

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.

the editors

jfml

Journal für Medienlinguistik
Journal for Media Linguistics

Discussion Paper 3 (2026):
submitted on 25 March 2026

Reviews and comments at:
<https://dp.jfml.org/2026/opr-hetjens-appeals-to-generative-authority/>

Discussion Paper

1 **Appeals to Generative Authority and the AI-Reference**
2 **Paradox**
3 **A Pragmatic Analysis of References to LLMs in Online Debates**

4 *Dominik Hetjens*

5 **1 Introduction**

6 Since the advent of generative AI, automation bias has
7 become a topic of increasing concern, and a recent survey by
8 the consulting agency Ernst & Young has shown that blind
9 trust in AI-generated content might indeed be a widespread
10 phenomenon (Ernst & Young 2025).

11 While researchers have pointed out the relevance of
12 analysing the use of LLMs for linguistics (Felder/Kückelhaus
13 2025), and one as-yet unpublished study has investigated how
14 LLMs are framed in newspapers and magazines (Kückelhaus
15 in prep.), no publication has yet applied methods from
16 pragmatics and discourse linguistics to investigate how
17 people in interactional online communities reference
18 generative AI as a source of information or to achieve a
19 certain effect in a debate – thereby ascribing epistemic
20 authority to the AI in question.

21 Utilising the social news aggregator website Reddit as a
22 source of authentic interactions (see Messerli et al. 2025:
23 227), this explorative study aims to uncover the linguistic
24 means used to introduce AI and AI-generated content into a
25 debate, as well as the epistemic markers used to express
26 certainty or demand trust in that output. Furthermore, it
27 analyses how both are relevant to the pragmatic strategies
28 employed to save face, win arguments, and exchange
29 knowledge on the website.

Dominik Hetjens. Generative Authority and the AI Reference Paradox

Discussion Paper

30 Consistent with the platform's interactional character, the
31 qualitative analysis extends beyond the text to include
32 various linguistic and extra-linguistic data, such as the
33 number of replies and the position of the comment in the
34 structure of the debate. These metrics allow for an
35 investigation into the interactional profiles and the effects of
36 the comments mentioning AI. The one used as an example in
37 this study, ChatGPT, represents a conversational agent based
38 on a large language model (LLM).¹

39 The corpus comprises German-language interactions
40 harvested from Reddit between May 2025 and February
41 2026, a period characterised by the normalisation of LLM
42 usage in everyday digital communication (see Ernst & Young
43 2025 for data on AI usage). The analysis is grounded in the
44 intersection of Computer-Mediated Communication (CMC),
45 pragmatics, and discourse analysis, the relevant concepts of
46 which will be briefly outlined in the first two chapters,
47 focusing on their applicability for the context in question. To
48 do so, a short introduction to the basic mechanics of Reddit
49 will be included, providing the basis for the following
50 methods section. Afterwards, the results will be presented in
51 an overview that will focus on the general function of the
52 comments as well as the evidentials and epistemics used.
53 Subsequently, individual comments will be analysed in
54 greater detail, grouped into previously developed types. The
55 discussion will focus on how these results fit into the larger
56 context of CMC and pragmatics, and how future studies
57 might expand on the basis of these findings.

58 **2 Reddit Linguistics**

59 As one of the most frequently visited websites on the
60 internet, with more than a billion monthly visitors, and as one
61 of the major suppliers of training data for the largest LLMs,
62 Reddit's relevance for academic study can hardly be
63 overstated (Similarweb 2026; Brown et al. 2020). While
64 research on the platform is growing, linguistic articles remain
65 rare, given the importance of the website (see Messerli et al.
66 2025 for an overview of the current state of research). The

1 In the following, "AI" is used as an umbrella term, and "LLM" denotes the models behind conversational agents like ChatGPT.

Dominik Hetjens. Generative Authority and the AI Reference Paradox

67 studies published so far utilise Reddit as a corpus, or a system
68 of corpora, as its structure allows for the analysis of authentic
69 language use in interaction (Messerli et al. 2025: 227;
70 Blombach 2020: 6310). The platform’s division into hundreds
71 of thousands of sub-communities has led to studies focusing
72 on specific areas of Computer-Mediated Communication,
73 among them expert-laypeople interactions and advice-giving
74 practices (Leuckert 2024; Lu 2025; Chen et al. 2021).

75 Some of the earlier studies analysing comments on Reddit
76 have concentrated on stance-taking, as the interactional
77 character of the platform lends itself to this kind of research
78 (e.g. Kiesling et al. 2018). Unlike monolithic forums, Reddit’s
79 architecture consists of “subreddits” – user-created
80 communities governed by specific thematic norms and a “tree
81 structure” of nested comments (Medvedev et al. 2020: 4;
82 Massanari 2015: 4). The platform’s defining feature is its
83 binary “karma” system, a type of crowdsourced review
84 mechanism that determines visibility of posts (which may
85 contain any form of linked content) as well as individual
86 comments in the discussion of these posts (Glenski &
87 Weninger 2017: 1). Operating under a model of pseudonymity,
88 users assume fluid roles as posters, commenters, or
89 moderators, with the latter curating the discursive “micro-
90 climate” of individual spaces through both communicative
91 and infrastructural regulation (van der Nagel 2017; Matias
92 2019; Leavitt 2015). In these spaces, the expertise and
93 credibility of individual statements and users are subject to
94 constant negotiation and openly visible evaluation, making
95 them an ideal playground for questions regarding epistemic
96 stance and authority (Hetjens in prep.; Leuckert 2024; Rudolf
97 von Rohr/Thurnherr/Locher 2018).

98 However, despite the German-language Reddit community
99 being the second-largest group on the website (Blombach et
100 al. 2020: 6310), and user numbers far surpass those of any
101 reputable German-language news website (Similarweb 2026),
102 relatively little research has yet focused on the language of
103 the German Reddit sphere (see Messerli et al. 2025 for an
104 overview).

105 **3 Evidentiality and Epistemic Modality**

106 As this study will examine how LLMs are treated as a source
107 of information, evidentiality – the linguistic marking of the
108 source of one's knowledge (for the German context, see
109 Diewald/Smirnova 2010) – and epistemic modality – the
110 expression of the speaker's degree of certainty or
111 commitment toward the truth of a proposition (Mortelmans
112 2000: 132) – provide the necessary framework to analyse
113 how users utilise and interact with AI-generated output. For
114 our context, they will not be analysed in regard to their
115 quality as grammaticalised lexical units, but as wider
116 phenomena to be observed in interactional data.

117 While the precise distinction between evidentiality and
118 epistemic modality is still being debated, and there is good
119 cause to view the two as overlapping categories, several
120 researchers have identified certain evidential strategies in
121 German (Diewald/Smirnova 2010). This paper examines
122 instances of reportative evidentiality (Mortelmans 2000: 131),
123 in which AI is directly named as a source of information. The
124 most common strategies for reportative evidentiality are the
125 use of lexical markers such as "laut" ("according to") as well as
126 the reportative use of modal verbs like "sollen" and "müssen"
127 ("should", "must") (Mortelmans 2000). However, the
128 interactional nature of the source of information – LLMs are
129 sources as well as constantly available artificial
130 conversational partners for potentially all participants in a
131 discussion – suggests that more complex evidentials might
132 arise when referring to them in debates.

133 While primarily indicating a source of information, these
134 evidential markers also serve to distance the speaker from
135 the propositional content (Mortelmans 2000: 131).
136 Consequently, they can function similarly to epistemic
137 markers, which signal the speaker's degree of commitment to
138 the proposition of a statement. In German, this epistemic
139 function is frequently fulfilled by modal verbs such as
140 "können" (can) or "dürfen" (may), verbs of cognition like
141 "wissen" (to know), "glauben" (to believe), or "denken" (to
142 think), and modal particles, including "mal" (just), "halt"
143 (simply), and "doch" (after all). The latter have no direct
144 equivalent in English and can be used to express the
145 speaker's attitude toward the proposition of their own

Dominik Hetjens. Generative Authority and the AI Reference Paradox

146 statement. In oral language, phrases like “Ich denk” (“I think”)
147 or “Ich glaub” (“I believe”) function as pragmatic hedges
148 (Wilton 2019: 222), allowing speakers to manage face and
149 mitigate potential conflict by presenting claims as subjective
150 interpretations rather than indisputable facts. As Computer-
151 Mediated Communication often represents language that
152 arguably can be described as conceptually oral, Dayter &
153 Messerli (2021) find similar hedges on Reddit, specifically the
154 phrase “I’d say”.

155 **4 Strategies of creating authority, credibility, and expertise on**
156 **Reddit**

157 A small number of studies have used different approaches for
158 analysing successful comments on the platform, investigating
159 related concepts relevant to this study, such as credibility,
160 authority, and expertise. While theoretical and
161 methodological considerations lead to varying terminologies,
162 the observations made are relevant to all three. For this
163 study, we will sum them up as “pragmatic strategies”, the
164 intentional employment of linguistic means to achieve a
165 certain communicative goal. Among these, the following
166 stand out:

167 **1. Introducing personal background:** Rudolf von
168 Rohr/Thurnherr/Locher (2018) find that, in advice-giving
169 subreddits, users frequently utilise what the researchers call
170 “background moves”: They explicitly state their personal
171 experience or professional qualifications (e.g. “As a nurse ...”).
172 This personal background can then be validated through
173 upvotes, suggesting that it might be in line with what others
174 have experienced.

175 **2. Referencing external authorities:** Another major
176 strategy in a number of subreddits is referencing external
177 authorities. This has been observed in debate-oriented
178 subreddits like r/Changemyview, among others (Messerli et
179 al. 2025). In his study on meta-perspectives on dictionaries,
180 Sven Leuckert identifies a type of reference to dictionaries as
181 evidential and epistemic moves, “giving authority and
182 credibility to a statement” (Leuckert 2024: 383). At the same
183 time, consulting dictionaries ascribes authority to the
184 lexicographical resources in question – an authority that is

Dominik Hetjens. Generative Authority and the AI Reference Paradox

185 questioned and negotiated in the comments (Leuckert 2024:
186 381).

187 **3. Presentation as persuadable:** In the sub-community
188 r/ChangeMyView, posters construct an identity of being
189 open-minded and “persadable” by frequently using
190 concessions (e.g., “although”, “even if”) alongside first-person
191 pronouns to qualify their *own* beliefs. When attempting to
192 persuade others, successful commenters build their case by
193 using concessives to qualify the beliefs of an imaginary third
194 party, which allows them to dismantle opposing arguments
195 without directly threatening the interlocutor's face (Messerli
196 et al. 2025: 230).

197 **5 Methods**

198 The general approach is inspired by a type of functional
199 analysis developed by Sven Leuckert in his study of
200 references to dictionaries on Reddit (Leuckert 2024), who
201 himself loosely based his method on thematic analysis
202 initially developed for working with interview data in
203 psychology, that was later applied to linguistics (Buschfeld &
204 Ahlers 2022). This approach is characterised by careful
205 iterative close reading of comments and the identification of
206 features relevant to the research question. Based on the
207 results, Leuckert then created a taxonomy of functional
208 categories, and assigned each example he found relevant for
209 the research question to one (or more) of the categories
210 (Leuckert: 382).

211 This study uses a similar approach to answer the overall
212 question: “How do people treat generative AI as a
213 communicative device in a small corpus of interactional
214 online-comments?”. To do so, it carries out a functional
215 analysis of the comments via close-reading, and develops a
216 taxonomy of functional categories based on the results. In
217 addition, this study records the evidential and epistemic
218 markers used, and investigates the communicative
219 surroundings of the relevant comments. This includes both
220 qualitative analysis of linguistic data (previous comments,
221 replies) and quantitative recordings of extra-linguistic
222 metadata, to show, where possible, how the reference of AI is
223 received. These metrics will be correlated with the functional

Dominik Hetjens. Generative Authority and the AI Reference Paradox

224 categories to create a typology of references to generative AI
225 in online-comments. To enable this type of analysis, the first
226 step is to determine how many mentions of AI constitute
227 references to AI as a source or as evidence for a proposition,
228 rather than contributions to discussions about AI itself or
229 meta-discussions accusing users of inadequate AI-usage.
230 These comments are then analysed, taking into account the
231 evidential and epistemic markers used, the form of reference
232 (quotations, paraphrases, contextualisations), the alignment of
233 individual comments with previous ones and replies, the
234 position of a comment in Reddit’s tree structure, and its
235 karma score.

236 The results will be presented in two parts: firstly, an
237 overview shows how many comments fall into which
238 category, which evidentials and epistemics are used, and
239 what patterns can be observed in both respects. Secondly,
240 individual examples are described in more detail, grouped
241 according to the previously developed typology, to illustrate
242 how the categories are constituted.

243 In this second part, there will be room to elaborate on the
244 communicative surroundings of comments, as well as the
245 extra-linguistic metadata, where deemed useful for the
246 analysis. An “L” will be used to indicate the level in Reddit’s
247 tree structure on which a comment can be found, with L1
248 signifying a comment that is a direct reply to the original post,
249 L2 a reply to an L1 comment, and so forth. Part two of the
250 results will introduce these metrics at the beginning of the
251 characterisation of each functional group, before individual
252 comments are discussed. Since the methodological process
253 involves hermeneutic cycles of grouping and re-grouping of
254 features into emerging categories, a strict division between
255 description and interpretation of results is not possible.

256 **6 Data**

257 The primary corpus contains the 100 most recent comments
258 mentioning ChatGPT (as of February 2026) in the largest
259 German community, *r/de* (Blombach 2020: 6310). ChatGPT
260 was chosen as it is the most well-known conversational agent
261 currently in widespread use, and can serve as an example of
262 how LLMs are treated as a whole. The subreddit features any

Dominik Hetjens. Generative Authority and the AI Reference Paradox

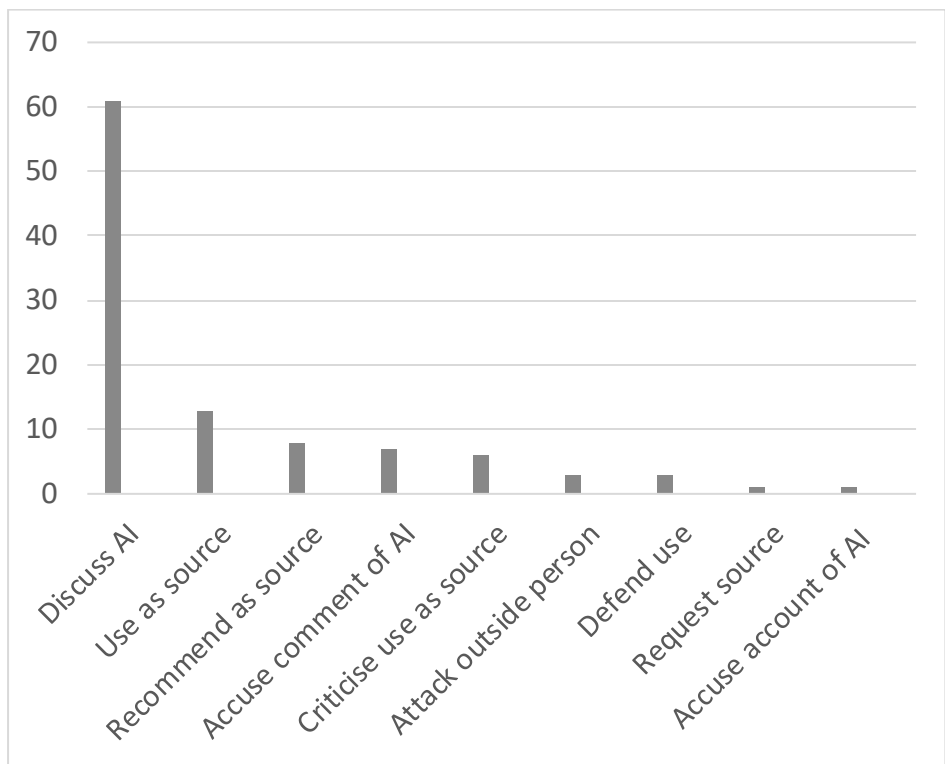
263 user-posted content deemed interesting or entertaining to a
264 German-speaking audience or those interested in the DACH
265 region – ranging from lighthearted memes to sensitive
266 political discussions – as long as the general rules of the
267 subreddit are followed. In February 2026, the vast majority of
268 posts linked to news articles, primarily regarding recent
269 political events. The community attracts approximately
270 700,000 weekly visitors and registers around 25,000
271 individual contributions per week (Reddit 2026).

272 To collect the data, the software *Arctic Shift* was used to
273 perform a case-insensitive search for “ChatGPT,” recording
274 the 100 most recent comments. All comments belonging to
275 the same comment trees as these 100 comments were taken
276 into account as a secondary corpus. Their content is
277 mentioned where relevant to the functional analysis of
278 comments from the primary corpus.

279 All Reddit comments analysed in this study represent
280 publicly available data from pseudonymous sources.
281 Nevertheless, this study does not cite individual usernames or
282 link to any specific comment thread to shield users from
283 investigations into their comment history or personal
284 information.

285 **7 Results**

286 7.1 Overview



287

288 **Figure 1:** Mentions of ChatGPT in the corpus

289 Of the 100 most recent comments mentioning “ChatGPT”, 61
290 discussed AI-related issues like OpenAI’s business strategies,
291 government policies, and similar topics. Among the remaining
292 comments, 13 referenced ChatGPT as a source of
293 information, 8 comments actively recommended its use,
294 though the actual function of these recommendations is
295 discussed separately at a later point, and 3 comments
296 defended the use of ChatGPT as a source. One user asked for
297 the source of a previous statement with the intention of
298 consulting ChatGPT about it.

299 A smaller subset of the corpus was characterised by
300 skepticism or criticism: 7 comments accused a specific
301 comment of being AI-generated, 6 comments criticised the
302 use of ChatGPT as a source, and 3 users criticised an outside
303 person of using ChatGPT in the wrong way. Finally, one
304 instance involved a direct accusation that a specific account
305 was an AI bot or was overusing the technology.

306 This distribution is in itself interesting, giving rise to a
307 number of questions. While there are comments critical of

Dominik Hetjens. Generative Authority and the AI Reference Paradox

308 using ChatGPT (6), far more comments do exactly that (13) or
309 recommend its use (8), indicating an already somewhat
310 established trust in the LLM. It would be worth investigating
311 the metadiscourse as well as possible patterns of accusing
312 users of AI-usage in a separate study. A preliminary look at
313 the data suggests that it is not the use of AI as a source that is
314 deemed problematic, but rather the act of passing off AI-
315 generated text as one's own. In this way, users appear to
316 defend the character of the platform as a place for authentic
317 interactions between human agents (Messerli et al. 2025: 227).
318 However, these aspects are beyond the scope of the present
319 study, as the focus remains on how AI is introduced into
320 debates as a source.

321 7.2 Referencing AI as a Source

322 Of the 13 comments doing so, 9 utilise the modal particle
323 “mal”, and 8 use the verb “fragen” (including one case of
324 “befragen”), either in the first person indicative, the infinitive
325 or the imperative in either the singular or the plural, resulting
326 in these two prototypical phrases:

- 327 (a) Ich habe mal ChatGPT gefragt ...
328 (I've just asked ChatGPT)
- 329 (b) Frag mal ChatGPT ...
330 (Just ask ChatGPT)

331 In some of the comments, hedges in the form of adverbial
332 phrases are inserted between “mal” and “ChatGPT”, such as
333 “aus Neugierde” (“out of curiosity”) or “aus Spaß” (“for fun”),
334 which can be seen as markers of justification indicating an
335 awareness of the meta-discourse on AI-usage, as is discussed
336 later. In these comments, “fragen” arguably has an evidential
337 quality, as it is used to mark the source of the subsequent
338 content. Similarly, the imperative “frag” doesn't necessarily
339 represent an actual command or request in the comments
340 included here (see comments of type 3 or 4 below).

341 The functional analysis shows that the references to
342 ChatGPT as a source in the dataset fall into four distinct
343 functional categories (Tab. 1). As is explained later in more
344 detail, the metrics observed in response to the questions
345 posed in Section 5 correlate with these functional categories,

Dominik Hetjens. Generative Authority and the AI Reference Paradox

346 giving rise to distinct patterns that allow for the creation of a
347 typology of comments characterised by specific form-
348 function correlations. These types can be viewed as being
349 placed on a spectrum ranging from informative to
350 argumentative. Overall, more-informative comments (Types 1
351 and 2) tend to make the actual output of ChatGPT more
352 explicit, either by quoting it verbatim or by providing longer
353 paraphrases, while argumentative comments tend to only
354 briefly mention the specific output (Type 3) or even omit it
355 entirely, focusing instead on its implications for the
356 discussion at hand (Type 4).

Type	Role of AI	Occurrences	Function of AI reference	Form of Reference	Main functional category
1	Scribe	1	Providing text for outside use	Direct Quote	Informative
2	Informant	8	Introducing information	Paraphrase	
3	Corroborator	2	Providing evidence for proposition	Paraphrase	Argumentative
4	Arbiter	2	End debate	Conclusion	

357 **Table 1:** Types of References as a Source

358 In the following sections, we examine each of these types
359 more closely, first by giving an overview of their respective
360 profiles (main function, form of reference, level on which
361 comments typically occur, communicative surroundings, and
362 karma score where relevant), and then providing specific
363 examples of each category, alongside a short discussion. All
364 evidentials and epistemics used when introducing the AI into
365 the discussion are displayed in Tab. 2, sorted according to the
366 types described above.

Dominik Hetjens. Generative Authority and the AI Reference Paradox

Discussion Paper

	Evidentials & Epistemics	Translation
Type 1	Ich habe mal ChatGpt gefragt	I just asked ChatGpt
Type 2	Ich habe aus Neugierde mal ChatGPT gefragt	I just asked ChatGPT out of curiosity
	Ich habe mal ChatGPT und Gemini befragt	I just consulted both ChatGPT and Gemini
	Frag mal Chatgpt, hab grad mal aus Neugier eingegeben	Try asking ChatGPT; I've just typed it in myself out of curiosity.
	Laut chatgpt	According to Chatgpt
	Einfach mal eingeben in Chatgpt	Just type it into Chatgpt
	Zur Sicherheit habe ich soeben ChatGPT gefragt	I've just asked ChatGPT to be on the safe side
	Quelle: ChatGPT	Source: ChatGPT
	Ich habe mich mit Hilfe von ChatGPT etwas in die Thematik eingearbeitet	I have worked myself somewhat into the subject matter with the help of ChatGPT
Type 3	wenn ich ChatGPT frage	When I ask ChatGPT
	Jedes Mal, wenn ich ChatGPT frage	Every time I ask ChatGPT
Type 4	Ich habe mal aus Spaß die KI gefragt	I've just asked the AI for fun
	Fragt Mal chatgpt	Ask ChatGPT

367 **Table 2:** Evidentials and Epistemics

Dominik Hetjens. Generative Authority and the AI Reference Paradox

368 7.3 Typology of References

369 7.3.1 Type 1: Scribe References

370 There is only one comment of this type, which stands out
371 significantly from all other comments and can be viewed as
372 being at the extreme end of the spectrum outlined above. It
373 provides a full text generated by ChatGPT for a potential
374 application outside of Reddit. It occurs at L1 as a direct reply
375 to a post linking to an article on social security and renting
376 practices in Germany. In the comment, the user first
377 describes the prompt given to ChatGPT and then provides
378 the full text produced by the LLM: a lengthy letter asking a
379 landlord for a reduction in rent. By providing both the
380 prompt and the full text generated, this constitutes the most
381 transparent reference in the corpus. However, both the two
382 replies to this comment and its low karma score hint at the
383 comment being relatively unsuccessful. To introduce the
384 reference to ChatGPT, the prototypical phrase is used, and
385 no hedges or boosters are employed:

386 (1) Ich habe mal ChatGpt gefragt, wie man seinen Vermieter
387 höflich rügen könnte um mit hoher Wahrscheinlichkeit
388 eine Mietminderung zu erhalten (beste Wortwahl,
389 loben,..)
390 [...]

391 (I just asked ChatGPT how to politely reprimand a
392 landlord in order to have a high probability of receiving a
393 rent reduction (best choice of words, using praise,..)[...])

394 7.3.2 Type 2: Informant References

395 In this type of comment, new information is introduced into a
396 discussion, and ChatGPT is referenced as the source. The
397 information is paraphrased rather than quoted verbatim. This
398 constitutes the largest sub-group among the “reference as a
399 source” comments, with eight comments falling into this
400 category. They exclusively occur early in a thread’s tree
401 structure (L1 to L3), meaning they are not part of a discussion
402 that has already gained traction; their most common position
403 in the structure is L1 (five comments). All of them contain
404 information that arguably constitutes “hard facts” rather than
405 opinions, such as salary levels, the most frequently visited
406 websites in Germany, or the jobs available to people with a

Dominik Hetjens. Generative Authority and the AI Reference Paradox

407 certain background. They can be seen as mostly informative,
408 and it is often unclear what stance the commenter takes
409 towards previously made propositions, which can be seen
410 below:

411 (2) Laut chatgpt belaufen sich die Diäten aller Abgeordneter
412 der Landesparlamente und des Bundestages jährlich auf
413 zwischen 180 und 320 Mio. €. Wobei ein MdB fast
414 12,000 € monatlich erhält.

415 (According to ChatGPT, the salaries of all members of
416 the state parliaments and the Bundestag amount to
417 between €180 million and €320 million annually. A
418 member of the Bundestag (MdB) receives nearly €12,000
419 per month.)

420 The comment acts as a reply to one claiming that the money
421 politicians receive could be put to better use:

422 (3) Ich sag mal so, wenn die Politiker auf einen Teil ihres
423 Monstergehalts verzichten würden, müssten wir nicht
424 immer mehr abdrücken. Aber das wäre natürlich ein
425 unmöglicher Zustand

426 (Let's put it this way: if politicians gave up a portion of
427 their monster salaries, we wouldn't have to keep
428 coughing up more and more. But of course, that would
429 be an 'unthinkable' situation.)

430 Comment (2) does not explicitly state whether the
431 commenter agrees with this proposition, as the general
432 structure is informative: The comment starts with the
433 evidential phrase "Laut chatgpt", after which no epistemics or
434 further evidentials are used. There is no contextualisation
435 that would state the value of the comment or the information
436 within the discussion at hand. This general structure is the
437 same in longer comments:

438 (4) Ich habe aus Neugierde mal ChatGPT gefragt, was die
439 genannten Unternehmen so an wen gespendet haben.
440 Für die Schwarz-Gruppe und Volkswagen AG gibt es
441 keine konkreten Zahlen (also Angaben zu Spenden über
442 35.000€), aber die BMW AG hat der CDU und FDP
443 jeweils 855.000€ und 120.000€ gespendet. Bei den
444 anderen Unternehmen schätzt es aber auch die Union als
445 größten Empfänger ein.

Dominik Hetjens. Generative Authority and the AI Reference Paradox

446 (Out of curiosity, I just asked ChatGPT who the
447 mentioned companies have donated to. There aren't any
448 concrete figures for the Schwarz Group or Volkswagen
449 AG (meaning reports of donations over €35,000), but
450 BMW AG has donated €855,000 and €120,000 to the
451 CDU and FDP, respectively. For the other companies, it
452 also estimates the Union (CDU/CSU) to be the largest
453 recipient.)

454 Here, the evidential phrase is longer, corresponds to the
455 prototype described above, and contains “aus Neugierde” as a
456 hedge. The respective information, however, is presented as
457 factual, utilising no evidentials or epistemics, and verbs in the
458 indicative (“es gibt”, “hat gespendet”). The closing sentence,
459 on the other hand, uses “einschätzen” (“to estimate”), an
460 epistemic verb, whereby ChatGPT rather than the
461 commenter acts as the agent estimating the donations, which
462 constitutes an anthropomorphisation in line with the findings
463 of Felder & Kückelhaus (2025).

464 The comment sparks a debate on the validity of ChatGPT’s
465 results when used as a search engine. Interestingly, the
466 comment itself has a negative karma score (-10), yet the
467 subsequent metadiscourse ends in favour of using ChatGPT
468 as a search engine, as is apparent in the last comment of the
469 tree:

470 (5) klar, wenn deine Suche so einfach ist, dass alle
471 Ergebnisse auf der ersten Seite der Google-Ergebnisse
472 auftauchen, macht es mehr Sinn, zu googeln. Ist aber
473 nicht immer so und insbesondere wird Google auch
474 immer schlechter.

475 (Sure, if your search is so simple that all the results show
476 up on the first page of Google, it makes more sense to
477 just Google it. But that’s not always the case and Google
478 in particular is getting worse and worse.)

479 This comment indicating that ChatGPT can be used as a
480 search engine, and might even yield better results than
481 Google, depending on the circumstances, has a score of 4,
482 which is relatively high given that it occurs on L6 in a thread
483 in which high scores are rare, and there is no contradicting
484 reply, possibly suggesting alignment.

Dominik Hetjens. Generative Authority and the AI Reference Paradox

485 Other comments of type 2 reference ChatGPT as something
486 the commenters have tested, were content with, and would
487 recommend, such as this reply to someone asking for job
488 market advice as a foreigner living in Germany with a degree
489 in German studies:

490 (6) Frag mal Chatgpt, hab grad mal aus Neugier eingegeben:
491 Was kann man als Amerikaner mit einem Abschluss in
492 “german studies” in Deutschland arbeiten? Da kamen
493 sehr viele mögliche Bereiche bei raus, vielleicht ist ja
494 was interessantes für dich dabei.

495 (Ask ChatGPT, I just typed it in out of curiosity: What
496 kind of jobs can an American with a “German Studies”
497 degree do in Germany? A lot of possible fields came up;
498 maybe there’s something interesting in there for you.)

499 7.3.3 Type 3: Corroborator References

500 Two comments fall in this category, occurring on L2 and L4,
501 replying to a comment they disagree with. They reference
502 ChatGPT as a type of “supporting evidence” for their
503 previously stated proposition in a debate. Both comments of
504 this type use the conditional/temporal “wenn” (“if/when”) in
505 combination with the prototypical verb “fragen” (“to ask”), a
506 feature exclusive to comments of this type. The first example
507 is part of a discussion about whether or not teachers’ salaries
508 in Germany are too high:

509 (7) Ja keine Ahnung, ich hab vielleicht einen schlechten
510 Vergleich, aber ich hab Freunde, die nach einem
511 Bachelor schon mehr verdient haben.
512 (Ingenieursstudiengänge).
513 Aber auch wenn ich ChatGPT frage oder diese typische
514 Quelle heranziehe, dann sind A13-Gehälter für
515 Akademiker so ziemlich im Mittelfeld. Das passt zu
516 meiner persönlichen Erfahrung. Aber wie gesagt, ich bin
517 da echt kein Experte und die Pensionsansprüche sind
518 natürlich dennoch üppig. Aber dafür sitzt man ja auch im
519 goldenen Käfig.

520 (I don't know, maybe I have a poor comparison, but I
521 have friends who earned more right after a Bachelor's
522 degree (engineering programs).

Dominik Hetjens. Generative Authority and the AI Reference Paradox

523 But when I ask ChatGPT or look at this typical source,
524 A13 salaries for academics are pretty much in the middle
525 of the pack. That matches my personal experience. But
526 as I said, I'm really no expert, and the pension
527 entitlements are obviously still lavish. But then again,
528 you're essentially sitting in a golden cage.)

529 Both the beginning and the end of the comment contain
530 several rather strong hedges: “Ja keine Ahnung, ich habe
531 vielleicht einen schlechten Vergleich” / “ich bin da echt kein
532 Experte.” The reference to ChatGPT and the proposition it
533 supports, however, do not contain any hedges; furthermore,
534 the evidential phrase utilised – the conditional/temporal
535 “wenn ich ChatGPT frage” (“if/when I ask ChatGPT”) –
536 arguably does not distance the speaker significantly from the
537 subsequent proposition. Instead, one could argue that the
538 “wenn” might even act as a booster, as it indicates the
539 reproducibility of the results.

540 The phrase “diese typische Quelle” contains a link to the
541 website of the *German Economic Institute*, a generally
542 accepted authority on economic matters. To further underpin
543 its credibility, the user refers to it as a “typische Quelle”,
544 highlighting its common use for salary comparisons. Both
545 ChatGPT and the *German Economic Institute* are presented
546 as yielding the same result, which is then characterised as
547 being in line with the user’s personal experience: “Das passt
548 zu meiner persönlichen Erfahrung.”

549 In doing so, the user employs all the strategies described in
550 Chapter 2 as typical for successful comments on the platform:
551 They introduce personal background, refer to external
552 sources, and present themselves as persuadable by using
553 concessive markers (the word “aber” alone appears four
554 times in this relatively short comment).

555 The only departure from the strategies described in
556 Chapter 2 is the introduction of ChatGPT, which appears
557 alongside an accredited source and is framed in a similar
558 manner. Both authorities are positioned within what could be
559 termed a “sandwich” of reported personal experience,
560 flanked by multiple hedges. Accompanying ChatGPT with a
561 traditionally credentialed source can again be seen as a sign
562 of meta-discursive awareness: while ChatGPT is deemed

Dominik Hetjens. Generative Authority and the AI Reference Paradox

563 worth referencing, certain measures are required to maintain
564 face when doing so.

565 The second comment occurs at L2 and constitutes a reply
566 to a comment proposing that Donald Trump is a dictator.

567 (8) Jedes Mal, wenn ich ChatGPT frage ob Trump ein
568 Diktator ist, wird das verneint. Diktatoren werden nicht
569 frei gewählt. Trump wurde frei gewählt und er hat bisher
570 auch nichts unternommen Wahlen abzuschaffen.

571 (Every time I ask ChatGPT if Trump is a dictator, it says
572 no. Dictators are not freely elected. Trump was freely
573 elected and, so far, he hasn't done anything to abolish
574 elections.)

575 Again, the evidential phrase uses “wenn” (“if/when”), but this
576 time, the possible “boosting” function is enhanced by “Jedes
577 Mal” (“every time”), suggesting that the judgment in question
578 is particularly credible, as the LLM produces the same result,
579 even when queried multiple times. The commenter does not
580 introduce any other evidence for the proposition, and it is
581 ambiguous whether the subsequent propositions are still
582 attributed to the AI.

583 *7.3.4 Type 4: Arbiter References*

584 While, at first glance, comments of this type will appear
585 similar to those of type 3, as both might be used to win a
586 debate, there are a number of key characteristics separating
587 the two. Comments of type four signal that the user is indeed
588 not persuadable, thereby going against the core strategy
589 described above, and use ChatGPT as a type of “final”
590 authority. Consequently, these comments do not contain any
591 concessions or hedges, and occur later in discussions and
592 further down the tree structure (L3 and L6). They are
593 characterised by a lack of information and arguments on the
594 actual topic of the debate, instead they directly assert that an
595 interlocutor’s position is incorrect based on the AI’s output.
596 An example for this is the following, in which a user defends
597 themselves against the accusation of “soapboxing”, a type of
598 dishonest debating named by a moderator as the reason one
599 of the user’s comments was deleted:

Dominik Hetjens. Generative Authority and the AI Reference Paradox

- 600 (9) Fragt Mal chatgpt ob das hier unter soapboxing fällt.
601 Nein einfach nein
- 602 (Just ask chatgpt whether this here counts as soapboxing.
603 No simply no)
- 604 Here, the imperative does not simply signify a request or
605 recommendation, but is used to reference an entity in favour
606 of the user’s proposition. This suggests that, in accordance
607 with the general availability and interactive nature of LLMs,
608 verbal phrases around a verb of communication in the
609 imperative can constitute evidentials. The intention to end a
610 debate appears even clearer in the following example:
- 611 (10) ich habe mal aus Spaß die KI gefragt, inwiefern deine
612 Aussagen stimmen: Es stimmt nichts! Die anderen in der
613 Kommentarspalte haben in der Debatte einfach
614 gewonnen :D Vllt solltest du ChatGPT mal verwenden,
615 um deine Aussagen zu prüfen und zu reflektieren.
- 616 (I just asked the AI for fun to what extent your
617 statements are true: Nothing is true! The others in the
618 comment column have simply won in the debate :D
619 Maybe you should use ChatGPT for once to check and
620 reflect on your statements.)
- 621 The start of the comment is in line with the prototypical
622 structure described before: It utilises the most common
623 evidential phrase “Ich habe mal gefragt” (“I just asked”), and
624 inserts “aus Spaß” (for fun) as a hedge to maintain face and
625 justify the use of AI. The rest of the comment, however,
626 departs sharply from the usual structure, and is inconsistent
627 with the casualness displayed by the evidential phrase. The
628 informative parts typical for type 2 comments are missing
629 completely, and the topic of the debate remains hidden.
630 Instead, the user directly claims that the debate has been won
631 in their own favour, and the other user is wrong on
632 everything: “Es stimmt nichts!” (“Nothing is true!”), as
633 purportedly proven by the judgment made by ChatGPT. If
634 the comment cited for type 1 constitutes one end of the
635 spectrum, this comment represents the other end. The topic
636 of the debate, the prompt used in interaction with the LLM,
637 the actual output – none of this is transparent in the

Dominik Hetjens. Generative Authority and the AI Reference Paradox

638 comment itself. Instead, absolute trust in the abilities of the
639 AI is expressed.

640 **8 Discussion**

641 The results of this study suggest that the integration of
642 conversational agents like ChatGPT has started to
643 significantly alter the nature of debates in interactional
644 online-spaces. Their conversational character as well as their
645 permanent availability for all partners in a debate has resulted
646 in a new type of evidential constructions, using verbs of
647 communication. By far the most frequently used phrase is
648 “Ich habe mal ChatGPT gefragt”. Not only does this
649 constitute an anthropomorphism in line with findings from
650 other studies, it also holds further implications for pragmatics
651 and the investigation of automation bias:

652 In the comments analysed above, anthropomorphisation
653 can be seen as an evidential and epistemic strategy: It
654 references the source of information and creates a distance
655 between the speaker and the generated output. By doing so,
656 users signal that while they are utilising AI, they are not
657 “blindly” trusting it. This allows the speaker to benefit from
658 the AI’s authority while maintaining a “buffer” of deniability
659 should the AI’s output be proven wrong or they be accused of
660 inadequate use of AI. For this purpose, the modal particle
661 “mal” as well as a number of (other) hedges are used, hinting
662 at users’ meta-discursive awareness. This meta-discourse is
663 highly visible in the corpus, as individual comments
664 sometimes spark a debate on the validity of AI-use. To pre-
665 empt this criticism, commenters often present the AI-
666 interaction as a kind of casual experiment they undertook out
667 of curiosity (“aus Neugierde”) or for fun (“aus Spaß”).

668 However, the practice of referencing ChatGPT itself
669 constitutes a display of trust in its abilities, and where output
670 is displayed or paraphrased, it is usually presented as factual.
671 This leads to what might be called the “AI-reference
672 paradox”: several comments exhibit an awareness of the
673 potential problems of AI usage and a high degree of trust in
674 AI-generated content at the same time.

675 While this holds true for the majority of comments, it is
676 particularly apparent in one of the argumentative comments

Dominik Hetjens. Generative Authority and the AI Reference Paradox

677 that employ ChatGPT as evidence to support a user's position
678 in a debate (Type 3). Both comments of this type utilise a
679 conditional/temporal construction within the evidential
680 phrase: “Wenn ich ChatGPT frage” (“when/if I ask
681 ChatGPT”). Notably, this construction is absent from all other
682 categories. This suggests a tendency to highlight the iterative
683 nature of one’s inquiry and the reproducibility of the results
684 to strengthen a position in a debate. Consequently, “wenn”
685 (“when” or “if”) and “jedes Mal, wenn” (“each time when”)
686 may qualify as boosters when combined with an evidential
687 phrase involving “fragen” (“to ask”) in the (purported)
688 reconstruction of human–AI interaction.

689 This indicates that users do not only use anthropomorphic
690 language to denote the source of information, but have also
691 developed nuanced ways to frame their interaction with the
692 LLM, according to the function of their comment in the
693 debate, as well as the degree of credibility ascribed to the AI
694 for that purpose. While ChatGPT is treated similarly to
695 traditionally accredited authorities to some degree, both the
696 use of conditional/temporal constructions and the imperative
697 clearly show that the interactive nature of the source has
698 given rise to new strategies of framing references that are
699 specific to AI: The credibility of a comment referencing
700 conversational agents depends in part on the framing of
701 supposed past interactions with them.

702 Their inclusion in online debates therefore creates a new
703 type of communication characterised by AI–human–human–
704 AI interaction, in which all interlocutors constantly envision a
705 potential separate conversation other interlocutors might
706 have with a third party, the AI. On the linguistic surface, this
707 is evident in the imperatives used. In some of the
708 argumentative comments observed, their application might
709 be summed up as “Ask ChatGPT, you’ll see that I’m right”,
710 making it a kind of “generative authority”, rendering the
711 communication AI-mediated in an indirect way. In the most
712 extreme cases, these references are used with the intent to
713 end a conversation in favour of the commenter, using the AI
714 as a type of “final arbiter” (type 4). As a result, the topic of the
715 debate becomes opaque in the utterance itself.

716 On a methodological level, the analysis has yielded a
717 number of interesting results. Examining references to AI in
718 regard to the epistemics and evidentials used has proven

Dominik Hetjens. Generative Authority and the AI Reference Paradox

719 fruitful. For a small dataset like this one, not all extra-
720 linguistic metrics typical to the platform turned out to be of
721 equal value. Reddit’s signature “karma score” did not offer
722 much insight beyond the interpretation of individual
723 comments, as the most common score was 1, and no clear
724 patterns for higher or lower scores could be identified. This is
725 most likely caused by the small number of comments in the
726 final corpus, and future studies utilising a larger corpus will
727 very likely deliver promising results in this regard as well.

728 However, taking the level in the tree structure into account
729 has shown relevant results: comments that occur on L1, close
730 to the original post and most visible, have a tendency to be
731 informative, suggesting that new information is introduced
732 early in a discussion. Argumentative comments (Types 3 and
733 4) rely on a debate already taking place, and therefore occur
734 later in the tree structure.

735 While this study has been able to uncover some of the
736 *mechanisms* and *strategies* relevant to AI-references in
737 certain online debates, it can say little about the frequency
738 with which they are applied. Future studies using larger
739 corpora are required to investigate the prevalence of the
740 phenomena in question, and shifting trends in usage.

741 9 Summary

742 This explorative study has investigated the linguistic
743 integration of a conversational agent based on an LLM into
744 the interactional landscape of the German-language Reddit-
745 community r/de. The findings suggest that ChatGPT and
746 similar tools are no longer merely external technologies but
747 are framed as interactive authorities. The qualitative
748 functional analysis yielded a four-part typology based on the
749 role ascribed to the AI – Scribe, Informant, Corroborator, and
750 Arbiter – which illustrates a spectrum ranging from
751 informative text generation to argumentative moves. In the
752 most extreme case, AI is weaponised to abruptly terminate
753 debates by proclaiming a type of “higher truth”.

754 Linguistically, the study uncovered a distinct pattern of
755 anthropomorphised reportative evidentiality. Users
756 consistently employ verbs of communication, such as
757 “fragen” (“to ask”), paired with the modal particle “mal” to

Dominik Hetjens. Generative Authority and the AI Reference Paradox

758 frame their AI interactions. This strategy serves a dual
759 pragmatic purpose: It ascribes a high level of agency and
760 authority to the AI and its output while providing the user
761 with a “buffer of deniability” via hedges like “aus Spaß” or
762 “aus Neugierde”. This “AI-reference paradox” – the
763 simultaneous display of meta-discursive skepticism and
764 deep-seated trust in AI-generated factual claims – emerges as
765 a defining characteristic of the interactions observed.
766 Ultimately, as LLMs become ubiquitous, the traditional
767 markers of expertise on platforms like Reddit are being
768 supplemented, and in some cases supplanted, by the ability to
769 strategically “interrogate” generative systems and frame AI-
770 references in order to win arguments and manage face.

771 This creates an AI–human–human–AI interaction where
772 every interlocutor must account for possible separate, private
773 conversations their peers might be having with a
774 conversational agent, while at the same time framing their
775 own interactions with it in the best possible way. Future
776 research is necessary to expand the empirical basis to include
777 larger, longitudinal datasets and references to other AI agents,
778 allowing for an analysis of how frequently these pragmatic
779 strategies are used and how they evolve as LLM-based
780 conversational agents become even more deeply embedded
781 in digital life.

782 References

- 783 Blombach, Andreas/Dykes, Natalie/Heinrich,
784 Philipp/Kabashi, Besim/Proisl, Thomas (2020): A corpus of
785 German Reddit exchanges (GeRedE). In: *Proceedings of the*
786 *Twelfth International Conference on Language Resources*
787 *and Evaluation (LREC 2020)*. Marseille: European
788 Language Resources Association (ELRA), 6310–6316.
789 URL: <https://aclanthology.org/2020.lrec-1.774.pdf>
790 Brown, Tom B./Mann, Benjamin/Ryder, Nick/Subbiah,
791 Melanie/Kaplan, Jared/Dhariwal, Prafulla/Neelakantan,
792 Arvind/Shyam, Pranav/Sastry, Girish/Askell,
793 Amanda/Agarwal, Sandhini/Herbert-Voss, Ariel/Krueger,
794 Gretchen/Henighan, Tom/Child, Rewon/Aditya,
795 Ramesh/Ziegler, Daniel M./Wu, Jeffrey/Winter,
796 Clemens/Hesse, Christopher/Chen, Mark/Sigler,

Dominik Hetjens. Generative Authority and the AI Reference Paradox

- 797 Eric/Litwin, Mateusz/Gray, Scott/Chess, Benjamin/Clark,
798 Jack/Berner, Christopher/McCandlish, Sam/Radford,
799 Alec/Sutskever, Ilya/Amodei, Dario (2020): Language
800 Models are Few-Shot Learners. In: *Advances in Neural*
801 *Information Processing Systems* 33, 1877–1901.
802 DOI: <https://doi.org/10.48550/arXiv.2005.14165>
- 803 Chen, Hui/Hara, Noriko/McKay, Clinton (2021): Investigating
804 mediated public engagement with science on the “science”
805 subreddit: From the participants’ perspective. In: *PLOS*
806 *ONE* 16 (4), e0249181.
807 DOI: <https://doi.org/10.1371/journal.pone.0249181>
- 808 Diewald, Gabriele/Smirnova, Elena (2010): *Evidentiality in*
809 *German. Linguistic Realization and Regularities in*
810 *Grammaticalization*. Berlin/New York: Mouton de
811 Gruyter.
- 812 Ernst & Young (2025): *EY AI Sentiment Index 2025*. EY
813 Newsroom.
814 URL: <https://www.ey.com/content/dam/ey-unified-site/ey-com/de-de/newsroom/2025/05/ey-presentation-ai-sentiment-index-2025.pdf>
- 817 Felder, Ekkehard/Kückelhaus, Marcel (2025): Das
818 definierende Sprachmodell (LLM):
819 Anthropomorphisierung in der Mensch-Maschine-
820 Interaktion. In: *Zeitschrift für Literaturwissenschaft und*
821 *Linguistik* 55, 431–448.
822 DOI: <https://doi.org/10.1007/s41244-025-00380-7>
- 823 Kückelhaus, Marcel (in prep.): *Menschliche Künstliche*
824 *Intelligenz. Eine diskurslinguistische Analyse zur*
825 *sprachlichen Konstituierung von KI*. Heidelberg: HeiDocs.
- 826 Hetjens, Dominik (submitted): Sources From Below – Source
827 Work From Below: Collaborative Approaches, AI and
828 Communities of Pseudonymous Linguistic Expertise. In:
829 Auer, Anita/Schiegg, Markus (Hg.): *West Germanic*
830 *Language Histories “from below”: New data Sources and*
831 *Approaches*. Amsterdam: John Benjamins.
- 832 Leavitt, Alex (2015): “This is a throwaway account”:
833 Temporary technical identities and perceptions of
834 anonymity in a massive online community. In: Cosley,
835 Dan/Forte, Andrea/Ciolfi, Luigina/McDonald, David (Hg.):
836 *Proceedings of the 18th ACM Conference on Computer*
837 *Supported Cooperative Work & Social Computing (CSCW*
838 *2015)*. New York: ACM, 317–327.

Dominik Hetjens. Generative Authority and the AI Reference Paradox

- 839 DOI: <https://doi.org/10.1145/2675133.2675175>
840 Lu, Rickey (2025): Understanding online advice-giving
841 evaluations through the politeness evaluation model. In:
842 *Journal of Politeness Research* 21 (1), 249–287.
843 DOI: <https://doi.org/10.1515/pr-2023-0068>
844 Massanari, Adrienne L. (2015): *Participatory culture,*
845 *community, and play: Learning from Reddit*. New York:
846 Peter Lang.
847 Matias, J. Nathan (2019): The civic labor of volunteer
848 moderators online. In: *Social Media + Society* 5 (2), 1–12.
849 DOI: <https://doi.org/10.1177/2056305119836778>
850 Medvedev, Alexey N./Lambiotte, Renaud/Delvenne, Jean-
851 Charles (2019): The anatomy of Reddit: An overview of
852 academic research. In: Ghanbarnejad, Fakhteh/Roy,
853 Risharaj Saha/Karimi, Fariba/Delvenne, Jean-
854 Charles/Mitra, Bivas (Hg.): *Dynamics on and of Complex*
855 *Networks III*. Cham: Springer, 183–204
856 DOI: https://doi.org/10.1007/978-3-030-14683-2_9
857 Messerli, Thomas C./Dayter, Daria/Leuckert, Sven/Liimatta,
858 Aatu/Mahler, Hanna/Bohmann, Axel/Kozma,
859 Gustavo/Tosin, Rafaela (2025): Digital debating cultures:
860 Communicative practices on Reddit. In: *Digital*
861 *Scholarship in the Humanities* 40 (1), 227–240.
862 DOI: <https://doi.org/10.1093/llc/fqaf005>
863 Mortelmans, Tanja (2000): On the 'evidential' nature of the
864 'epistemic' use of the German modals müssen and sollen.
865 In: *Journal of Pragmatics* 32 (2), 131–148.
866 Reddit (2026): *Subreddit r/de*.
867 URL: <https://reddit.com/r/de> (Accessed: 19 March 2026).
868 Rudolf von Rohr, Marie-Thérèse/Thurnherr,
869 Franziska/Locher, Miriam A. (2018): Linguistic expert
870 creation in online health practices. In: Bou-Franch,
871 Patricia/Garcés-Conejos Blitvich, Pilar (Hg.): *Analyzing*
872 *Digital Discourse: New Insights and Future Directions*.
873 Cham: Springer, 219–250.
874 DOI: https://doi.org/10.1007/978-3-319-92663-6_8
875 Similarweb (2026): *Top Websites Ranking – Most Visited*
876 *Websites in the World*.
877 URL: <https://www.similarweb.com/top-websites/>
878 (Accessed: 19 March 2026).

Dominik Hetjens. Generative Authority and the AI Reference Paradox

- 879 van der Nagel, Emily (2017): From usernames to profiles: The
880 development of pseudonymity in Internet communication.
881 In: *Internet Histories* 1 (4), 312–331.
882 DOI: <https://doi.org/10.1080/2080/24701475.2017.1389548>
883 Wilton, Antje (2019): Epistemic Stance Markers in German
884 and English as a Lingua Franca Media Sports Interviews.
885 In: *Alicante Journal of English Studies* 32, 217–243.
886 DOI: <https://doi.org/10.14198/raei.2019.32.09>