

German anaphoric demonstrative pronouns shape discourse expectations about contrast and animacy

The anaphoric use of the two German demonstrative pronouns (*der* and *dieser*) differs from the use of the personal pronoun *er* in that they signal topic change, but the difference between the two demonstratives is still an open research question. We present results from a corpus study showing that the demonstratives interact with and shape discourse expectations downstream and upstream in two previously unexplored ways. First, we find an interaction between antecedent animacy and presence of a competitor: *dieser* prefers preceding contexts with potential competing antecedents more than *der* and *er*, but without competitors, *dieser* manifests a preference for inanimate antecedents *der* and *er* do not have. Second, antecedents of *er* persist longer in following discourse than those of the demonstratives, but competitors to antecedents for *dieser* persist longer than those for *der* and *er*.

German has two demonstrative pronouns, *der* (we use only the masculine form since it distinguishes case) and *dieser*, both of which behave similarly in their anaphoric use when compared to the personal pronoun *er* in contexts with more than one antecedent. Research has extensively compared the demonstrative *der* to the personal pronoun, but fewer studies focus on differences between the two demonstratives (Fuchs & Schumacher, 2020; Patil et al., 2020, 2023; Buchholz & von Heusinger, 2024a, b). We assume that *dieser* expresses a contrast similar to the anti-uniqueness effect of demonstrative DPs (Çokal and von Heusinger, 2024), and there are indications that antecedent animacy plays an important role (Bader et al., 2023). Regarding reference persistence, it is known that both *er* and the demonstrative *der* (at least in certain genres, cf. Repp 2024) can form longer topic chains, while *dieser* cannot occur twice in direct succession in a coreferential chain (*dieser...dieser*). There is also a tentative finding that referents picked up by *dieser* might persist longer (Fuchs & Schumacher, 2020).

In order to explore these experimental findings in naturally occurring written discourse, we investigated the two demonstrative pronouns and the personal pronoun with 8 predicates each, for a total of 24 pronoun-predicate combinations (*er/der/dieser* – *hat/ist/hatte/war/hätte/wäre/müsste/könnte* ‘has’/‘is’/‘had’/‘was’/‘had’ (conjunctive)/ ‘would have’/‘must’(conjunctive)/‘could’). We created a corpus of $24 \times 50 = 1200$ randomly selected short texts (5-7 sentences) centered around sentences with the pronoun in subject position followed by the predicate, with 2 sentences as preceding and 3-5 sentences as continuation context, sourced from German newspaper texts published in Germany after 1950 in the *W* archive of DeReKo (Deutsches Referenzkorpus DeReKo 2024-1). (1) gives a representative example for a text matching the criteria of the search plus further manual selection, e.g. the exclusion of pronoun occurrence in direct speech.

- (1) Bayerns Wirtschaftsminister₁ kritisierte am Morgen Äußerungen seines Rivalen Seehofer₂. Dieser₂ hatte sich₂ darüber beschwert, Huber₁ und Bayerns Innenminister Günther Beckstein₃ (CSU), [...], würden die Auseinandersetzung um den CSU-Vorsitz „über die Medien spielen“.
Bavaria's Minister of Economic Affairs₁ criticised comments made by his rival Seehofer₂ this morning. The latter₂ [= *dieser*] had complained that Huber₁ and Bavaria's Interior Minister Günther Beckstein₃ (CSU), [...], were ‘playing the media’ in the dispute over the CSU chairmanship. (SOL07/JAN.02462 Spiegel-Online, 25.01.2007; Stoiber dementiert Rücktritt vom Rücktritt)

The texts were manually annotated by our student assistants for a number of features, including presence of a competitor in the preceding context and tracking of it in the following context. Antecedent referent and potential competitors were annotated for animacy on a four-part scale: human individual, animate collective > concrete inanimate > abstract inanimate.

We compared whether a competitor was present in the texts with *er*, *der* and *dieser* and crossed

this with the animacy of the antecedent referent. As Figure 1 shows, competitors were more frequently present with *dieser* (77% of all cases) than with *der* (66%) and *er* (58%). At the same time, the ratio of inanimate antecedents for *dieser* increased from contexts with competitor to those without competitor (42% vs. 75%), while this wasn't the case as much for *der* (40% vs. 55%) and not at all for *er* (12% vs. 8%). This suggests that there is a tripartite distinction between the three pronoun types along these dimensions: *dieser* cues a contrast and is thus more frequently used in the presence of a competitor, but without competitor, an additional preference for inanimates emerges (while *er* and *der* are known to differ along further dimensions). *Dieser* thus shows traces of a tendency to track inanimate referents (Himmelmann, 1996). We also compared the number of times the antecedent referent and the competitor referent (if present) were taken up again in the subsequent context. Results indicate that while referents taken up by *er* (mean number of reuptakes=2.32, sd=2.18) persist longer on average than those taken up by the demonstratives (*der*: m=0.78, sd=1.27; *dieser*: m=0.74, sd=1.35), the competitors (when present) to referents taken up by *dieser* also seem to persist somewhat longer (mean number of reuptakes=0.84, sd=1.44) than for the other two pronouns (*der*: m=0.64, sd=1.14; *er*: m=0.56, sd=1.24). This is still a very tentative result. We currently extend our study via an LLM annotation, using our human annotation as benchmark. By thus scaling up the amount of data analyzed by more than one order of magnitude, we aim to more reliably explore subtle effects of pronoun choice on discourse expectations like these.

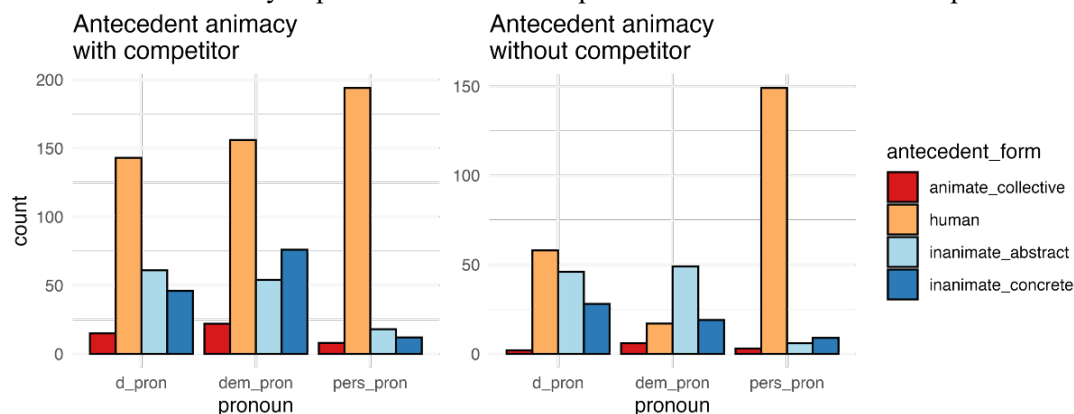


Figure 1: Antecedent animacy per pronoun type (*der*=d_pron, *dieser*=dem_pron, *er*=pers_pron), separated by contexts with a competitor (left, $N_{der} = 265$, $N_{dieser} = 308$, $N_{er} = 232$, $N_{total} = 805$) and without a competitor (right, $N_{der} = 134$, $N_{dieser} = 91$, $N_{er} = 167$, $N_{total} = 392$).

References

- Bader, Markus, Jacopo Torregrossa & Esther Rinke. 2023. Pinning down the interaction between animacy and syntactic function in the interpretation of German and Italian personal and demonstrative pronouns. *Discourse Processes* 60. 655–673. doi:<https://doi.org/10.1080/0163853X.2023.2252699>.
- Buchholz, Timo & Klaus von Heusinger. 2024a. German demonstrative pronouns differ in their sensitivity to discourse and sentence topics. *Frontiers in Communication* 9:1369290. doi: <https://doi.org/10.3389/fcomm.2024.1369290>.
- Buchholz, Timo, & Klaus von Heusinger. 2024b. German demonstratives and topic questions. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46(0), 5143–5150.
- Çokal, Derya & Klaus von Heusinger. 2024. German demonstrative pronouns in contrast. *Dialogue and Discourse* 15(1). 45–76. doi:<https://doi.org/10.5210/dad.2024.102>.
- Fuchs, Melanie & Petra B. Schumacher. 2020. Referential shift potential of demonstrative pronouns – evidence from text continuation. In Ashild Næss, Anna Margetts & Yvonne Treis (eds.), *Demonstratives in discourse*, 185–213. Berlin: Language Science Press.
- Himmelmann, Nikolaus. 1996. Demonstratives in Narrative Discourse: A Taxonomy of Universal Uses. In Barbara Fox (Ed.), *Studies in Anaphora*, 205–254. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.33.08him>
- Patil, Umesh., Bosch, Peter, & Stefan Hinterwimmer. 2020. Constraints on German diese demonstratives: Language formality and subject-avoidance. *Glossa*, 5(1). <https://doi.org/10.5334/gjgl.962>
- Patil, Umesh. Hinterwimmer, Stefan, & Petra B. Schumacher. 2023. Effect of evaluative expressions on two types of demonstrative pronouns in German. *Glossa*, 8(1), Article 1. <https://doi.org/10.16995/glossa.9577>
- Repp, Magdalena 2024. *D-pronouns in narrative texts: Examining referential behavior and neural processing* [PhD Thesis, Universität zu Köln]. <https://d-nb.info/1345151667/34>