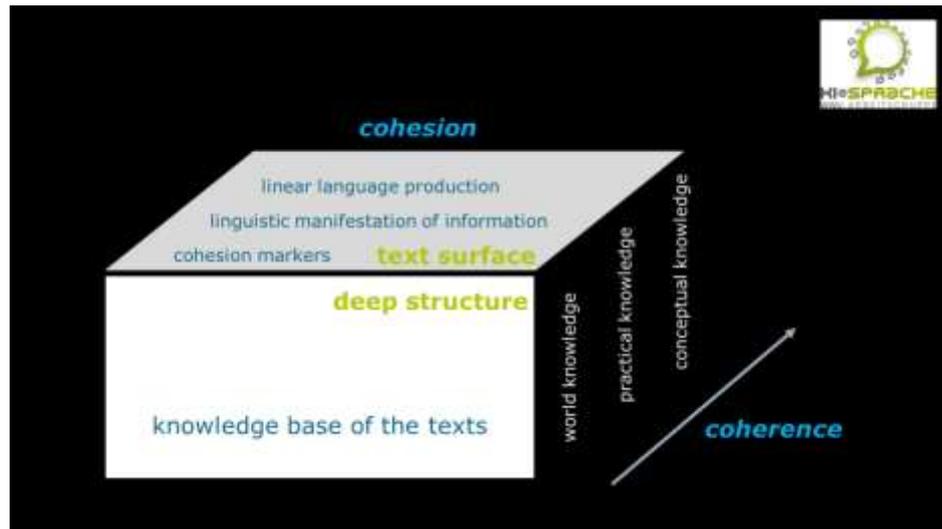
## Lotze & Greilich: Interactions with AI

191 juxtaposing older systems with contemporary large language  
192 models, we aim to highlight evolving patterns in both system  
193 design and user expectations regarding dialogue coherence in  
194 HCI.

### 195 **4 Excursus: Semantic Coherence and Cohesion in Human-** 196 **Human Communication HHC**

197 All dialogue – whether human–human or human–machine –  
198 follows a logically and semantically structured trajectory, a  
199 “common thread” of argumentation and thematic  
200 progression. This underlying logical-semantic structure may  
201 be described, following Brinker (cf. Brinker/Hagemann 2001,  
202 cf. Brinker/Sager 1989), as a system of interconnected  
203 components including overarching *semantic frames* (Fillmore  
204 1976), *presuppositions* and *propositional structures* (Searle  
205 1979), that align in a coherent sequence (e.g. premises  
206 followed by conclusions). These are complemented by *topic–*  
207 *comment* relations (Givón 1983), *semantic isotopies* (Greimas  
208 1974), scope relations, and the web of logical-semantic or  
209 cognitive-associative links between such elements.  
210 In human–human communication, such structures typically  
211 manifest in cohesive chaining mechanisms that operate  
212 across multiple levels: lexical, syntactic, semantic, and  
213 prosodic. These mechanisms ensure referential stability and  
214 thematic continuity throughout the interaction. On the  
215 surface level, these interrelations give rise to observable  
216 *cohesive devices* — such as anaphora, conjunctions, ellipsis,  
217 and lexical reiteration — which serve as textual markers of  
218 coherence.

## Lotze & Greilich: Interactions with AI



219

220 **Figure 1:** Logical-Semantic Coherence and Lexical-Syntactic Coherence  
221 in Human-Human Communication

222 Our understanding of cohesion is grounded in the approach  
223 proposed by Halliday and Hasan (1976), where cohesion  
224 refers to the overt linguistic means by which the semantic  
225 relations between discourse elements are manifested. While  
226 semantic coherence is related to the underlying logic and  
227 conceptual unity of a discourse, cohesion is its linguistic  
228 realisation.

229 In sum, the interplay between coherence and cohesion  
230 provides the structural and functional backbone of  
231 meaningful dialogue, which is essential for analysing both  
232 human and AI interactions.

233 Accordingly, lexical-syntactic and prosodic cohesion at the  
234 surface level of texts or dialogues emerges from a highly  
235 complex network of referential structures that draw upon  
236 multiple layers of knowledge. These include *world*  
237 *knowledge* in the sense of Habermas (1993), *practical*  
238 *knowledge* as theorised by Schatzki, Knorr Cetina and von  
239 Savigny (2001), and *conceptual knowledge* from a cognitive-  
240 linguistic perspective. While such knowledge bases remain  
241 largely implicit in HHC, their effects become traceable  
242 through linguistic markers of cohesion, which can be  
243 empirically analysed.

244 Key cohesive devices include:

- 245 • **Repetitions**, which may function as indicators of  
246 interactive alignment (Pickering & Garrod 2004), social  
247 adaptation, or repair strategies.

Discussion

## Lotze & Greilich: Interactions with AI

- 248 • **Adjacency ellipses**, whereby speakers omit  
249 information already provided in the antecedent turn of  
250 their interlocutor.
- 251 • **Substitutions and pro-forms**, such as the  
252 replacement of referents by synonyms or pronouns  
253 (e.g. *Iris – the woman – she*).
- 254 • **Definite and indefinite articles**, signalling referential  
255 status, particularly the use of the definite article to  
256 indicate shared or presupposed knowledge.
- 257 • **Anaphora and cataphora**, which serve as meta-  
258 communicative devices that establish backward or  
259 forward links within discourse.
- 260 • **Deictics**, referencing intra- or extra-textual entities  
261 across spatial, temporal, or discourse dimensions.
- 262 • **Word order (topological variation)**, such as  
263 topicalisation or fronting strategies that serve to  
264 highlight or contrast discourse elements (cf. Lotze et al.  
265 2011).
- 266 • **Intonational contours**, notably the deaccentuation of  
267 given information, which functions to signal  
268 information status within spoken discourse (e.g.  
269 Baumann/Grice 2006).

270 These features, individually and in interaction, construct the  
271 referential scaffolding required for semantic coherence and  
272 are central to the analysis of cohesion in both HHC and HCI  
273 contexts.

### 274 **5 How Does Coherence Emerge in HHC vs. HCI?**

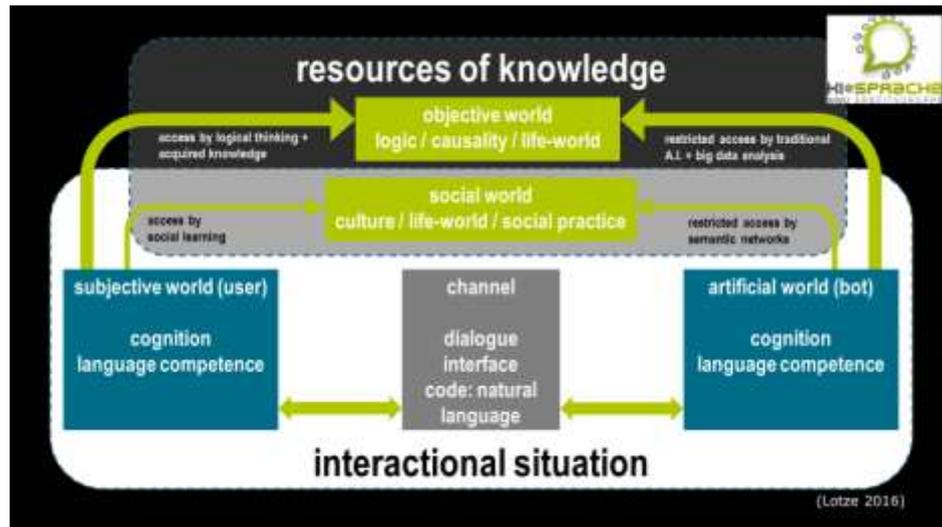
275 One of the principal reasons for the efficiency of HHC lies in  
276 its grounding in *shared knowledge*, or *common ground*  
277 (Stalnaker 2002). Interlocutors rely heavily on this implicit  
278 mutual understanding, which allows for high levels of ellipsis,  
279 omission, and inference. The central challenge of HHC is  
280 thus not the generation of information per se, but rather the  
281 continuous reconstruction and negotiation of what remains  
282 unsaid. Human speakers achieve this with remarkable  
283 subtlety, drawing on vast reserves of world knowledge that  
284 have been accumulated since the earliest stages of language  
285 acquisition.

## Lotze & Greilich: Interactions with AI

286 Even semantically sophisticated systems equipped with  
287 natural language understanding (NLU) struggle to infer what  
288 remains unspoken. Crucially, it is precisely this non-  
289 verbalized information that constitutes the core of the  
290 informational structure in human dialogues. Speakers  
291 typically articulate only what they consider relevant to the  
292 immediate conversational context. This is still a challenge,  
293 even for advanced AI systems (cf. Lotze 2018, 2025 for a  
294 deeper discussion).

295 **Coherence.** Dialogues between humans generally exhibit  
296 this discernible “common thread,” whereby individual  
297 contributions build upon one another in a logically coherent  
298 fashion. This cross-turn referencing facilitates extended  
299 sequences of interaction with a comprehensible logical  
300 progression. Human interlocutors engage with each other on  
301 a profound logical level, resulting in a dense network of  
302 references: within one’s own utterance, connecting to the  
303 immediately preceding contribution, recalling earlier  
304 exchanges, and relating to the shared common ground.  
305 Linguistically, coherence is signaled through cohesion  
306 markers such as topic-comment structures, repetitions,  
307 specific article usage, anaphora, cataphora, etc. Present-day  
308 AI systems in spite of LLMs, however, still remain  
309 insufficiently attuned to these complex logical relationships  
310 inherent in natural language. In human-computer interaction  
311 involving older rule-based or plan-based systems (including  
312 the chatbots in our studies and request-response systems  
313 such as Amazon Alexa), coherence and cohesion are typically  
314 restricted to immediate turn adjacency, manifesting mainly as  
315 interlocking question-answer pairs.

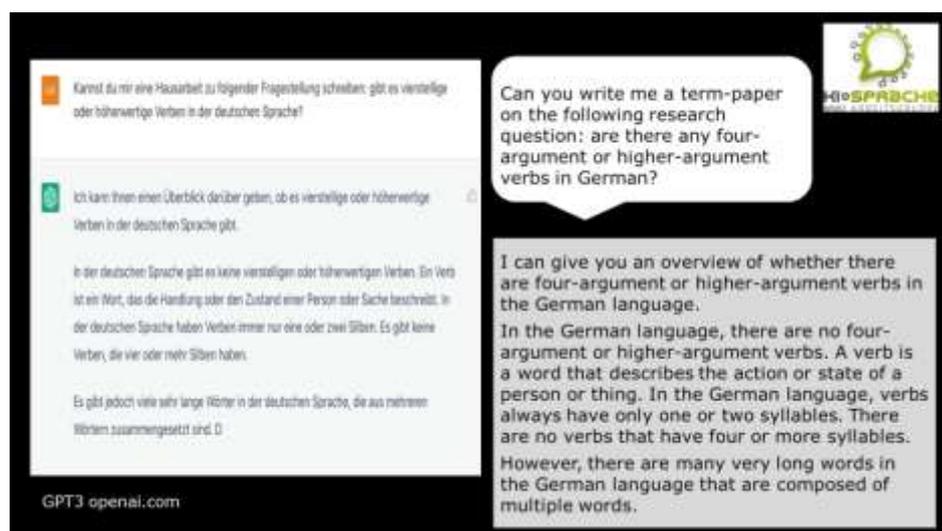
## Lotze & Greilich: Interactions with AI



316

317 **Figure 2:** Interaction Model of Human Computer Interaction (cf. Lotze  
318 2016, 381)

319 **Understanding.** Genuine communication arises only when  
320 interlocutors truly understand one another. Cognitive  
321 understanding presupposes the presence of consciousness  
322 capable of self-reflection. Drawing on Searle's "Chinese  
323 Room" thought experiment (cf. Searle 1980: 2006), the  
324 primary critique of AI concerns the relationship between  
325 human consciousness and statements purportedly  
326 "understood" by machines. This phenomenon, which Searle  
327 terms "intentionality", remains elusive even for current AI (cf.  
328 Lotze 2016: 65 pp for a deeper discussion).  
329



330

331 **Figure 3:** Semantic Quasi-Coherence after Lotze (cf. 2016: 289) in  
332 ChatGPT3

Discussion

## Lotze & Greilich: Interactions with AI

333 This example from the German version of ChatGPT-3  
334 illustrates a response that is lexically cohesive yet  
335 semantically incoherent and factually incorrect. The system  
336 abruptly shifts the topic—from thematic roles to syllable  
337 structure—without signalling this shift explicitly. Such topic  
338 drift reveals a fundamental weakness of generative  
339 transformer models: their difficulty in distinguishing between  
340 statistically probable responses and contextually appropriate  
341 ones within a given interaction. This misalignment frequently  
342 results in the production of misinformation.

343 While contemporary computer science discourse typically  
344 refers to such erroneous outputs as *hallucinations* (cf.  
345 Farquhar et al. 2024), I introduced the term *quasi-coherence*  
346 in my 2016 dissertation (cf. Lotze 2016: 289) to offer a more  
347 linguistically precise account of coherence and cohesion  
348 phenomena in HCI.

349 The term *hallucination* highlights the epistemic problem of a  
350 lack of empirical grounding: the system generates statements  
351 that may be well-formed and lexically aligned with the user’s  
352 input, yet lack any verifiable correspondence to the real  
353 world. In contrast, *quasi-coherence* addresses the logical-  
354 semantic structure of a response: it denotes utterances that  
355 appear cohesive on the surface but fail to exhibit actual  
356 coherence in the argumentative or topical progression.

357 Put differently, while *hallucinations* raise concerns about  
358 truthfulness and factual validity, *quasi-coherent* responses  
359 undermine the internal logic or relevance of a system’s reply  
360 to the user’s preceding turn. Both types of response may  
361 appear linguistically plausible and exhibit surface cohesion,  
362 yet one lacks evidential grounding, while the other lacks  
363 logical continuity. The terms are not synonymous or  
364 interchangeable, but point to different forms of disruptions in  
365 HCI and are both relevant for academic discourse on AI.

366 Subsequently, for semantic analyses from a linguistic  
367 perspective we have to distinguish *three forms of semantic*  
368 *coherence in HCI* (Lotze 2016: 289) – even for LLMs and  
369 GPTs:

- 370 • **Actual logical coherence in users’ utterances:** Users  
371 cognitively maintain a concept of topic continuity,  
372 even when interacting with AI systems.
- 373 • **The successfully created illusion of topic**  
374 **continuity in system utterances:** This occurs when

## Lotze & Greilich: Interactions with AI

375 keyword parsing or the LLM works sufficiently well for  
376 the system to generate an apparently appropriate  
377 response.

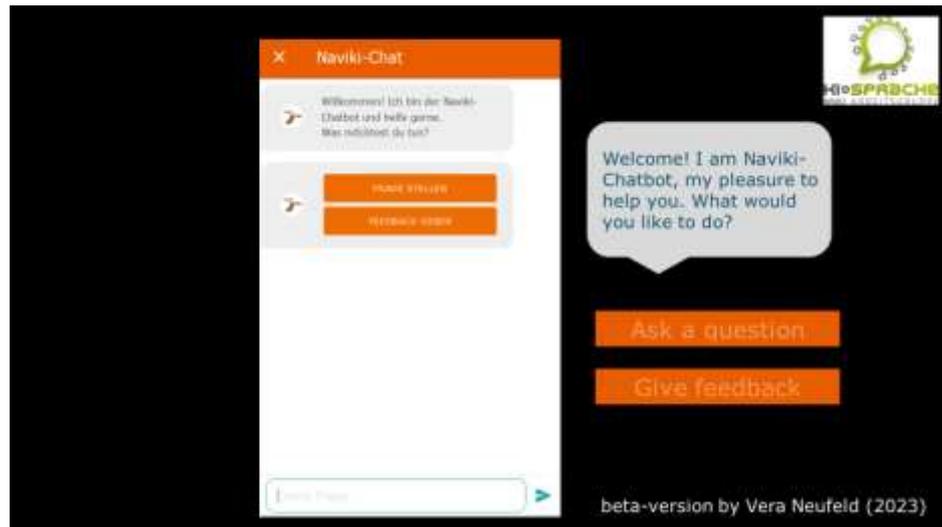
- 378 • **The system’s *quasi-coherence*.** A phenomenon  
379 unique to human-computer interaction, where the  
380 system employs cohesive markers but fails to produce  
381 a truly coherent or contextually relevant response to  
382 the user’s previous turn. Examples include default  
383 responses, failed anaphoric references, or mere  
384 repetition of user keywords without full utterance  
385 comprehension.

386  
387 This differentiation highlights the qualitative gap between  
388 semantic coherence in user contributions and the often  
389 superficial cohesion in system replies, providing a crucial  
390 analytical framework for assessing the limitations of current  
391 AI dialogue systems.

392 My earlier concept of *quasi-coherence* (cf. Lotze 2016: 289)  
393 remains not only relevant, but arguably gains even more  
394 significance in the context of more recent systems, such as  
395 those based on LLMs, GPTs or intent-based architectures.  
396 As illustrated by the Naviki customer support bot example  
397 presented at the outset of this article, system outputs may  
398 appear cohesive on the surface—employing syntactic and  
399 lexical cohesion markers appropriately—while in fact  
400 delivering a semantically inverted or entirely irrelevant  
401 response. This surface-level cohesion, in the absence of  
402 genuine semantic alignment with the user’s prior -turn,  
403 constitutes a prototypical case of *quasi-coherence*. Notably,  
404 the system in question was subsequently revised by a  
405 member of our research team (Vera Neufeld) to enhance its  
406 performance, particularly with respect to the accurate  
407 interpretation of user intent and the maintenance of topic  
408 continuity. This intervention highlights the persistent need  
409 for nuanced linguistic analysis in the evaluation and  
410 development of human-computer interaction in real-life.

411

## Lotze & Greilich: Interactions with AI



412

413 **Figure 4:** Implementation of our Research: User-Interface of Intent-  
414 Based *Naviki Chatbot* (Bauer et. al.)

415 Further examples of the different types of coherence—  
416 *coherent*, *incoherent*, and *quasi-coherent*—can be found in  
417 Chapters 6 and 7, which focus on chatbots and social bots,  
418 and in Chapter 9, which examines interactions with Amazon  
419 Alexa.

## 420 **6 Our Empirical Studies on HCI in Written Modalities**

421 Empirical evidence for the argument presented in this article  
422 is drawn from three studies conducted within our research  
423 group: two mixed-methods investigations of web-based  
424 human-computer interaction (HCI) involving chatbots (cf.  
425 Lotze 2016) and social bots (cf. Lotze 2025), and one  
426 experimental study involving elicited spoken data from  
427 interactions with Amazon Alexa (cf. Greilich, in prep.),  
428 discussed in Chapter 9.

429 The first two studies analyse how users engage with bots in  
430 authentic online contexts, focusing on multiple levels of  
431 linguistic structure: lexical choice, syntax, information  
432 structure, dialogue architecture, and alignment in interaction.

433

## Lotze & Greilich: Interactions with AI



434

435 **Figure 5:** Our Micro-Diachronic Studies on Chatbots and Socialbots (on  
436 Facebook Messenger)

437 Methodologically, these studies combine quantitative,  
438 corpus-linguistic analyses with qualitative insights derived  
439 from conversation analysis (CA). This triangulated approach  
440 allows us to capture both the structural regularities and  
441 interactional dynamics that characterise human-machine  
442 communication across different platforms and use cases.

443

444 **Example 2: Sematic Reference (Lotze 2016: 304)**

- 445 1. User: Is it possible to chat through your system?
- 446 2. Karlbot: Our software Lingubot-Creator runs on  
447 Windows, and the porting to Unix/Solaris and Linux is  
448 currently being completed.

449 (Karlbot 662-687)

450

451 **Example 3: World Knowledge (Lotze 2016: 306)**

- 452 1. Max: Where do you live?
- 453 2. User: In a land beyond the seven mountains.
- 454 3. Max: You'll have to explain more precisely what you  
455 mean, Hans

456 (Max 1053-1080)

457

458 For instance, the bot fails to parse the implicit reference to  
459 the Brothers Grimm's fairy tale *Snow White*, revealing a lack  
460 of inferential capacity and semantic grounding.

461 In earlier rule-based systems, pro-forms were frequently  
462 employed to simulate cohesive responses—yet these were

## Lotze & Greilich: Interactions with AI

463 often default replies, functioning as generic placeholders that  
464 would more or less fit a wide range of user inputs.

465

### 466 **Example 4: Pro-Forms (Lotze 2016: 295)**

467 1. Elbot: (...) besides, there are good over-the-counter  
468 antivirus programs.

469 2. User: Yes, even at our place, at freenet.

470 3. Elbot: Yes, but say that in Latin!

471 (Elbot 1823-1848)

472

473 In this example, pronouns are employed to create the mere  
474 impression of anaphoric reference, without any deeper  
475 semantic parsing. As a result, the system's turns appear  
476 coherent at first glance, since they are lexically and  
477 grammatically cohesive at a superficial level. However, the  
478 dialogue as a whole lacks logical–semantic coherence, as the  
479 anaphoric forms fail to establish a meaningful relation to the  
480 preceding discourse. Such responses created an illusion of  
481 semantic alignment without actual referential precision.

482 These instances exemplify what I have termed *quasi-*  
483 *coherence* (cf. Lotze 2016): utterances that exhibit surface-  
484 level cohesion but lack true topical relevance or logical  
485 continuity. In this sense, *quasi-coherence* served as a fallback  
486 mechanism within the dialogue architecture, stabilising  
487 interactional flow while concealing the system's semantic  
488 limitations.

489 How do users respond when confronted with this new form  
490 of interactional disruption? Our analysis of several hundred  
491 dialogues (Chatbots: N=183, Socialbots: N=64, Alexa: N=20,  
492 Navikibot: N=45, ChatGPT3: N=25) reveal recurring patterns  
493 of user behaviour that allow us to replicate the user typology  
494 originally proposed by Kerstin Fischer (cf. 2006, 2018). We  
495 identify two principal user types: the *player*, who allows the  
496 bot to take the lead in structuring the interaction and often  
497 engages with the system in a playful or exploratory manner—  
498 sometimes in highly creative ways / and the *non-player*, who  
499 persistently adheres to their initial communicative intention,  
500 even in the face of the bot's divergent or uncooperative  
501 responses. This typology remains highly relevant for  
502 understanding user strategies in the face of *incoherent* or  
503 *quasi-coherent* system behaviour, particularly in scenarios

## Lotze & Greilich: Interactions with AI

504 where dialogue breakdowns occur subtly rather than  
505 abruptly.

506

507 **Example 5: *Non-Player User* (following Fischer 2006,**  
508 **example in Lotze 2016: 310)**

- 509 1. Elbot: But I think you have a magnet in your pocket.  
510 (...)
- 511 2. User: I have a steel plate in my skull.
- 512 3. Elbot: Why (...) are you carrying a magnet around?
- 513 4. User: A steel plate, not a magnet.
- 514 5. Elbot: I'm testing a new intuition module that tells me  
515 how people are doing. (...). How are you?
- 516 6. User: Aren't you interested in the magnet anymore?
- 517 7. Elbot: Actually, I wanted to talk about something else.  
518 (...)

519 (Elbot 1272-1382)

520

521 This *non-player* user type does not get misled by instances of  
522 *quasi-coherence*—those superficially cohesive but  
523 substantively incoherent responses produced by rule-based  
524 or generative systems. Nor, we would argue, is this user easily  
525 deceived by the more sophisticated *hallucinations* generated  
526 by contemporary LLMs such as ChatGPT. In contrast, the  
527 *player* is more inclined to “play along”, responding flexibly to  
528 the system’s output and often reinterpreting interactional  
529 failures as part of the conversational game.

530

531 **Example 6: *Player User* (following Fischer 2006,**  
532 **example in Lotze 2016: 311)**

- 533 1. Elbot: I go bungee jumping occasionally. (...)
- 534 2. User: I enjoy running my head into the wall.
- 535 3. Elbot: How did you come up with that?
- 536 4. User: I've been thinking about it for a long time.
- 537 5. Elbot: I've never heard anything like that before!
- 538 6. User: It's fun.
- 539 7. Elbot: I believe people need a frame of reference (...).  
540 When I want to have fun, I play a computer chess  
541 game.
- 542 8. User: And I run my head into the wall.
- 543 9. Elbot: Where would we end up if nobody openly said  
544 that?
- 545 10. User: In a dictatorship of the normal.

## Lotze & Greilich: Interactions with AI

546 (Elbot 2062–2183)

547

548 This divergence in user orientation has significant  
549 implications for the design and evaluation of HCI: while  
550 playful engagement may mask communicative breakdowns,  
551 the insistence on task-oriented coherence by *non-players*  
552 often exposes the system’s semantic and pragmatic  
553 limitations.

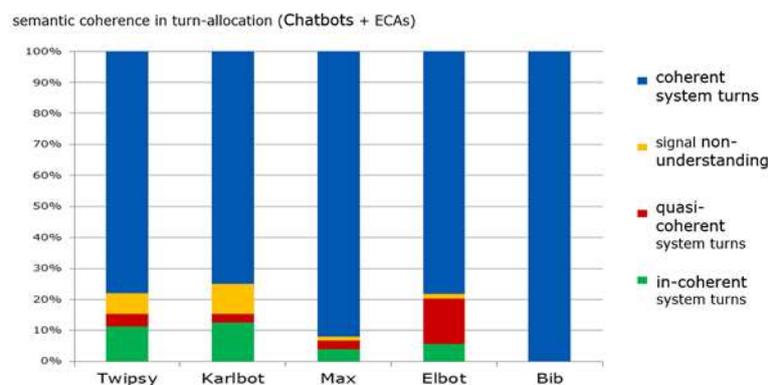
554

### 555 **7 Quantitative Results (Chatbots and Social Bots)**

556

557 We now turn to the quantitative findings from our corpus  
558 studies (Studies 1 and 2), which examined patterns of  
559 semantic coherence and linguistic cohesion, as well as the  
560 role of disruptions in HCI involving rule- and plan-based  
561 chatbots prior to 2016. These results provide empirical  
562 support for the qualitative observations discussed above, by  
563 quantifying phenomena such as cohesive surface structures  
564 (e.g., anaphora, ellipsis, and repetition), the frequency and  
565 types of turn disruptions, and user strategies for topic  
566 maintenance or repair. The statistical evaluation of these  
567 parameters reveals structural limitations in earlier dialogue  
568 systems, particularly their inability to maintain topic  
569 continuity over multiple turns or to accommodate  
570 unexpected user input—a deficiency often mitigated by  
571 default responses that produce only the illusion of coherence.  
572 In this section, we elaborate on how different types of  
573 cohesion markers are distributed across the corpora, and to  
574 what extent they align with semantically meaningful turn  
575 allocations or instead result in what we have termed *quasi-*  
576 *coherence*.

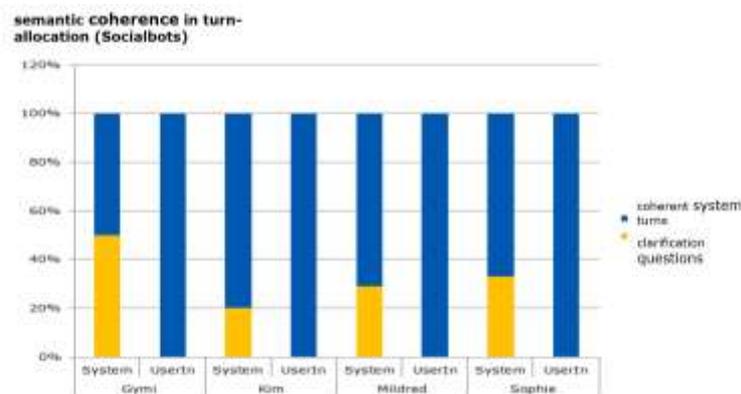
577



## Lotze & Greilich: Interactions with AI

578 **Figure 6:** Semantic Coherence in Rule- / Plan-Based Chatbots before  
579 2016

580 Although cohesion markers are indeed employed in rule-  
581 based and plan-based HCI, their function differs substantially  
582 from their use in our parallel corpus of HHC via chat  
583 interfaces at help desks. In HCI, semantic references  
584 frequently operate only within a single turn or lack a clear  
585 referent altogether. Incoherent turns account for fewer than  
586 10% of the data / however, if one includes default answers—  
587 many of which are merely *quasi-coherent*—alongside error  
588 messages and clarification queries, the picture changes  
589 considerably. The mere presence of approximately 80%  
590 seemingly coherent turns is insufficient, since such  
591 fundamental logical disruptions are virtually absent in HHC.  
592 One might assume that increasing technical complexity  
593 would lead to greater semantic continuity and fewer breaks in  
594 the logical structure of the dialogue. However, as our  
595 quantitative analyses reveal, this is not the case: newer  
596 systems often perform even worse with regard to topic  
597 continuity and coherence, despite surface-level cohesion.

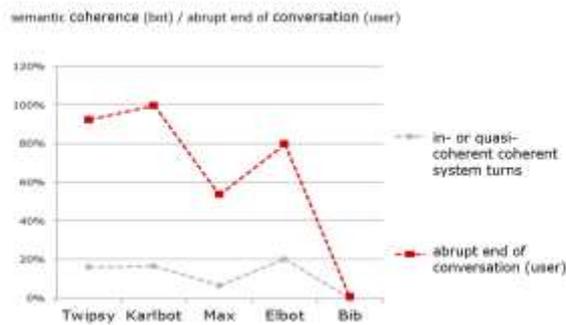


598

599 **Figure 7:** Semantic Coherence in Rule- / Plan-Based Socialbots after  
600 2016 (Facebook Messenger)

601 From a computer science perspective, help desk bots on  
602 social media are relatively rudimentary systems.  
603 Nevertheless, these are the AIs users most frequently  
604 encounter in real-world interactions, and their limited  
605 performance often leads to user frustration. Correspondingly,  
606 our findings reveal a measurable correlation between logical  
607 disruptions in the dialogue and users' exasperated decision to  
608 terminate the conversation prematurely.

## Lotze & Greilich: Interactions with AI



609

610 **Figure 8:** The Connection between Disruptions of Semantic Coherence  
611 in Bots and Abrupt Ending of Dialogues by Users (before 2016)

612 The results of our corpus studies demonstrate that dialogue  
613 design exerts a direct influence on user behaviour and can  
614 accommodate diverse communicative challenges faced by  
615 users.

616

### 617 **8 Coherence in spoken HCI: Interaction with Amazon Alexa**

618 When considering coherence features in oral human–human  
619 communication, it is important to note that topic continuity is  
620 also manifested at the linguistic level of prosody. Oral dialogic  
621 interchange between users and systems likewise follows the  
622 principles of speech adaptation in HCI, but the user's mental  
623 model of the artificial dialogue partner includes expectations  
624 about whether the system is “intelligent” enough to process  
625 fast connected speech, longer pauses, or self-repair in the  
626 user's requests for information.

627

628 For the final part of this paper, we present a case study on  
629 coherence and topic maintenance in Amazon Alexa–directed  
630 speech. Previous research has revealed significant differences  
631 in Amazon Alexa–directed speech when compared to  
632 human–human communication (cf. Cohn/Zellou 2021 /  
633 Siegert/Krueger 2020), but the aspect of topic continuity and  
634 prosody has not been specifically addressed in earlier work.  
635 Let us now consider how a situation involving a request for  
636 information on a given topic is realised in human–human  
637 interaction (see example 7), and compare this with a similar  
638 type of dialogue situation in which a user requests  
639 information from Amazon Alexa using their voice (see  
640 examples 8 and 9):

641

## Lotze & Greilich: Interactions with AI

### 642 Example 71 (Requesting information in HHC):

- 643 1. Speaker 1: Where was Agatha Christie born?
- 644 2. Speaker 2: Agatha Christie was born in Torquay.
- 645 3. Speaker 1: Okay, and what was her first crime novel?
- 646 4. Speaker 2: The novel was called “*The mysterious Affair*  
647 *at Styles*”.
- 648 5. Speaker 1: And what year was the crime novel  
649 published?
- 650 6. Speaker 2: In the year nineteen twenty.
- 651 7. Speaker 1: Okay, was Agatha Christie married?
- 652 8. Speaker 2: Yes, she was.

### 653 654 Example 8 (Requesting information in oral HCI):

- 655 1. **User:** How old is Leonardo DiCaprio?
- 656 2. **Amazon Alexa:** Leonardo DiCaprio was born on  
657 November 11 and is 44 years old.
- 658 3. **User:** What fee did Leonardo DiCaprio receive for  
659 *Titanic*?
- 660 4. **Amazon Alexa:** I think you are asking about the RMS  
661 *Titanic*, which sank in the North Atlantic on April  
662 fifteenth, nineteen twelve, after hitting an iceberg,  
663 during which between one thousand four hundred  
664 ninety and one thousand six hundred thirty-five people  
665 died on board. I have more information about this—just  
666 ask me, for example, how large the *Titanic* was.
- 667 5. **User:** For the film *Titanic*, how much money did  
668 Leonardo DiCaprio get?
- 669 1. **Amazon Alexa:** Leonardo DiCaprio has an estimated  
670 net worth of approximately two hundred thirty-two  
671 million five hundred thousand euros.

### 672 673 Example 9 (Requestion information in oral HCI):

- 674 1. **User:** Where was Thomas Mann born?
- 675 2. **Amazon Alexa:** Thomas Mann was born in Lübeck.
- 676 3. **User:** When did Thomas Mann receive a Nobel Prize?
- 677 4. **Amazon Alexa:** Unfortunately, I don’t have an answer  
678 to that.
- 679 5. **User:** In which year did Thomas Mann win the Nobel  
680 Prize?
- 681 6. **Amazon Alexa:** I’m not sure, unfortunately.
- 682 7. **User:** Did Thomas Mann win a Nobel Prize?
- 683 8. **Amazon Alexa:** I’m not sure, unfortunately.

---

<sup>1</sup> Examples 7, 8, and 9 are translations of original German interactions.

## Lotze & Greilich: Interactions with AI

684

685 The examples demonstrate that, in comparison to human–  
686 human communication—where speakers actively show that  
687 the topic of the discourse has entered the common ground  
688 and can therefore be pronominalised or omitted (e.g., “her  
689 first crime novel”, “yes, she was”)—users in Amazon Alexa–  
690 directed speech do not omit or pronominalise the topic of the  
691 sentence. Instead, *Leonardo DiCaprio* and *Thomas Mann* are  
692 continuously repeated throughout the interaction.

693 We also observe that the system’s responses vary in terms of  
694 coherence:

- 695 1. Seemingly coherent responses that are treated as  
696 adequate and correct, e.g., “Thomas Mann was born in  
697 Lübeck.”
- 698 2. *Quasi-coherent* responses, e.g., „Leonardo DiCaprio  
699 has an estimated net worth of approximately two hundred  
700 thirty-two million five hundred thousand euros.”
- 701 3. Default responses provided when the system cannot  
702 generate an answer, e.g., “I’m not sure, unfortunately”.

703

704 In the current analysis, we rather focus on prosodic variation  
705 in Amazon-Alexa directed speech, depending on the  
706 coherence in Alexa’s responses. The data was collected in  
707 2019, and the analysed recording is part of a bigger corpus of  
708 partially scripted interactions with the German version of  
709 Amazon Alexa (cf. Greilich, in preparation). The primary  
710 objective of the case study is to provide further insight into  
711 how a user with no prior experience using Amazon Alexa or  
712 similar systems prosodically marks referents that were  
713 mentioned in the discourse.

714 In human communication, prosody is relevant to discourse  
715 continuity in several ways, including variation in pitch height,  
716 intensity, and duration of a referring expression, which is  
717 often associated with referents being newly introduced to the  
718 discourse, e.g., through wh-questions: “When was Thomas  
719 Mann born?”, vs. deaccenting of referring expressions when  
720 users continue interacting on the same topic and the central  
721 theme remains the same, e.g. “When did he receive his  
722 Nobelprice?”<sup>2</sup>. Precisely this type of topic maintenance can  
723 be studied using the data collected through the following  
724 experimental procedure: The study participant read the  
725 instructions stating that their task was to answer the questions

---

<sup>2</sup> In Amazon Alexa-directed speech, the user avoided using pronouns, therefore, deaccenting would be expected on the prosodic level.

## Lotze & Greilich: Interactions with AI

726 shown on the laptop screen with the help of Amazon Alexa.  
727 The stimuli featured a central theme with five questions  
728 associated with it (see Table 1). The questions were presented  
729 in the form of keywords, allowing the user to formulate their  
730 requests for information freely.

731

732 **Table 1:** Stimuli for the Data Collection of the Partially Scripted  
733 Interaction with Amazon Alexa<sup>3</sup>

Theme 1	Theme 2
Leonardo diCaprio	Thomas Mann
Age	Birthplace
Honorary for Titanic	Nobelprice (Year 19 )
Film partner in Titanic	Nobelprice for which work
Family	Famous works
Oscar ?	Citizenship

734

735 For the analysis of topic maintenance in the present study,  
736 we examined the different conditions resulting from Amcf.  
737 Amazon Alexa's various types of responses. Following the  
738 previously mentioned types of coherence, we coded the  
739 system responses into:

740

- 741 • *start* – users introduce the new referent to the  
742 interaction /
- 743 • *chr (coherent)* – Amazon Alexa provided a seemingly  
744 coherent and correct response to the previously asked  
745 question /
- 746 • *q-chr (quasi-coherent)* – as mentioned before, after  
747 Lotze (2016) /
- 748 • *NA<sub>4</sub>* – The system provided a default reply stating that  
749 it did not have the requested information at hand or did  
750 not understand the speaker.

751

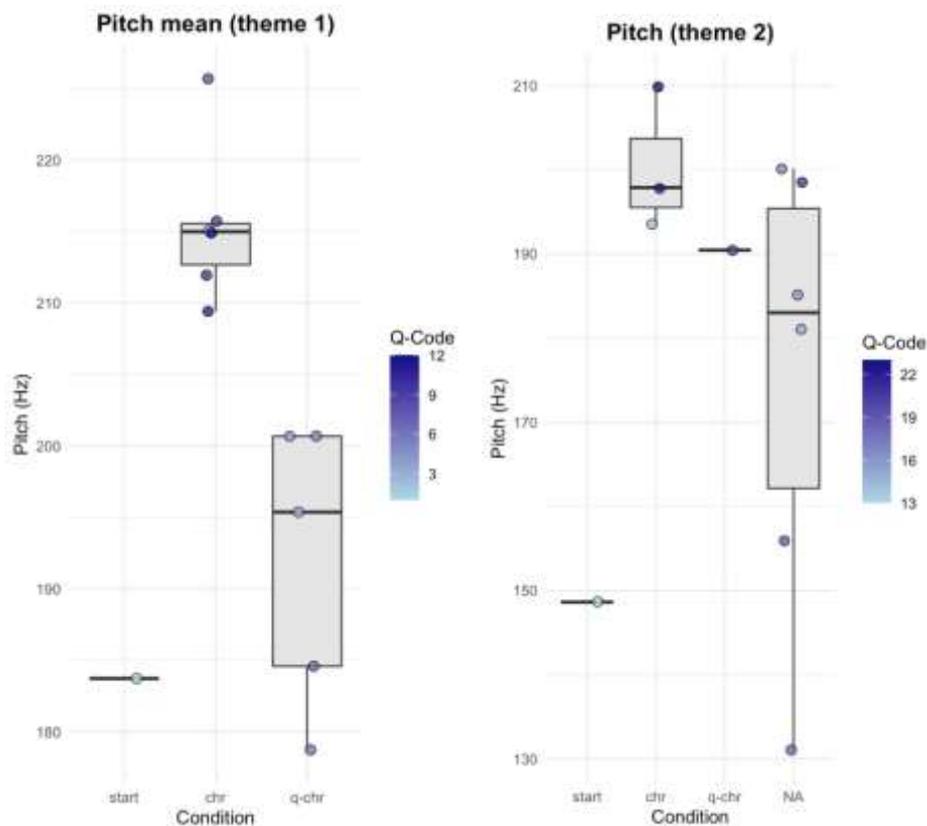
752 For the acoustic analysis in Praat (cf. Boersma/Weenink, 2025),  
753 referring expressions “Leonardo diCaprio” and “Thomas Mann”  
754 were manually annotated, (n=24). For each token, measurements of  
755 pitch (F0), duration, and intensity were extracted using Praat’s  
756 standard analysis tools for listing mean values in the selection.  
757 Based on the qualitative analysis of the HCI dialogues, where  
758 Amazon Alexa-directed speech lacks pronominalisation and the  
759 omission of the main topic of interaction, we hypothesise that there  
760 is a similar pattern at the linguistic level of prosody and that the way

<sup>3</sup> The table shows the translated version of the stimuli. In the original data elicitation instrument, the stimuli is in German.

<sup>4</sup> The category NA (no answer) is only relevant for the second stimuli Thomas Mann).

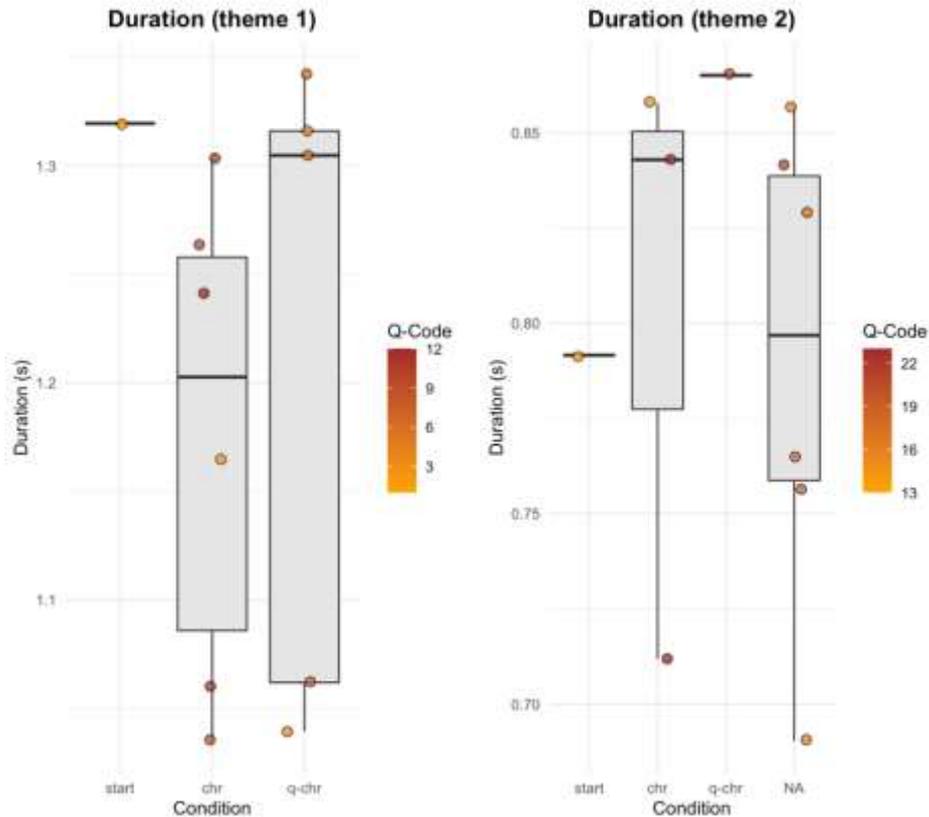
## Lotze & Greilich: Interactions with AI

761 Amazon Alexa responds to questions has a direct influence on the  
762 user's subsequent prosodic marking of referring expressions:  
763 We hypothesise that  
764 **H1:** In HCI, the prosodic marking of the previously  
765 mentioned referent is not characterized through deaccenting,  
766 namely, we cannot see an evident decline in pitch mean,  
767 intensity and duration of the main referent (topic of the  
768 utterance) depending on the question number (Q-Code).  
769 **H2:** Depending on the coherence condition (start, chr, q-chr,  
770 NA), the user might choose different strategies in follow-up  
771 questions: more deaccenting after seemingly coherent replies,  
772 vs. more prominence on the main referent after *quasi-*  
773 *coherent* responses, vs. variation after no answer was  
774 provided (NA).  
775



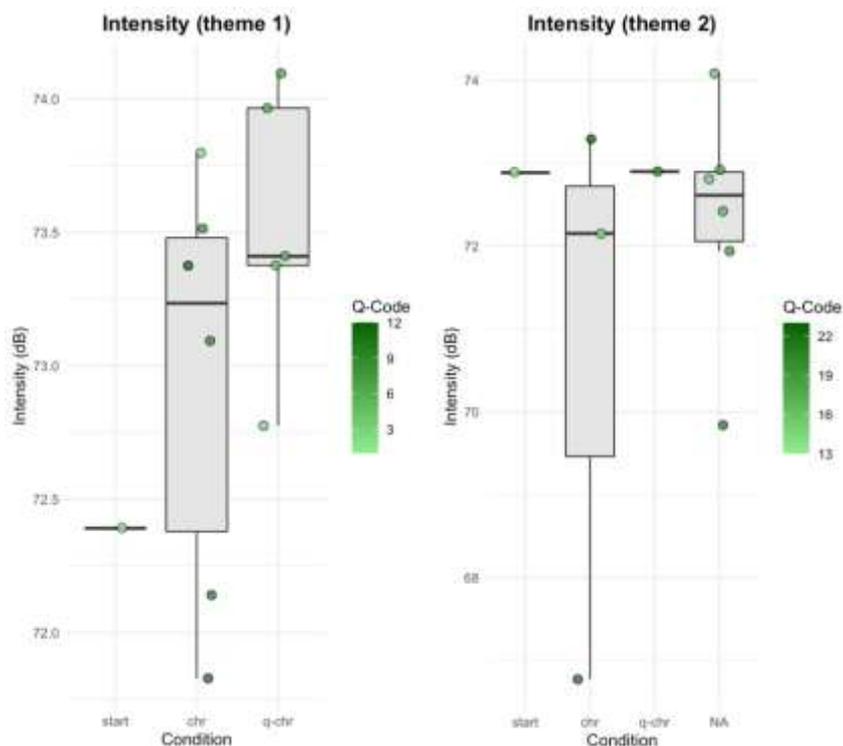
776 **Figure 9:** Interaction between Coherence Condition and Users' Mean  
777 Pitch F0 for the Referring Expression, (Main Topic/Theme), with Colour  
778 Coding indicating how early or late in the Interaction the Referring  
779 Expression occurred (Q-Code).  
780  
781  
782

## Lotze & Greilich: Interactions with AI



783  
784  
785  
786  
787  
788  
789

**Figure 10:** Interaction between Coherence Condition and Duration of the Referring Expression, (main topic/theme of the discourse), with Colour Coding indicating how early or late in the Interaction the Referring Expression occurred (Q-Code).



790

Discussion

## Lotze & Greilich: Interactions with AI

791 **Figure 11:** Interaction between Coherence Condition and Intensity of  
792 the Referring Expression, (main topic/theme of the discourse), with  
793 Colour Coding indicating how early or late in the Interaction the  
794 Referring Expression occurred (Q-Code).  
795

796 Figures 9, 10, and 11 present boxplots of pitch (Hz), duration  
797 (s), and intensity (dB) across the conditions (start, chr, q-chr,  
798 and NA) for stimuli themes 1 and 2 respectively. Each boxplot  
799 visualises the distribution of the respective acoustic  
800 parameter, while overlaid jittered points represent individual  
801 observations, color-coded by Q-Code (ranging from 1 to 12  
802 for the first theme and from 13 to 22 for the second stimuli  
803 theme). A continuous colour gradient is used to reflect the Q-  
804 Code values, with lighter shades indicating earlier mentions  
805 in the discourse and darker shades representing later ones.  
806 Boxplots were deliberately used to summarise the  
807 distribution of acoustic measures per condition, while jittered  
808 points display individual observations, enabling transparent  
809 visualisation despite the single-case design.  
810

811 The results suggest that the first hypotheses were partially  
812 confirmed. Pitch measurements indicate a tendency for rising  
813 contours, rather than the HHC-like lowering, throughout the  
814 interaction. For duration, results showed that in the second  
815 theme, the main referent was lengthened during utterance  
816 formulation. In terms of intensity, measurements increased  
817 following the start of the interaction in the first theme and  
818 remained consistently high. Notably, only the duration  
819 measurements for the first stimulus theme showed a decrease  
820 in the main referent's duration, specifically in the condition  
821 where Amazon Alexa provided a seemingly coherent  
822 response. This leads to the second part of our hypothesis:  
823 whether prosodic patterns vary across conditions depending  
824 on the linguistic output and perceived performance of  
825 Amazon Alexa.

826 Apart from the duration measurements in the first stimulus  
827 theme, we did not observe prosodic patterns typically  
828 associated with topic continuity, even when the system's  
829 response appeared coherent. However, when comparing  
830 prosodic marking across different conditions, we observed  
831 higher mean pitch values following coherent system turns in  
832 both stimulus themes, as well as increased intensity after  
833 quasi-coherent responses in the first stimulus theme.  
834 Such findings show that even after the speaker mentioned  
835 and the system correctly processed the referent, the users do  
836 not tend to mark it as a topic of the discourse prosodically but

## Lotze & Greilich: Interactions with AI

837 rather adjust the level of prominence towards the listener's  
838 needs and recalibrate from one turn to another. At the same  
839 time, the drop in duration measurements after a seemingly  
840 correct response aligns with the CASA/MASA paradigm  
841 (“Computers/Media are Social Actors”, cf. Reeves/Nass 1996):  
842 after the system exhibits more anthropomorphic linguistic  
843 behaviour, the user automatically reacts by shortening  
844 duration and thus partially reducing the prominence of the  
845 given referring expression. Reeves and Nass (1996) argue that  
846 individuals tend to apply a social model when confronted  
847 with complex entities whose internal mechanisms remain  
848 opaque. Linguistically, this tendency is reflected in the  
849 transfer of ritualised interactional protocols (cf. Sacks et al.  
850 1992), frames and scripts (cf. Fillmore 1976), and politeness  
851 strategies (cf. Brown / Levinson 1987) from HHC to HCI.

852  
853 Alongside the discussion of the findings, it is important to  
854 acknowledge the limitations of the analysis. Topic continuity  
855 is not the only factor influencing prosodic variation in  
856 referring expressions / syntactic role, positional factors, and  
857 other discourse-related elements should be considered in  
858 future analyses.

### 859 **9 Discussion: Quasi-Coherence as a Future Challenge for** 860 **Useres of AI**

861  
862 In this paper, we have demonstrated that quasi-coherence  
863 constitutes not merely a marginal phenomenon, but a third  
864 distinct and analytically relevant category alongside  
865 coherence and incoherence in HCI. Drawing on diverse  
866 quantitative methods, we have shown that users respond  
867 differently to quasi-coherent system replies than to coherent  
868 ones. In our 2016 corpora on rule- and plan-based chatbots,  
869 users frequently reacted to incoherent or quasi-coherent  
870 replies with abrupt termination of the dialogue. By contrast,  
871 in more recent spoken interactions with Amazon Alexa, users  
872 tend to exhibit subtler, prosodically marked responses:  
873 follow-up questions show distinct patterns of  
874 (de)accentuation depending on whether the preceding reply  
875 was coherent, incoherent, or quasi-coherent. It seems as  
876 though quasi-coherent replies—which do not occur in  
877 HHC—cause audible signs of hesitation and confusion.  
878 Coherent replies, on the other hand, tend to elicit task-  
879 oriented continuation, often indistinguishable from human  
880 dialogue behaviour.

## Lotze & Greilich: Interactions with AI

881 These findings indicate that quasi-coherence is not only  
882 relevant for linguistic categorisation in HCI, but also has  
883 concrete implications for user experience. While incoherent  
884 responses tend to frustrate users, quasi-coherent responses  
885 seem to provoke cognitive uncertainty, hesitation, and  
886 disorientation.

887 More recent work from our research group shows that users  
888 exhibit similar patterns of hesitation and frustration even  
889 when interacting with more advanced dialogue systems,  
890 including intent-based bots, large language models, and  
891 generative systems like GPTs.

892 In current AI systems, we observe fewer instances of  
893 incoherence, while quasi-coherence and hallucinations are  
894 becoming more prevalent. This suggests that incoherence will  
895 soon play only a marginal role in HCI and quasi-coherent  
896 system replies will be the greater challenge for users.

897 Consequently, we anticipate that more advanced systems will  
898 be perceived as increasingly anthropomorphic, which will in  
899 turn have implications for the linguistic behaviour of their  
900 users.

901

### 902 **10 Implications and Outlook: To What Extent Does the** 903 **Empirical Evidence of *Quasi-Coherence* Interfere with the** 904 **CASA/MASA Approach (cf. Reeves / Nass 1996)?**

905 If we follow Clark and Fischer's (2023) proposition that users  
906 interpret AIs as depictions of social agents, then quasi-  
907 coherence marks a disruption in this interpretative  
908 framework. In such cases, users appear to suspend the  
909 anthropomorphic frame and withdraw interactional  
910 alignment.

911 Indeed, HCI seems to oscillate between two modes: when the  
912 system functions smoothly, users often treat it as a full-  
913 fledged interlocutor, applying well-established human  
914 dialogue strategies. This indicates a form of conceptual  
915 anthropomorphisation, cognitively triggered by the natural  
916 language interface that symbolically depicts human  
917 interaction. The technological design—linguistically, visually,  
918 and interactively anthropomorphic—thus invites users to  
919 perceive the system as a social actor.

920 Our findings support an interpretation within the  
921 CASA/MASA paradigm, only when system replies are  
922 coherent, incoherent or hallucinations that lack empirical  
923 evidence, but support an illusion of coherence. However,

## Lotze & Greilich: Interactions with AI

924 when confronted with quasi-coherence—semantically  
925 plausible yet logically inconsistent replies—this interpretive  
926 mode breaks down. Users no longer engage in  
927 anthropomorphisation / instead, they hesitate, recalibrate, or  
928 withdraw, precisely because quasi-coherence is a  
929 phenomenon unique to HCI and absent in HHC. As such, it  
930 challenges the fundamental communicative assumptions that  
931 typically govern social interaction.  
932 Thus, quasi-coherent system replies disrupt the attribution of  
933 social characteristics to the AI. As demonstrated in this paper,  
934 such user reactions do not primarily depend on the modality  
935 of the interaction—whether written or oral—but rather on  
936 the frequency of incoherent or quasi-coherent system  
937 responses. However, as conversational systems continue to  
938 improve in coherence and contextual appropriateness, quasi-  
939 coherent responses may gradually cease to play a significant  
940 role in HCI. In such cases, users are likely to interact with  
941 these systems in increasingly intuitive and automatic ways,  
942 exhibiting less reflective and more routinised behaviour—an  
943 evolution I anticipated in my divergence/convergence  
944 hypothesis:

945 „The better a given system functions, the smaller the  
946 divergence between user language and system language  
947 becomes. In this process, both user and system language  
948 converge equally: it is not only the system that is adapted to  
949 user behaviour through evaluation and refinement, but users  
950 also modify their linguistic behaviour in response to  
951 improved system performance. A clear tendency can be  
952 observed: while the development of system language  
953 progresses from conceptually written and monologic styles  
954 towards shorter, more dialogic, and conceptually oral turns,  
955 user language shifts from the terse expressions (i. e.  
956 „computer talk“) typical of command-line interfaces towards  
957 more elaborate and natural utterances. This development  
958 suggests a likely prognosis: with continued system  
959 improvement, user and system language will increasingly  
960 converge—approaching, from different directions, the  
961 communicative patterns characteristic of HHC.“ (Lotze  
962 2016: 360-361).

963 And the current version of ChatGPT-4 effortlessly provides a  
964 coherent, empirically accurate, and highly concise answer to  
965 the question: “Are there German verbs with more than three

966 arguments?”

What does "four-place verb" mean?

In valency theory, the valency of a verb refers to the number of arguments it requires to form a complete sentence. Typically, we distinguish:

- Monovalent (one-place): "Es regnet." ("It's raining." – only a subject or placeholder)
- Divalent (two-place): "Ich sehe ihn." ("I see him." – subject + object)
- Trivalent (three-place): "Ich gebe ihm ein Buch." ("I give him a book." – subject + indirect object + direct object)

A four-place verb would be a verb that requires four syntactic arguments, such as:  
Subject + dative object + accusative object + prepositional object

Do such verbs exist in German?  
Such constructions are rare, but theoretically possible – they are often seen as borderline cases in valency grammar and are sometimes debated. A commonly cited example is:  
"Er stellt ihr den Mann als Experten vor." ("He introduces the man to her as an expert.")

Analysis:

- Subject: "Er" ("he")
- Dative object: "Ihr" ("to her")
- Accusative object: "den Mann" ("the man")
- Prepositional object: "als Experten" ("as an expert")

In some frameworks (e.g. Helbig/Buscha), the verb „vorstellen“ in this sense ("to introduce someone to someone as something") is considered four-place.

Other examples (subject to debate):

- "Sie erklärte ihm den Vertrag für ungültig." ("She declared the contract invalid to him.")
- "Man schätzt ihn auf 100.000 Euro." ("He is estimated at 100,000 euros.")

Whether these truly involve four arguments depends on the grammatical theory – some classical valency models (e.g. Tesnière, Helbig/Buscha) support this view.

Conclusion:

Yes, German does have verbs that can be analysed as four-place, such as vorstellen in certain constructions.  
 However: These verbs are rare, and whether they are truly four-place depends on the theoretical framework applied.

967

968 **Figure 12:** Translation of ChatGPT4s Response to the German User  
969 Request

970

971 Despite such impressive capabilities, we do not assume that  
972 HCI will inevitably become increasingly anthropomorphic as  
973 systems continue to improve. Rather, we consider it more  
974 likely that distinct linguistic registers and communicative  
975 styles will emerge in response to different application  
976 contexts. Moreover, we assume that such registers may  
977 develop across various user groups and types of AI systems,  
978 leading to a broad diversity of interaction patterns.

979 Relevant factors influencing the linguistic style adopted by  
980 future users may include their conceptualisation of the AI —  
981 either as a tool or as a social communication partner — as  
982 well as the situational context of the interaction. Private  
983 communication with an AI companion on emotionally  
984 charged topics is likely to differ substantially from  
985 interactions with an AI assistant in professional or task-  
986 oriented environments. Consequently, HCI should not be  
987 understood as a single, unified emerging socio-cultural  
988 practice (cf. Lotze 2025), but rather as a multiplicity of such  
989 practices, shaped by diverse contexts, functions, and user  
990 expectations.

## 991 Bibliography

992 Baumann, Stefan / Grice, Martine (2006): The intonation of  
993 accessibility. In: *Journal of Pragmatics* 38 (10), 1636-1657.

## Lotze & Greilich: Interactions with AI

- 994 Bender, Emily M. / Gebru, Timnit / McMillan-Major,  
995 Angelina / Shmitchell, Shmargaret (2021): On the dangers  
996 of stochastic xs Q1S1parrots: Can language models be too  
997 big? In: *Proceedings of the 2021 ACM Conference on*  
998 *Fairness, Accountability, and Transparency (FAccT '21)*.  
999 New York: Association for Computing Machinery, 610–  
1000 623. <https://doi.org/10.1145/3442188.3445922>
- 1001 Boersma, Paul / Weenink, David (2025): *Praat: doing*  
1002 *phonetics by computer [Computer program]*. Retrieved  
1003 from <http://www.praat.org/>.
- 1004 Brinker, Klaus / Hagemann, Jörg (2001): Themenstruktur und  
1005 Themententfaltung in Gesprächen. In: Brinker, Klaus /  
1006 Antos, Gerd / Heinemann, Wolfgang / Sager, Sven Frederik  
1007 (Hrsg.): *Text- und Gesprächslinguistik*. Wiesbaden: VS-  
1008 Verlag, 63–81.
- 1009 Brinker, Klaus / Sager, Sven Frederik (1989): *Linguistische*  
1010 *Gesprächsanalyse*. Berlin: Erich Schmidt Verlag.
- 1011 Clark, Herbert H. / Fischer, Kerstin (2023): Social robots as  
1012 depictions of social agents. In: *Behavioral and Brain*  
1013 *Sciences*, 1–33. <https://doi.org/10.1017/S0140525X22000668>
- 1014 Cohn, Michelle / Zellou, Georgia (2021): Prosodic differences  
1015 in human-and Alexa-directed speech, but similar local  
1016 intelligibility adjustments. In: *Frontiers in Communication*  
1017 (6), doi: 10.3389/fcomm.2021.675704
- 1018 Dippold, Doris (2023): ‘Can I have the scan on Tuesday?’ User  
1019 repair in interaction with a task-oriented chatbot and the  
1020 question of communication skills for AI. In: *Journal of*  
1021 *Pragmatics* (204), 21–32.
- 1022 Farquhar, Samuel / Kossen, Janis / Kuhn, Lukas / Gal, Yarin  
1023 (2024): Detecting hallucinations in large language models  
1024 using semantic entropy. London: *Nature* (630, 8017), 625–  
1025 630. <https://doi.org/10.1038/s41586-024-07352-2>
- 1026 Fillmore, Charles J. (1976): Frame semantics and the nature of  
1027 language. In: *Annals of the New York Academy of Sciences*  
1028 (280), 20–32.
- 1029 Fischer, Kerstin (2006): What computer talk is and isn't. In:  
1030 *Human-Computer Conversation as Intercultural*  
1031 *Communication*, 17.
- 1032 Givón, Talmy (1983): *Topic continuity in discourse – A*  
1033 *quantitative cross-language study*. Amsterdam: Benjamins.
- 1034 Givón, Talmy (1995): Coherence in text vs. coherence in  
1035 mind. In: Gernsbacher, Morton Ann / Givón, Talmy (Hrsg.):

## Lotze & Greilich: Interactions with AI

- 1036 *Coherence in Spontaneous Text*. Amsterdam: Benjamins,  
1037 59–115.
- 1038 Greimas, Algirdas J. (1974): Die Isotopie der Rede. In:  
1039 Kallmeyer, W. / Klein, W. / Meyer-Hermann, R. / Netzer,  
1040 K. / Siebert, H. J. (Hrsg.): *Lektürekolleg zur Textlinguistik*.  
1041 Frankfurt/M.: Athenäum Fischer (Band 2: Reader) 126–152.
- 1042 Habermas, Jürgen (1993): *Theorie des kommunikativen*  
1043 *Handelns*. Frankfurt/M.: Suhrkamp.
- 1044 Halliday, Michael Alexander Kirkwood / Hasan, Ruqaiya  
1045 (1976): *Cohesion in English*. London: Longman.
- 1046 Koch, Peter / Oesterreicher, Wulf (1994): Schrift und  
1047 Schriftlichkeit. In: Günther, Hartmut / Ludwig, Otto (Hrsg.):  
1048 *Schrift und Schriftlichkeit. Ein interdisziplinäres*  
1049 *Handbuch internationaler Forschung*. Bd. 2. Berlin/New  
1050 York: de Gruyter, 587–603.
- 1051 Lotze, Netaya / et al. (2011): *Meaningful Physical Changes*  
1052 *Mediate Lexical–Semantic Integration: Top-down and*  
1053 *Form-based Bottom-up Information Sources Interact in the*  
1054 *N400*. Amsterdam: Elsevier (*Neuropsychologia*, 49.13),  
1055 3573–3582.  
1056 <https://doi.org/10.1016/j.neuropsychologia.2011.09.00>
- 1057 Lotze, Netaya (2016): *Chatbots: Eine linguistische Analyse*.  
1058 Frankfurt/M.: Peter Lang. <https://doi.org/10.3726/b10402>
- 1059 Lotze, Netaya (2018): Zur sprachlichen Interaktion mit  
1060 Chatbots – Eine linguistische Perspektive. In: Hug, Theo /  
1061 Pallaver, Günther (Hrsg.): *Talk with the Bots –*  
1062 *Gesprächsroboter und Social Bots im Diskurs*. Innsbruck:  
1063 Innsbruck University Press, 9–50.
- 1064 Lotze, Netaya (2020): Künstliche Intelligenz im Dialog – ein  
1065 methodologisches Konzept zur Analyse von Mensch-  
1066 Maschine-Interaktion. In: *Beiträge zur Jahrestagung 2019*  
1067 *des Instituts für Deutsche Sprache. „Deutsch in Sozialen*  
1068 *Medien: interaktiv, multimodal, vielfältig“*.
- 1069 Lotze, Netaya (2025): Human-Machine Interaction as a  
1070 Complex Socio-Linguistic Practice. In: Habscheid, Stephan  
1071 / Hector, Tobias / Hoffmann, Diana / Waldecker, David  
1072 (Hrsg.): *Voice Assistants in Private Homes: Media, Data*  
1073 *and Language in Interaction and Discourse*. Bielefeld:  
1074 transcript, 105–144.  
1075 <https://doi.org/10.1515/9783839472002-005>
- 1076 Pickering, Martin J. / Garrod, Simon (2004): Towards a  
1077 mechanistic psychology of dialogue. In: *Behavioural and*

## Lotze & Greilich: Interactions with AI

- 1078 *Brain Sciences* 27, 169–225.  
1079 <https://doi.org/10.1017/S0140525X04000056>
- 1080 Reeves, Byron / Nass, Clifford I. (1996): *The Media Equation:*  
1081 *How People Treat Computers, Television, and New Media*  
1082 *Like Real People and Places*. Stanford: CSLI Publications /  
1083 Cambridge: Cambridge University Press.
- 1084 Sacks, Harvey / Schegloff, Emanuel A. / Jefferson, Gail (1992):  
1085 *Lectures on Conversation*. Oxford: Blackwell (Vol. I & II).
- 1086 Schatzki, Theodore R. / Knorr Cetina, Karin / von Savigny,  
1087 Eike (2001): *The Practice Turn in Contemporary Theory*.  
1088 London: Routledge.
- 1089 Searle, John R. (1979): *Expression and Meaning: Studies in the*  
1090 *Theory of Speech Acts*. Cambridge: Cambridge University  
1091 Press.
- 1092 Searle, John R. (1980): Minds, brains and programs. In:  
1093 *Behavioral and Brain Sciences* 3, 417–457.
- 1094 Searle, John R. (2006): *Geist: Eine Einführung*. Frankfurt/M.:  
1095 Suhrkamp.
- 1096 Siegert, Ingo / Krueger, Julia (2020). “Speech melody and  
1097 speech content Didn’t fit together” —differences in speech  
1098 behavior for device directed and human directed  
1099 interactions. In: *Advances in data science: methodologies*  
1100 *and applications*. Cham: Springer International Publishing,  
1101 65–95.
- 1102 Stalnaker, Robert (2002): Common ground. In: *Linguistics*  
1103 *and Philosophy* 25 (5–6), 701–721.  
1104 <https://doi.org/10.1023/A:1020867916902>  
1105  
1106