frontiers | Frontiers in Language Sciences

TYPE Brief Research Report
PUBLISHED 21 August 2024
DOI 10.3389/flang.2024.1433482

OPEN ACCESS

EDITED BY
Moreno I. Coco,
Sapienza University of Rome, Italy

REVIEWED BY
Paula Rubio-Fernández,
University of Oslo, Norway
Óscar Loureda,
Heidelberg University, Germany

*CORRESPONDENCE
Derya Çokal
dcokal@uni-koeln.de

RECEIVED 15 May 2024
ACCEPTED 25 July 2024
PUBLISHED 21 August 2024

CITATION
Çokal D and von Heusinger K (2024) The role of alternatives in the cognitive processing of German demonstratives: insights from online and offline processing.
Front. Lang. Sci. 3:1433482.
doi: 10.3389/flang.2024.1433482

The role of alternatives in the cognitive processing of German demonstratives: insights from online and offline processing

Derya Çokal * and Klaus von Heusinger

Institute for German Language and Literature, University of Cologne, Köln, Germany

This study, employing eye-tracking reading and sentence completion experiments, explores the impact of competing antecedents on the German demonstratives *der* and *dieser*. It challenges prior assumptions, revealing that in competitive alternative antecedent contexts, processing *dieser* initially posed challenges, indicating sensitivity to alternatives. *Dieser* exhibited less processing difficulties than *der,* potentially influenced by a register effect. Consistent with previous findings, in the offline task, references to the non-prominent entity were similar for both demonstratives, but our online experiment shows functional differences in cognitive processes between the two in reading. Our results suggest that *Thematic Role* accounts better explain antecedent preferences for *der* and *dieser* than *Centering Theory*.

KEYWORDS

demonstratives, anaphora, online-processing, competing antecedents, discourse processing

1 Introduction

Referential expressions serve as "procedural instructions"—pertaining to locating referents in memory (Cornish, 2008)—to construct/modify mental models in an unfolding discourse within the minds of readers/speakers. Such continuous updates involve competing alternative antecedents (i.e., local foci) and establish topic/foci (cf. Strube and Hahn, 1999; Poesio et al., 2004; Karamanis et al., 2009). In S4 below, "an inline skater" and "an elderly person" are competing alternative antecedents for the German demonstratives *der*/*dieser* (he).

(1) S1. During my round through the park, I saw **an inline skater** (CFs) and **an elderly person** (CFs). S2. **The inline skater** (CP = topic establishment) was desperately trying to brake on the slippery road. S3. Today, after many rain showers, it was very muddy. S4. **The elderly person** (CP = topic establishment) was narrowly avoided by the inline skater. *Der*/*Dieser* (He).

Anaphors exhibit varying degrees of sensitivity to factors such as salience and word order (Brown-Schmidt et al., 2005; Kaiser and Trueswell, 2008; Çokal et al., 2016). Demonstratives can differ in their ability to evoke alternatives regarding the index argument and are preferred when one or two alternative antecedents have been introduced in the previous discourse (Saha et al., 2023; Buchholz and von Heusinger, 2024). Demonstratives—specifically *dieser*—serve to "single out" one referent from multiple competing antecedents (Ahrenholz, 2007). However, two main questions have not yet been examined: (1) How does the competition between multiple antecedent options affect the processing of *der*/*dieser* in online reading? (2) Is *dieser* more sensitive to local competing antecedents than is *der*?

These two questions lead to a comparison between two theories: (3) Are the tenets of Centering Theory (CT; i.e., subject > object) or Thematic Role (TR; i.e., agent > patient) used in the reference resolution process of *der*/*dieser* when multiple antecedents exist in the local context? These inquiries, specifically regarding the comparison of two German demonstratives (i.e., *der*/*dieser*), have not received sufficient attention. The present study aims to bridge this gap by offering insights into the cognitive processes associated with *der*/*dieser* in both competing and non-competing alternative antecedent contexts in reading and sentence production experiments (cf. Peeters et al., 2021 for a review of demonstratives and the lack of reading studies on demonstratives).

German includes various types of demonstratives that can be used anaphorically. The most prevalent ones are the demonstratives from the *der* paradigm (i.e., *der*, *die, das*—also known as "d-pronouns") and those from the *dieser* paradigm (i.e., *dieser*, *diese*, *dieses*—"dem-pronouns"). In the current study, we examine both *der* ("that one") and *dieser* ("this one"). A few studies found nuanced differences between *der*/*dieser*, including: *der* avoids reference to topics (Bosch and Umbach, 2007), while *dieser* can refer to entities at the beginning of sentences, with a preference for object reference (Patil et al., 2020). *Dieser* refers to the most recently mentioned entity references (e.g., "inline skater"; Ahrenholz, 2007) and functions as inducing topic persistence (Cokal and von Heusinger, 2024). Other offline studies indicate that there are no clear functional differences in their anaphoric use (e.g., Fuchs and Schumacher, 2020). However, compared to *er/he*, references to the non-agent with *der* were preferred (Schumacher et al., 2017). Further investigations reveal that both *der* and *dieser* refer to the second-mentioned referent (i.e., less-prominent entity and/or recently-mentioned: Fuchs and Schumacher, 2020). Contrary to convention, with an object experiencer verb, majority of references with *der* and *dieser* were to the subject referent. On the other hand, the antecedent preferences changed with subject-experiencer verbs (Bader et al., 2022). However, this claim has not been further tested through online-reading and sentence production experiments focusing on two demonstratives.

Two main approaches—Centering Theory (CT) and Thematic Role (TR) accounts—are proposed to explain which entity will be prominent to be referred by anaphora (i.e., specifically for pronouns) at any time in discourse. According to CT, the local focus includes a set of forward-looking centers (CFs). Consider the following example, repeated below: "During my round through the park, I saw **an inline skater** and **an elderly person.**" Both "an inline skater" and "an elderly person" are CFs: Some CFs acquire particular prominence, referred to as Preferred Center (CP), which corresponds to the concept of "topic" (Chafe, 1976). The "inline skater" in S2 and "elderly person" in S4 are CPs because they are both in the subject position [see example (1)]. According to CT, the subject position (i.e., grammatical role) is the most prominent one.

Contrary to CT, TR accounts propose that an agent of the action is more prominent than a non-agent (Stevenson et al., 1994; Schumacher et al., 2017). In the following sentence, prominence establishment is reversed: "The elderly person was narrowly avoided by the inline skater." The "inline skater" is prominent rather than the "elderly person" because the skater is the agent of the action (i.e., the one who avoids hitting the elderly person). Notably, both CT and TR accounts employ distinct strategies for assigning prominence in anaphora resolution. Based on previous findings on demonstratives, if less prominence is a factor in the processing and production of *der*/*dieser*, then in CT, demonstratives would avoid subject references and a reference to the "inline skater" would be preferred. However, in TR accounts, referring to the agent (i.e., inline skater) would be avoided, and "elderly person" would be preferred for demonstratives.

The current study aims to deepen our understanding of the cognitive processes that *der*/*dieser* signal in written text, as well as German speakers' production and comprehension of these expressions. Specifically, we examine whether the use of *der*/*dieser* depends on the presence of alternative competitive or no-competitive antecedents. In addition, we explored whether participants' preferences in an online reading experiment were the same as in an offline production experiment. To investigate these, we conducted an eye-tracking reading experiment to index German readers' use of information during comprehension. In addition, we ran a sentence- completion experiment to explore whether writers employ thematic role or grammatical role assignment in anchoring antecedents of these expressions.

2 Experiment 1

We designed a 2 x 2 within subject experiment, crossing two levels of anaphora (*dieser* vs. *der*) and two levels of competitor type in the previous discourse (competitor vs. no- competitor).

Context: S1. Bei meiner Runde durch die Parkanlage habe ich einen Inlineskater und einen Rentner gesehen. **S2**. Der Inlineskater versuchte verzweifelt auf der rutschigen Straße zu bremsen. **S3**. Heute war es nach vielen Regenschauern sehr matschig.

Competitor conditions S4. (2a and 2b): Der Rentner wurde deswegen von dem Inlineskater wirklich nur sehr knapp umfahren. **S5. (2a)** ***Dieser*****/(2b)** ***Der*** fiel auf seinen Knieschoner und fluchte lautstark.

No-competitor conditions S4. (2c and 2d): Einige Hecken wurden deswegen von dem Inlineskater wirklich nur sehr knapp umfahren **S5. (2c)** ***Dieser*****/(2d)** ***Der*** fiel auf seinen Knieschoner und fluchte lautstark.

Context: S1. During my round through the park, I saw an inline skater and an elderly person. **S2**. The inline skater was desperately trying to brake on the slippery road. **S3**. Today, after many rain showers, it was very muddy.

Competitor conditions S4. (2a and 2b): The elderly person was, therefore, narrowly avoided by the inline skater. **S5. (2a)** ***Dieser (this one)*****/(2b)** ***Der*** **(*****that one*****)** fell on his knee pad and cursed loudly.

No-competitor conditions S4. (2c and 2d): Some hedges were, therefore, narrowly avoided by the inline skater. **S5. (2c)** ***Dieser (this one)*****/(2d)** ***Der*** **(*****that one*****)** fell on his knee pad and cursed loudly.

In all conditions, S1 introduces two characters (e.g., an inline skater and an elderly person). S2 focuses on one of the characters (i.e., the inline skater). While **in the competitor context,** the critical sentence (S4) has two arguments (the inline skater vs. the elderly person), **in the no-competitor context**, S4 has only one argument

(the inline skater). In both conditions, S4 has a passive construction to balance the prominence of two arguments (the elderly person = the inline skater). While in the English translation above, the "inline skater" in S4 comes just before *der/dieser*, that is not the case in German stimuli, since we use an adverbial phrase and the participle between potential antecedents and demonstratives. Our manipulations— adverbial phrase, participle, and passive construction— are unlike those in previous studies, which reported last- mentioned entity preferences for demonstratives, where the distance between demonstratives and their recent antecedents was kept short (e.g., Fuchs and Schumacher, 2020). S5 began with a demonstrative (*dieser* or *der*) and referents of *der/dieser* were disambiguated with an object (i.e., a knee pad) more likely to be used by one of the characters (e.g., a skater would more likely have a knee pad rather than an elderly person). The two competing alternatives were always congruent in gender and number with the demonstratives. We controlled having no matching third antecedent in a previous context.

In the eye-tracking reading experiment, we predicted that: In the anaphora region, any significant difference between *der* and *dieser* would be due to the length difference between these two words. However, if, in the anaphora region, *dieser* in the competitor context in [2a] leads to processing difficulty, high odds ratios in regressions-out, or longer total time than *dieser* in a non-competitor context, then *dieser* is sensitive to competing antecedents (i.e., subject non-agent vs. non-subject agent). The crucial prediction is for an interaction in the disambiguating region, which would demonstrate that the two referring expressions differ in their antecedent preferences. Schumacher et al. (2017) predict that demonstrative pronouns are preferred for only non-agent references (i.e., the elderly person). *Dieser* has two additional restrictions: (1) reference to locally accessible antecedents; (2) references to the most recent entity (i.e., the inline skater). Therefore, we predicted that in the no-competitor condition, *dieser* would access only one local referent, and thus it would lead to low odds ratios in regression-out and less total time. However, in the competitor condition, there would be a conflict between the two locally available antecedents (non- agent/elderly person vs. most recent/inline skater). Even if there were references to the most recently mentioned entity (i.e., the inline skater), suppressing the alternative antecedents would take time. In both conditions, *der* in the disambiguation region, however, accesses to the globally available non-agent (i.e., elderly person), leading to higher odd ratios in regression-out and longer total time (i.e., reading difficulty) than *dieser*. According to CT, references to the "inline skater" with *der*/*dieser* would lead to less processing (i.e., an avoidance of subject). On the other hand, according to TR accounts, there would processing difficulty when *der*/*dieser* referred to the "inline skater" (i.e., an avoidance of agent).

2.1 Methods

2.1.1 Participants

Fifty-two paid native German-speakers from the University of Cologne participated in the experiment (ages 21–24, M = 22; SD = 1.126). All were unaware of the study's purpose.

2.1.2 Apparatus

We used an Eyelink 1000 eye-tracker (SR Research Ltd, Canada) in tower-mounted mode, with a chin rest to stabilize each participant's head.

2.1.3 Materials

Forty items were created based on Example 2 above. Each item appeared in the four conditions, crossing competitor type (competitor and no-competitor) with anaphora type (*der* and *dieser*). The 40 stimuli were distributed into four lists, following a Latin Square procedure. In all four lists, each item appeared in only one condition and each condition appeared an equal number of times. There were 73 fillers and three practice items, all of which were similar in length to the experimental sentences.

2.1.4 Pre-testing the stimuli

We ran the acceptability judgement test on Qualtrics with 62 native speakers of German. We asked participants to rate their acceptability on a scale from 1 to 5. Each participant saw only one condition for each item. The conditions in the current study are grammatically acceptable but participants preferred the use of *dieser* to *der* (No-competitor: *der*: $M = 3.22$, *dieser*: $M = 3.35$; competitor: *der*: $M = 2.95$, *dieser*: $M = 3.15$). We also tested participants' object preferences for individuals. We presented the initial text (e.g., The elderly person was, therefore, narrowly avoided by the inline skater. Der fell on his knee pad) and asked participants to choose "Who would use the object [e.g., knee pad]?"

(a) Inline skater
(b) Elderly person

Eighty-percentage of cases were the non-subject agent of action (e.g., skater) and 20% was the subject non-agent entity (e.g., elderly person). The participants' object preferences were in line with our manipulation.

2.1.5 Procedures

We presented the 116 texts in Times New Roman 18 font, in fixed random order, with no experimental items adjacent. Comprehension questions never probed the referents of der/dieser (please see Supplementary Section 1 for details).

2.1.6 Data analysis

Before conducting data analysis, blinks from trials were removed using Eyedoctor (developed by UMASS Eye-tracking lab). For data analysis, we used "measures.pl" written by Patrick Sturt and extended by Amit Dubey to allow use of parameter file. Texts were divided into three regions (i.e., anaphora, verb, and disambiguation). Twenty-one percent of the excluded data points in the pronoun region involve instances where the pronoun was skipped (see Rayner et al., 2011 for the skipping rate of functional words). Data points for each measure across regions are as follow: regression-out: pronoun = 1,302; verb = 1,644; disambiguation = 2,000; total times: pronoun = 2,077; verb =

2,077; and disambiguation = 2,077 (see Supplementary Section 1.1 for data points in each condition). All participants scored at least 90% correct on comprehension questions.

We report results for regression-out (i.e., the proportion of trials where readers looked back from the region to an earlier piece of the text between the time when the region was first entered from the left to the time when the region was exited to the right). Lexical semantic information is processed during regression-out and used in recovery from processing difficulty from the previous text where the current word/text does not meet readers' expectations. For completeness, we also report total time (i.e., the sum of all fixations in the region) as a general measure of processing, even though this does not provide information about initial processing. In cases where the region received no fixations (total time), the trial was treated as missing data and excluded from analysis. We employed various packages: lme4 for logistic mixed effects regression (GLMER) models on regression-out, Sjplot for odds ratios and random effects calculation, emmeans library for standard errors and confidence intervals, and ggplot2 for visualizing proportion estimates by condition. The analysis focused on whether participants exhibited regression-out in the region, retaining trials without such regressions in the analysis (coded as 0). Since the data were categorical, logistic mixed effects regression models with random slopes and intercepts were run. For each region and total time measure, linear mixed effects regression (LMER) was constructed, incorporating all fixed effects and interactions in a single step. An additional package (plyr) was used to compute mean values.

For both linear and logistic mixed effects models, factor labels were transformed into numerical values, and centered prior to analysis, and we used binary contrast for both factors. All analyses reported below incorporated crossed random intercepts for participants and items. Random slope parameters (anaphora type), corresponding to the two experimental factors (competitor and no-competitor) and their interactions in the slopes (anaphor type * context type +1| subject), were included in the maximal model for both participants and items (Barr et al., 2013; Bates et al., 2015). To aid convergence, and to avoid spurious over-estimates of correlations, random correlation parameters were excluded from the model. The resulting maximal model [e.g., (anaphora type * context type) + (1| participant)] converged in most cases. The results include coefficients, standard errors, and t-values for each fixed effect and interaction. A given co-efficient was judged to be significant at $\alpha = 0.05$ if the absolute t-value/z-value exceeded two (Baayen et al., 2008). Data and scripts for all experiments are available at: https://osf.io/5pj8q/?view_only=2f26fbff1a9a49be94393e79c59d6ac3.

2.2 Results

In the anaphora region, regression-out showed a significant two-way interaction between the context with competitor type and anaphora type, with higher odd-ratios for *dieser* in the competitor condition than in the no-competitor condition (Table 1) (see Supplementary Figures 1, 2 and Supplementary Table 2 for means and standard errors). Due to the length difference between *der* and *dieser*, we approach such interaction cautiously. In the competitor context, *dieser* led to more regression-out than *dieser* in the no-competitor context, which was also significant in the pairwise comparison ($\beta = 1.271$, $SE = 0.572$, $z = 2.221$, $p = 0.026$). In the same region, total time revealed main effects of anaphora and competitor types. Compared to *der*, *dieser* led to longer reading times (again length effect). Processing of competitor context resulted in longer reading times than no-competitor.

In verb and disambiguation regions, regression-out and total time revealed a main effect of anaphora, with higher odd ratios and longer reading times for *der* than *dieser* (see Figures 1, 2; see Supplementary Table 2 for means and standard errors). However, there was no main effect of context with competitor type or a significant two-way interaction between the two factors.

Overall, our results show that *dieser* with competitor, in the anaphora region, led to more processing difficulties compared to *dieser* in the no-competitor context. This indicates that *dieser* signals competing two local antecedents and readers realize that they need to determine an antecedent. Since the pattern for *dieser* was seen in only one region for one eye-movement measure, we approach the sensitivity of *dieser* to competitor cautiously. While in later regions (i.e., verb and disambiguation regions), the initial preference for *dieser* changed, *der* was always difficult to process. References to the last-mentioned entity/agent in non-subject position (i.e., the inline skater) with *dieser* led to less processing difficulties than with *der*. The reading difficulty associated with *der* in both verb and disambiguation regions may stem from the preference for non-agent references with *der* (Schumacher et al., 2017), which aligns with the TR account. Our findings regarding *dieser* seem to support the CT account, suggesting an avoidance of the subject and consequently, less processing difficulty due to references to the last-mentioned entity (i.e., the inline skater/non-subject). However, a definitive conclusion eludes us. Hence, to ascertain which competing antecedents would be preferred for *der*/*dieser*, we conducted Sentence Completion Experiment 2.

3 Experiment 2

Experiment 2 tested participants' alternative antecedent preferences using a sentence completion method.

3.1 Methods

3.1.1 Participants

Thirty-one paid native German-speakers from the University of Cologne participated in the experiment (ages: M = 22; SD = 1.126). None had participated in Experiment 1.

3.1.2 Materials

We used the same sentences as in Experiment 1 (40 items, 60 fillers, 2 x 2 design). Unlike Experiment 1, each participant was provided with an initial text and asked to provide a completion for the sentence fragment starting with *Der* or *Dieser hat* in a manner consistent with the previous text (see Supplementary Section 2 for a sample item).

TABLE 1 Results of mixed-effects analysis for regression-out and total time for Experiment 1.

Regions/parameters	Regression-out			Total time		
	Odds ratio	SE	*p*-value	*β*	SE	*t*
Anaphora						
Intercept	0.07 (0.05–0.10)	0.01	0.001*	5.619	0.037	151.50
Context	1.52 (0.99–2.34)	0.33	0.055	0.059	0.023	2.561*
Anaphora	0.83 (0.48–1.42)	0.23	0.489	0.060	0.026	2.296*
Context * anaphora	2.46 (1.04–5.80)	1.08	0.040*	−0.024	0.046	−0.526
Verb						
Intercept	0.16 (0.12–0.22)	0.02	0.001*	5.718	0.058	97.481
Context	1.03 (0.79–1.34)	0.14	0.843	0.006	0.022	0.288
Anaphora	0.57 (0.44–0.75)	0.08	0.001*	−0.184	0.022	−8.276*
Context * anaphora	0.82 (0.48–1.40)	0.22	0.472	−0.005	0.044	−0.130
Disambiguation						
Intercept	0.12 (0.09–0.17)	0.02	0.001*	6.059	0.059	102.454
Context	0.91 (0.70–1.19)	0.12	0.494	0.010	0.019	0.520
Anaphora	0.47 (0.36–0.62)	0.06	0.001*	−0.045	0.019	−2.300*
Context * anaphora	0.93 (0.55–1.59)	0.25	0.799	0.073	0.039	1.877

Regressions-out are reported with odds ratios, standard errors, and p-values, the results for total times also include coefficients, standard errors, and t-values for each fixed effect and interaction. Context corresponds to the condition either with or without a competitor. *Signals interaction between the two factors.

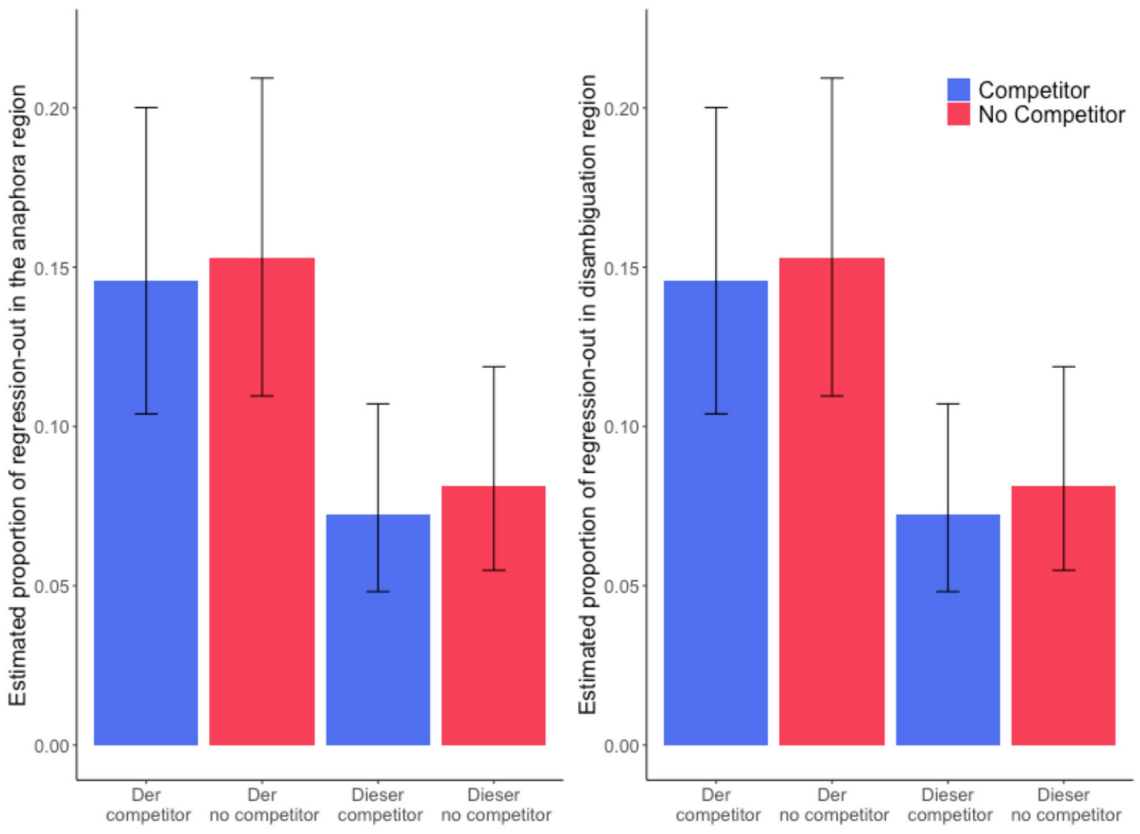

FIGURE 1
Regression-out in the anaphora and disambiguation regions. Error bars represent 95% confidence intervals.

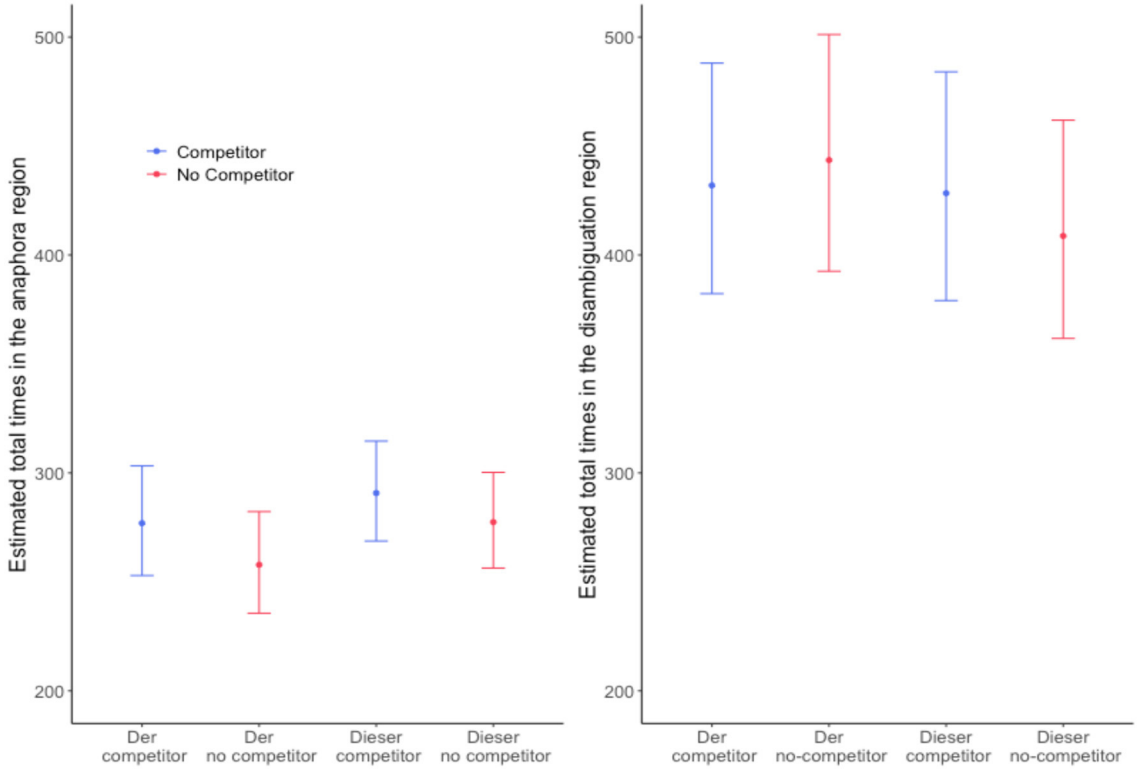

FIGURE 2
Total time in the anaphora and disambiguation regions. Error bars represent 95% confidence intervals.

3.1.3 Procedure

We used a two-block data collection process previously used in studies of plurals (Koh and Clifton, 2002; Çokal et al., 2023). In stage I, participants completed the given sentences in our lab. In stage II, immediately after the experimenter saved their completions, participants were asked to go over their completions and underline what *dieser* or *der* referred to. Filler sentences included referential expressions (e.g., *sie*, *das*) as well (see Supplementary Section 2 for coding).

3.2 Results

The data analysis reported in the manuscript was based on participants' underlined referential choices/interpretations. Since our data were categorical, we ran logistic mixed effects regression, taking anaphors (*dieser* vs. *der*) and competitor type (competitor and no-competitor context) as the fixed effects, and including crossed random intercepts and slopes for participants and items. Since the full model did not converge, we reduced the random effect structure until convergence was reached: response ~ (1 |Participant) + (1 |Item) + anaphor type * context type. In the logistic mixed effects regression, we used binary contrast for both factors.

There was a main effect of competitor type ($OR = 0.55$, $SE = 0.07$, $p = 0.001$) but not a main effect of demonstrative ($OR = 1.32$, $SE = 0.19$, $p = 0.053$), or an interaction between these two factors ($OR = 0.97$, $SE = 0.27$, $p = 0.899$; see Supplementary Table 1). Figure 3 presents that unlike CT and TR accounts, in the competitor context, *der/dieser* referred to both subject/non-agent and non-subject agent. The results in the no-competitor condition are surprising: only 60% refer to the non-subject agent in **S4** (i.e., the inline skater), while 40% refer to the entity mentioned in **S1** (i.e., the elderly person) even though it is not locally available (i.e., globally available). This suggests a tendency to avoid agent (even if non-subject), which would support TR account.

4 Discussion

Our first prediction was that *dieser* would be sensitive to competing antecedents (i.e., subject non-agent vs. non-subject agent). Results in the anaphora region for regressions-out showed that *dieser* in the competitor context leads readers to search for a local antecedent. Such an antecedent search for *dieser* might support assumptions in Ahrenholz (2007) and Saha et al. (2023).

Additionally, we observed higher odds ratios for regression-out in the verb region with *der,* which could perhaps be explained by the observation of *der* being used as a determiner phrase. A small corpus search from the Cosmas Tagged T2 supports this assumption. The percentage of determiner use of *der* is 95%, while the pronominal use is 5% [cases: *Der*-pronoun (n) = 13,352; *Der*-determiner (n) = 253,801]. In contrast, determiner use of *dieser* accounts for 50% of cases, with pronoun use also making up 50% [cases: *Dieser*-pronoun (n) = 9,402; *Dieser*-determiner (n =

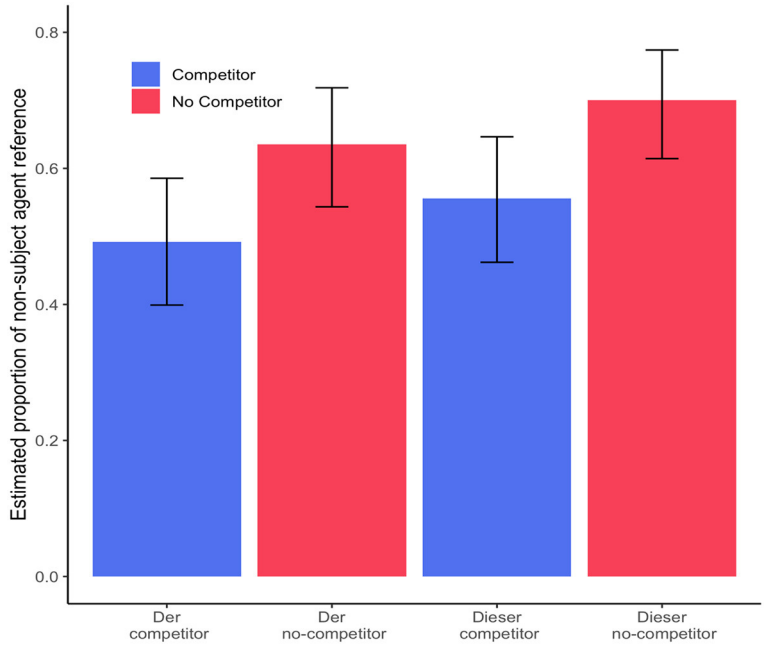

FIGURE 3
Estimated proportions of non-subject agent references out of total number of references (e.g., the inline skater vs. elderly adult). Error bars represent 95% confidence intervals.

9,525)]. In online reading, *dieser* led to less processing difficulties than *der*, perhaps due to a register effect (Patil et al., 2020).

Our second prediction was that in both conditions *der*—in the disambiguation region—accesses the globally available non-agent (i.e., the elderly person), leading to higher odd ratios in regression-out and longer total time (i.e., reading difficulty) than *dieser*. Our eye-tracking results supported our second prediction, showing longer reading times and regressions-out for *der*, in line with Schumacher et al.'s (2017) findings regarding non-agent antecedent for *der*.

In the disambiguation region, references to the last-mentioned entity with *dieser* led to lower odds ratios in regression-out and less total time. In the same region, processing with *der* was difficult due to a greater preference for the non-agent antecedent (i.e., the elderly person) than the agent references. While the processing of *dieser* aligns with the CT for demonstratives (i.e., less-prominent entity/recent entity), the processing of *der* supports TR (i.e., non-agent).

Our sentence completion experiment reveals that in the competitor condition, the probability of referring to non-agent subject or an agent non-subject entity is equal for *der/dieser*. In an offline task, participants use both semantic information of thematic role and syntactic information on grammatical role. Our results do not support previous findings from offline experiments (e.g., Bosch and Umbach, 2007's avoidance from subject/topic references). In the no-competitor condition, the probability of the local argument is even higher for both demonstratives (Ahrenholz, 2007). This is perhaps expected if there is only one argument. However, it is still surprising that demonstratives referred to the antecedent in the first sentence (i.e., the elderly person) but not in the previous sentence (S4).

Overall results may not be robust enough for both CT and TR accounts. However, it seems that there is a slight preference for the TR account over CT, which becomes evident for *der* in the online experiment. For the offline experiment, the modeling of demonstrative preference for a less prominent antecedent becomes achievable when assigning equal weight to the thematic role, grammatical role, and recency. In the competitor condition, this approach results in a balanced access to either the non-agent (subject) or the agent (non-subject, most recent) argument. However, in the no competitor context, demonstratives only refer to the local antecedent in 60% of cases. Forty-percentage references to an early mentioned entity might support the TR account. The differences between our results and previous findings can be explained by methodological differences. While our offline study shows the same antecedent distributions for *der/dieser* as in line with Fuchs and Schumacher (2020) and Bader et al. (2022), our online study shows subtle processing differences. One of our limitations is the absence of references to the subject/non-agent (i.e., the elderly person) in the eye-tracking reading experiment. However, our sentence completion experiment compensates for this limitation.

We think it is interesting to compare our results with studies that use the Visual World Paradigm (VWP). Unlike eye-tracking reading studies, the VWP shows immediate identification of a pronoun's antecedent, using the subject preference/first-mentioned entity as a disambiguation strategy (Ehrlich and Rayner, 1983; Arnold et al., 2000; Clifton et al., 2016; Brocher and von Heusinger, 2018). However, the subject preference does not hold true for demonstratives, as shown for Finnish by Kaiser and Trueswell (2008), and only applies to *der*—as noted in Wilson (2009). With this mind, if we run a VWP experiment using our paradigm,

we predict that hearing *der* would result in more looks to the subject non-agent (i.e., the elderly person), likely due to a spoken language/register effect in German. For *dieser*, fixations would be equally split between the subject non-agent (i.e., "the elderly") and non-subject agent (i.e., "the inline skater"), with increased saccades as participants would determine the referent (non-agent/elderly person vs. most recent/inline skater). When disambiguating information unfolds, fixations on the non-subject agent would increase for *dieser*, regardless of the competitor's presence. However, for *der*, participants would continue to prefer the subject non-agent antecedent, possibly avoiding the agent and focusing on the globally-available antecedent (i.e., "the elderly adult") over the locally available one (i.e., "the inline skater") in the no-competitor condition. Our current findings suggest that our new paradigm, distinct from prominence avoidance, should be further investigated using the VWP.

5 Conclusion

In conclusion, we propose that to disentangle the functions of *der* and *dieser*, the prominence avoidance paradigm, which most previous studies have used, is not useful in this situation. Greater contrast between *der* and *dieser* can be found by examining contexts where the antecedent(s) is prominent. Consequently, scholars should further investigate our paradigm using both eye-tracking and EEG methodologies.

Data availability statement

The raw data supporting in the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of the University of Cologne (Nr. 2016-09E2-200213). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

DÇ: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. KH: Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)-Project-ID 281511265-SFB 1252 Prominence in Language -in the project C04 "Conceptual and referential activation in discourse" at the University of Cologne, Department of German Language and Literature I, Linguistics.

Acknowledgments

We would like to express our gratitude to Robert Voigt and Nagihan Gökben Konuk for their invaluable assistance in crafting the stimuli and conducting piloting sessions with native speakers of German. We also thank Petra B. Schumacher, Stefan Hinterwimmer, Massimo Poesio, Markus Phillip, and Timo Buchholz for discussion.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/flang.2024.1433482/full#supplementary-material

References

Ahrenholz, B. (2007). *Verweise mit Demonstrativa im gesprochenen Deutsch. Grammatik, Zweitspracherwerb und Deutsch als Fremdsprache*. Berlin: de Gruyter.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., and Trueswell, J. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition* 76, 13–26. doi: 10.1016/S0010-0277(00)00073-1

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005

Bader, M., Portele, Y., and Schäfer, A. (2022). "Semantic bias in the interpretation of German personal and demonstrative pronouns," in *Proceedings of Linguistic Evidence 2020: Linguistic Theory Enriched by Experimental Data*, eds R. Hörnig,

S. von Wietersheim, A. Konietzko, and S. Featherston (Tübingen: University of Tübingen), 399–419.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bosch, P., and Umbach, C. (2007). "Reference determination for demonstrative pronouns," in *Proceedings of the Conference on Intersential Pronominal Reference in Child and Adult Language*, eds D. Bittner and N. Gagarina (Berlin: Zentrum für Allgemeine Sprachwissenschaft), 39–51.

Brocher, A., and von Heusinger, K. (2018). A dual-process activation model: processing definiteness and information status. *Glossa J. Gen. Linguist.* 3, 1–34. doi: 10.5334/gjgl.457

Brown-Schmidt, S., Byron, D. K., and Tanenhaus, M. K. (2005). Beyond salience: interpretation of personal and demonstrative pronouns. *J. Mem. Lang.* 53, 292–313. doi: 10.1016/j.jml.2005.03.003

Buchholz, T., and von Heusinger, K. (2024). German demonstrative pronouns differ in their sensitivity to discourse and sentence topics. *Front. Commun.* 9:1369290. doi: 10.3389/fcomm.2024.1369290

Chafe, W. (1976). "Givenness, contrastiveness, definiteness, subjects, topics, and point of view," in *Subject and Topic*, ed C. N. Li (New York, NY: Academic Press), 25–55.

Clifton, Jr. C., Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., et al. (2016). Eye movements in reading and information processing: Keith Rayner's 40-year legacy. *J. Mem. Lang.* 88, 1–19. doi: 10.1016/j.jml.2015.07.004

Çokal, D., Filik, R., Sturt, P., and Poesio, M. (2023). Anaphoric reference to mereological entities. *Discour. Process.* 60, 202–223. doi: 10.1080/0163853X.2023.2197682

Çokal, D., Sturt, P., and Ferreira, F. (2016). Processing of it and this in written narrative discourse. *Discour. Process.* 55, 272–289. doi: 10.1080/0163853X.2016.1236231

Cokal, D., and von Heusinger, K. (2024). German demonstrative pronouns in contrast. *Dialogue Discourse* 15, 45–76. doi: 10.5210/dad.2024.102

Cornish, F. (2008). How indexicals function in texts: discourse, text, and one neo-Gricean account of indexical reference. *J. Pragmat.* 40, 997–1018. doi: 10.1016/j.pragma.2008.02.006

Ehrlich, K., and Rayner, K. (1983). Pronoun assignment and semantic integration during reading: eye movements and immediacy of processing. *J. Verb. Learn. Verb. Behav.* 22, 75–87. doi: 10.1016/S0022-5371(83)80007-3

Fuchs, M., and Schumacher, P. B. (2020). "Referential shift potential of demonstrative pronouns—evidence from text continuation," in *Demonstratives in Discourse*, eds Å. Næss, A. Margetts, and Y. Treis (Berlin: Language Science Press), 185–213.

Kaiser, E., and Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: evidence for a form-specific approach to reference resolution. *Lang. Cogn. Process.* 23, 709–748. doi: 10.1080/01690960701771220

Karamanis, N., C. M., Poesio, M., and Oberlander, J. (2009). Evaluating centering for information ordering using corpora. *Comput. Linguist.* 35, 29–46. doi: 10.1162/coli.07-036-R2-06-22

Koh, S., and Clifton, C. J. (2002). Resolution of the antecedent of a plural pronoun: ontological categories and predicate symmetry. *J. Mem. Lang.* 46, 830–844. doi: 10.1006/jmla.2001.2829

Patil, U., Bosch, P., and Hinterwimmer, S. (2020). Constraints on German diese demonstratives: language formality and subject-avoidance. *Glossa J. Gen. Linguist.* 5, 1–22. doi: 10.5334/gjgl.962

Peeters, D., Krahmer, E., and Maes, A. (2021). A conceptual framework for the study of demonstrative reference. *Psychon. Bullet. Rev.* 28, 409–433. doi: 10.3758/s13423-020-01822-8

Poesio, M., Stevenson, R., Eugenio, B. D., and Hitzeman, J. (2004). Centering: a parametric theory and its instantiations. *Comput. Linguist.* 30, 309–363. doi: 10.1162/0891201041850911

Rayner, K., Slattery, T., Drieghe, D., and Liversedge, S. (2011). Eye movements and word skipping during reading: effects of word length and predictability. *J. Exp. Psychol.* 37, 514–528. doi: 10.1037/a0020990

Saha, A., Sag, Y., and Davidson, K. (2023). "Focus on demonstratives: experiments in English and Turkish," in *Proceedings of the 33rd Semantics and Linguistic Theory Conference*, eds J. Kim, B. Öney, Y. Zhang, and F. L. Zhao (Linguistic Society of America), 460–479.

Schumacher, P. B., Roberts, L., and Järvikivi, J. (2017). Agentivity drives real-time pronoun resolution: evidence from German er and der. *Lingua* 185, 25–41. doi: 10.1016/j.lingua.2016.07.004

Stevenson, R. J., Crawley, R. A., and Kleinman, D. (1994). Thematic roles, focus, and the representation of events. *Lang. Cogn. Process.* 9, 519–548. doi: 10.1080/01690969408402130

Strube, M., and Hahn, U. (1999). Functional centering: grounding referential coherence in information structure. *Comput. Linguist.* 25, 309–344.

Wilson, F. (2009). *Processing at the Syntax-Discourse Interface in Second Language Acquisition* (Ph. D. thesis). University of Edinburgh, Edinburgh, United Kingdom.