

ORIGINAL ARTICLE

Interpretation preferences in contexts with three antecedents: examining the role of prominence in German pronouns

Clare Patterson*  and Petra B. Schumacher

Department of German Language and Literature I, Linguistics, University of Cologne, Germany

*Corresponding author. Email: cpatters@uni-koeln.de

(Received 4 October 2019; revised 28 May 2021; accepted 11 June 2021; first published online 31 August 2021)

Abstract

This paper focuses on the relational notion of prominence, in which entities of equal type are ranked according to certain prominence-lending features. In German two demonstrative forms, “*der*” and “*dieser*”, can function like personal pronouns in English. It has been proposed that processing “*der*” involves computing a prominence hierarchy of the prior referents, and excluding the referent with the highest prominence rank. The demonstrative “*dieser*” has not been extensively tested. In the current study, personal and demonstrative pronominal forms were investigated following ditransitive contexts, where three potential antecedents are available, in two rating experiments. The personal pronoun showed flexibility in that it received equally high ratings for all three antecedents in canonical configurations. The ratings for *dieser* followed a graded sensitivity to thematic role prominence, with lowest scores when referring to prominent antecedents (agents) and the highest scores for the least prominent antecedents (patients), with scores for the medium prominence candidate (recipients) differing from both. *Der* followed a similar but not identical pattern, with a less marked difference between lower prominence candidates. Positional information also has a strong influence on demonstratives. In sum, final interpretation is sensitive to fine-grained differences in prominence hierarchies.

Keywords: pronoun resolution; demonstratives; reference; prominence

The resolution of personal pronouns has attracted a great deal of attention in the psycholinguistics literature. It has been established that pronoun resolution can be affected by various different factors, and that many of these factors contribute to the relative prominence of potential antecedents in the prior discourse (Arnold, 2010; Arnold et al., 2000; Cowles et al., 2007; Järvikivi et al., 2005; Kaiser & Trueswell, 2008; Rohde, 2019; Schumacher et al., 2016). Thus, we can generalize that the process of resolving a pronoun is driven, at least to some extent, by the prominence of its antecedent.

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

The establishment of coreference relations is here viewed within the prominence framework (Himmelmann & Primus, 2015; von Heusinger & Schumacher, 2019). Prominence is characterized as an organizational principle that is relational and dynamic and attracts structural operations. For the purposes of this paper, we focus on the relational notion, that is that entities of equal type (here reference to individuals) are ranked according to certain prominence-lending features. The criterion of dynamicity posits that linguistic units may shift their (prominence) status as discourse unfolds, that is the prominence status of an entity is not static and entities can be promoted or demoted with regard to their prominence status. The criterion of structural attraction assumes that prominent entities attract more operations than less prominent entities. This is, for instance, reflected in the availability of alternative structures for prominent entities, for example, prominent arguments license passive constructions or clefts or prominent referents express coreference with more referential forms.

A core prerequisite for the relational notion is that referents are represented as an ordered set in discourse representation (e.g., Grosz et al., 1995; von Heusinger, 2006). The ordering of referents at a given point in time relies on prominence-lending cues. While various factors have been identified to interact during the computation of prominence rankings (see below), we test the hypothesis that agentivity serves as a central prominence-lending cue for reference resolution (Schumacher et al., 2016, 2017).

We further assume that coreference resolution is affected by the prominence ranking. We thus utilize coreference via different referential expressions (i.e., personal vs. demonstrative pronouns) to observe the prominence ranking.¹ In contrast to accounts that posit a strict correspondence between referential expressions and the status of an entity in discourse (e.g., Ariel, 1990; Gundel et al., 1993), the current proposal adopts a more dynamic correspondence between form and function.

We thus ask how comprehenders choose from an ordered set during reference resolution. Some theories assume one backward looking center, such as Centering Theory (Grosz et al., 1995), which has been widely applied to the resolution of personal pronouns.² In its original proposal, the prominence framework proposed that one entity is singled out from a set of entities of equal type. Here, we test whether the role of prominence ranking is merely to single out one referent from a set or whether it makes available an ordered set that interacts with different referential expressions in specific ways.

Earlier approaches to pronoun resolution sought to identify the individual factors which determined the prominence of particular antecedents. Later, a more nuanced view of prominence developed, in which prominence is determined via the combination of several factors such as grammatical role, position, or thematic role (Bader & Portele, 2019; Järviö et al., 2005; Kaiser & Trueswell, 2004; Schumacher et al., 2015, 2016). More recent approaches have gone beyond this and identified how pronoun resolution is part of a larger process of establishing coherence in a text or discourse, but is at the same time affected by certain biases. For instance, Kehler and Rohde (2013) show that pronoun interpretation influenced by form-related production biases (how likely it is that a pronoun will be produced to refer to a particular entity) but also next mention biases (likelihood that a particular entity will be mentioned again); these biases are related in a Bayesian model (see also Kehler et al., 2008).

While much research in pronoun resolution has been devoted to third-person personal pronouns (he/she/they) in English, in recent years, a broader spectrum of languages and pronoun types have been investigated (e.g., Carminati, 2002; Çokal et al., 2018; Colonna et al., 2012; de la Fuente, 2015; Ellert, 2010; Fossard et al., 2012; Hemforth et al., 2010; Kaiser, 2011b; Kaiser & Trueswell, 2008; Keating et al., 2016). The current study continues this trend by investigating personal and demonstrative pronouns in German. Broadening research across a wider variety of languages not only increases the cross-linguistic validity of existing proposals, it also gives access to a wider range of pronoun types, which leads us to consider different questions with respect to pronoun resolution. For instance, studies in Finnish have shown that the pronouns *hän* and *tämä* are differently sensitive to syntactic role and word order/information structure (Kaiser, 2003, 2005; Kaiser & Trueswell, 2008). The contrast between null and overt pronouns (for instance in Spanish, Italian, Greek, Turkish) points to a division of labor between these two forms (e.g., Alonso-Ovalle et al., 2002; Carminati, 2002; Dimitriadis, 1996; Filiaci, 2010; Turan, 1995); moreover, this division of labor is subtly different in different languages and language varieties (e.g., Barbosa et al., 2005; Filiaci et al., 2014). A similar division of labor may also be observed in the contrast between personal and demonstrative forms in Germanic languages such as Dutch (e.g., Haeseryn et al., 1997; Kaiser, 2011a).

Our investigation of German personal and demonstrative pronouns addresses a major limitation in the research to date: namely that many experiments on pronouns are limited to contexts in which only two potential antecedents are available. This limitation gives the impression that prominence is binary: a potential antecedent is either prominent or not prominent for the purposes of pronoun resolution, leaving open the question of how pronoun resolution plays out when more than two antecedents are in the discourse context. Our study starts to address this limitation by investigating the resolution of pronouns in ditransitive constructions where three potential antecedents are available. The aim of our study is to test the sensitivity of the different pronouns to the relative prominence of the antecedent when more than two antecedents are available, and also to probe the division of labor between the demonstrative forms. This is a particularly relevant question given that demonstratives in German seem to target less prominent antecedents.

In the next section, we review previous research on personal and demonstrative pronouns in German and clarify our assumptions about the notion of prominence in pronoun resolution; following that we review the relevant aspects of ditransitive constructions and work on pronoun resolution in these contexts. We then present hypotheses about division of labor and targeting less prominent antecedents which are addressed in the current study.

Personal and demonstrative pronouns in German

German has a richer pronominal system than English for referring to animate entities. The personal pronouns, which are inflected for gender, number, and case, are complemented by several types of demonstrative pronoun. In the current paper, we focus on masculine and feminine forms of the third-person singular personal pronouns *er/sie* (“he/she”), contrasting them with two types of demonstrative,

der/die and *dieser/diese* (for convenience, we refer subsequently only to the masculine form of these pronouns).

The German demonstrative pronouns are regularly used to refer to human referents (as well as inanimate referents), in contrast to English *this/that*.³ The two demonstrative forms in German, *der* and *dieser*, do not convey distance-based information such as proximate/distal (but note that there is such a contrast between *dieser* and the more obsolete *jener*). *Dieser* further expresses contrast or delimitation (e.g., Bisle-Müller, 1991). The two demonstrative pronouns *der* and *dieser* are claimed to differ in their potential to shape the upcoming discourse (Ahrenholz, 2007; Ehlich, 1983; Weinrich, 1993; Zifonun et al., 1997), but interpretive differences between these forms have not been empirically demonstrated, with the exception of Fuchs & Schumacher (2020) who show discrete referential shift functions.

Previous studies on the interpretation of different pronominal forms have focused mainly on *der* in contrast to the personal pronoun *er*. *Dieser* has received relatively less attention. Findings on the interpretation of *er* have found that this form refers preferentially to a prominent antecedent from the prior discourse (Bader & Portele, 2019; Bosch et al., 2007; Bouma & Hopp, 2006, 2007; Ellert & Holler, 2011; Schumacher et al., 2016), while it retains flexibility in its interpretation and can also refer felicitously to less prominent antecedents, in particular when multiple prominence hierarchies come into play (Bosch et al., 2007; Schumacher et al., 2016, 2017; Wilson, 2009). Findings for *der* mainly agree that it refers to a less prominent antecedent and is more rigid in its interpretation (Bader & Portele, 2019; Bosch et al., 2007; Hinterwimmer & Bosch, 2016; Kaiser, 2011b; Schumacher et al., 2015, 2016, 2017; Wilson, 2009). However, there is no firm consensus across studies on the features contributing to prominence as it relates to pronoun resolution in German. Earlier studies proposed that grammatical role was the critical factor with personal pronouns preferring subject antecedents and demonstratives preferring nonsubjects (Bosch et al., 2003, 2007). This was disputed in later studies, which proposed instead that topichood was the critical factor (Bosch & Hinterwimmer, 2016; Bosch & Umbach, 2007; Wilson, 2009; see also Abraham, 2002). For instance, Bosch and Hinterwimmer (2016) proposed that the semantic representation of *er* and *der* differs by way of an “avoid topic” feature for *der*.

A different proposal is that thematic role is an important component of prominence when it comes to German pronouns, particularly when contrasted with the contribution of grammatical role. This has been supported by a number of studies that manipulated both thematic role hierarchy and grammatical role hierarchy by contrasting anaphoric resolution preferences in active-accusative verbs versus dative-experiencer verbs. Schumacher and colleagues (2016) showed that *er* was resolved most often to antecedents with the thematic role of (proto-) agent⁴, while *der* was resolved most often to antecedents with the role of (proto-) patient. These findings have been confirmed by ERP evidence and sentence completions (Schumacher et al., 2015) as well as visual world eye tracking (Schumacher et al., 2017) and acceptability ratings and eye-tracking during reading (Patterson & Schumacher, 2020). Nevertheless, these studies also showed that it was not the thematic role alone that determined resolution preferences, rather, it was thematic role in combination with grammatical role and linear order that determined these

preferences. Relatedly, Portele and Bader (2016) claim that a combination of syntactic prominence and givenness is a key component in differentiating the two pronouns, and Bader and Portele (2019) claim that d-pronouns favor the least prominent antecedent that can be determined by a combination of topichood, grammatical role, and linear position, while the personal pronoun in contrast appears to favor the subject of the previous clause.

In contrast with *er* and *der*, the interpretation preferences for *dieser* have received very little attention to date. Zifonun et al. (1997) claim that *dieser* is restricted to referring to the last-mentioned antecedent, but they do not present any empirical evidence for this claim. Preliminary studies suggest, however, that *dieser* patterns similarly to *der* in preferring an antecedent with lower prominence, even when the lower prominence antecedent was not the last-mentioned antecedent (Fuchs & Schumacher, 2020; Lange, 2016; Özden, 2016). In Özden's (2016) sentence completion study, two potential antecedents, with the roles agent and patient, were introduced.⁵ The order of antecedents was manipulated (agent-before-patient vs. patient-before-agent). Note that the patient-before-agent order gives rise to a change in information structure such that the patient is now the assumed topic. *Dieser* was mainly interpreted as referring to the patient, irrespective of antecedent order. Similarly, Lange (2016) found that *dieser* was usually interpreted as the (proto-)patient with dative-experiencer verbs, irrespective of antecedent order. These preliminary findings are not compatible with Zifonun et al.'s (1997) description that *dieser* refers to the last-mentioned antecedent and are suggestive instead of other prominence cues being involved in the interpretation of *dieser*.

Based on the evidence presented above, we can characterize the resolution preferences of German personal and demonstrative pronouns as follows. Personal pronouns are somewhat flexible in their reference, but show a tendency to resolve to more prominent antecedents, such as topics, subjects or those with the (proto-)agent thematic role. The demonstrative *der* strongly avoids referring to prominent antecedents, and has a strong tendency to refer to antecedents with a less prominent role such as (proto-)patient. The demonstrative *dieser* patterns with *der* in referring to less prominent antecedents, with preliminary evidence pointing away from a purely position-based account for *dieser*.

It is not our aim in this study to directly examine the precise nature of prominence-lending cues; however, it is important here to clarify our own assumptions for the purpose of the current study. We follow the majority of recent studies in rejecting the outdated notion that prominence is determined by a single factor. Rather, we follow Schumacher et al. (2015, 2016) and Bader and Portele (2019) (who limit this notion to demonstrative pronouns) in assuming that a single prominence hierarchy is determined by a combination of several factors including, but not limited to, thematic role and linear order. Furthermore, we follow Schumacher et al. (2015, 2016) in assuming that there is some commonality between factors, which determine the prominence hierarchy for personal pronouns and demonstrative pronouns, contra Kaiser and Trueswell (2008) who argue that different pronoun types are affected by different factors. Finally, we acknowledge the body of work by Kehler, Rhode and colleagues (e.g., Kehler et al., 2008; Kehler & Rohde, 2013; Rohde & Kehler, 2014), which suggests that the importance of prominence is relegated to a subject or topic resolution bias for personal pronouns (in English); this

overlays a general production bias toward mentioning entities that are rendered semantically prominent through coherence, discourse properties, world knowledge, etc. The semantic component is not specific to pronouns, but is predicated on the notion of which entity is likely to be mentioned next, in whatever form. While these claims are important to understand the underlying mechanisms supporting pronoun resolution, the current study does not support or contradict the claims made in this literature, and the focus of this study is not to apply the Bayesian model to German pronouns (for a discussion of the Bayesian model in German pronouns, see Bader and Portele (2019)). The focus of this study is rather to probe the behavior of personal and demonstrative pronouns in contexts with three antecedents, a question that has not been much addressed to date.

The evidence reviewed so far is limited to scenarios where two potential antecedents are presented. It is, therefore, unclear to what extent the above findings are generalizable to a broader set of contexts. This limitation may also mask the precise division of labor among the different pronoun types, particularly *der* and *dieser*. In the current study, we address this limitation by presenting pronouns following ditransitive constructions. By using ditransitive verbs, three potential antecedents are presented, which allows us to extend our understanding of the resolution of personal and demonstrative pronouns. Ditransitive constructions are particularly useful to address this issue because they introduce three potential antecedents that are all arguments of the same verb; these arguments are thought to have an internal hierarchy in terms of the prominence of their thematic roles. Other ways of introducing three potential antecedents, for instance, a matrix verb with a clausal complement such as “Marie thinks [that Louise likes Jasmine]”, have an additional hierarchical clause structure that may cause a confound when examining prominence hierarchies; using ditransitive constructions avoids this problem.

Using ditransitives allows us to test hypotheses that are not testable with two antecedents alone. Our question is, how does the comprehension system choose from an ordered set of potential antecedents? The prominence framework, as discussed at the beginning of this paper, originally proposed that one entity was singled out from a set of entities of equal type. Here, we test whether the role of prominence ranking is merely to single out one referent from a set, rendering the other referents “invisible” to the referential expression, or whether referential expressions interact with the complete (ordered) set of referents. Take, as a basis, a ditransitive context with three arguments, A, B, and C. These arguments form an ordered set, with A the most prominent and C the least prominent (setting aside discussion of precisely which factors determine the prominence ranking). Encountering a demonstrative pronoun involves selecting an antecedent from this ordered set. Now imagine that the “rule” for resolving a demonstrative pronoun involves excluding the highest-ranked candidate. In a transitive context when only two potential antecedents are available (A, B), excluding the highest ranked candidate, A, only leaves B available for reference. But in a ditransitive context, such a description is underspecified. Excluding A leaves both B and C available. Are these candidates equally likely to be chosen? Existing accounts of demonstrative resolution do not explicitly state their expectations with regard to three-candidate contexts, but we can identify two contrasting scenarios that are implied by the characterization of demonstrative resolution preferences discussed earlier, as well as proposing a third

alternative. The most common way of describing the preferences for demonstratives in German is to state that they avoid the most prominent antecedent; for instance, the “avoid topic” feature developed in Bosch et al. (2003) and Bosch and Hinterwimmer (2016). This implies that, once candidate A has been taken out of play, then any other available antecedents (B and C) are equal candidates for demonstrative reference. An alternative, contrasting, possibility is that the demonstrative seeks out the least prominent antecedent. This is most closely implied in Bader and Portele (2019, p. 185), who state that “a d-pronoun prefers the antecedent that is least favored by the structural biases”. Targeting a single (lowest prominence) antecedent, (C in our example) implies that A and B are unavailable or not chosen during resolution. In a three-antecedent scenario, then, the two remaining antecedents that are not the least prominent are equally unavailable.

There is one further possibility. There could be a graded preference that directly reflects the prominence ranking, whereby the least prominent antecedent is most favored (for demonstrative resolution), the most prominent antecedent is least favored, and the medium prominent antecedent sits between the two extremes. This final possibility, which assumes graded preferences, is not specifically excluded by existing accounts of demonstrative resolution. But neither is it proposed or discussed, probably because it is not possible to test when only two possible antecedents are presented. Directly accessing a set of potential referents with graded preferences represents an extension of the prominence framework. This possibility involves further specifying demonstrative preferences to include sensitivity to prominence ranking, beyond excluding the most prominent or targeting the least prominent candidate.

Our current study uses ditransitive constructions to present three potential antecedents in order to assess the resolution preferences of personal and demonstrative pronouns in German. Below we review existing relevant work on ditransitives.

Ditransitive constructions and pronoun resolution

Ditransitive constructions are those in which a predicate takes three arguments; the prototypical examples are transfer of possession events with the verb “give”, as in (1), which involves *the girl*, *the boy*, and *a book*:

- (1) The girl gave the boy a book.

This is expressed in German in a similar way, with the recipient argument (*the boy* in (1)) taking dative case and the theme (*a book* in (1)) taking the accusative case:

- (2) Das Mädchen gab dem Jungen ein Buch.

The girl gave the boy a book.

In English, the same event can also be expressed using a prepositional object for the recipient argument (*The girl gave a book to the boy*).⁶ There has been a great deal of debate in the theoretical literature about the motivation for this alternation. What is relevant for the current study, however, are the thematic roles assigned to the three

arguments in constructions such as (2) and their relative prominence. The nominative-marked argument is generally understood to have the thematic role of source or proto-agent, the accusative argument is the theme or proto-patient, and the dative argument is the proto-recipient. Proto-recipients (*the boy* in [1]) share features with the other two proto-roles, for example, in transfer of possession events, they are “proto-patient-like” by undergoing a change in possession relative to the proto-agent (*the girl*) but they are “proto-agent-like” in their function as possessor relative to the proto-patient (*the book*). Primus (1999) in her analysis of thematic roles assigns the three main roles the hierarchy PROTO-AGENT > PROTO-RECIPIENT > PROTO-PATIENT (for simplicity, we refer to these proto roles as agent, recipient and patient throughout the paper). Primus highlights the shared features between the agent and the recipient by placing the recipient higher in the hierarchy than the patient. A similar constellation holds for verbs with benefactives.⁷

Earlier psycholinguistic work on English pronoun resolution in ditransitive contexts focused mainly on the observation that there is a bias toward resolving pronouns to referents with the role of goal (recipient) in transfer of possession events (e.g., Stevenson et al., 1994). Arnold (2001) used sentence completion and corpus data to demonstrate that the bias arises from a general likelihood of a particular discourse referent being mentioned again; she suggested that the activation of the recipient in the discourse was increased because of the increased expectation that it would be mentioned again, making pronoun resolution to the recipient more likely. Data from Rohde et al. (2006) and Kehler et al. (2008) suggests that it is not the event itself, but the unfolding of the discourse via different coherence relations that in turn makes either the agent or the recipient more prominent with respect to pronoun resolution. Kehler and Rohde (2017) demonstrate that expectations about the unfolding discourse (in this case, manipulating the expected Question-under-Discussion) can influence the preference for source or goal interpretations of pronouns in real time before the critical information has even been encountered. Based on these findings, we took account of potential biases arising from the verb/scenario/discourse context in the current study by paying particular attention to the discourse coherence of the items, and by pretesting the materials.

Prior research on the behavior of pronouns in dative constructions in German is fairly limited by comparison. Bouma and Hopp (2006) tested preferences for resolving the personal pronoun *er* to the recipient and patient arguments in German (although ditransitive verbs were only a subset of their experimental items), using a referent-choice task. They did not find a strong preference for resolving *er* to the recipient or patient argument, regardless of argument order (recipient before patient or patient before recipient). In Bouma and Hopp (2007) there was again no preference, except in conditions when one of the arguments was topicalized by moving it to initial position, in which case, a slight preference for resolving to the recipient was observed. These results support the notion that the personal pronoun is flexible in its interpretation; however, because they did not test resolution to the agent argument in the ditransitive constructions, it is not possible to assess the true extent of this flexibility among three potential referents. Demonstrative pronouns were not included in these studies. Two previous studies that did include demonstratives in ditransitive contexts were conducted by Uzun (2015b, 2015a). In a forced-choice pronoun resolution task and a sentence

completion task, Uzun (2015a) compared *er*, *der*, and *dieser* in ditransitive contexts while manipulating the animacy of the patient argument and the word order. This resulted in a strong general preference to resolve **all** pronouns to the patient argument. This result is unexpected given the large number of previous studies in two-argument constructions, where clear differences emerge between preferences for personal and demonstrative pronouns.⁸ In Uzun's studies, the manipulation of both animacy and word order led to a very high number of experimental conditions; additionally, the ditransitive contexts with three animate arguments did not all sound very natural. In the current study, we hope to improve upon the design and materials by restricting the number of conditions and by making the experimental items sound more natural.

Current study

The aim of the current study is to test the resolution preferences of German personal and demonstrative pronouns in ditransitive constructions, when three potential antecedents are available. This enables us to examine how preferences for less prominent antecedents are worked out for demonstrative pronouns in particular, in nonbinary contexts. We test three alternative scenarios, as outlined above. One scenario for demonstratives is that the most prominent candidate antecedent is identified and rejected, but all other candidates are equally available. The second scenario is that the candidate antecedent with the lowest prominence rank is identified as the most suitable antecedent, meaning that reference to all other candidate antecedents is rejected. The final scenario is that the resolution preferences of the pronominal forms are graded, directly reflecting their prominence ranking. In addition to uncovering preferences for demonstratives in nonbinary contexts, we can also observe whether there is a division of labor between the two forms of the demonstrative, *der* and *dieser* (cf. Brown-Schmidt et al., 2005; Çokal et al., 2018; Kaiser & Trueswell, 2004, 2008; Wilson, 2009). We conducted three acceptability rating experiments (1a, 1b, and 2) with passages consisting of two sentences that included a referentially unambiguous pronoun. The use of acceptability ratings, rather than a task that asks the participant to choose a referent, ensures that we have sufficient and balanced data not only about the favored antecedent but also about less favored antecedents as well, which is important for the assessment of and comparisons between dispreferred candidates.

The purpose of Experiment 1 (1a and 1b) was to test resolution preferences for demonstrative pronouns (and personal pronouns as a point of comparison) in ditransitive contexts in which three candidate antecedents are available; Experiment 2 more closely examines the last-mention preference for *dieser* (Zifonun et al., 1997) by repeating Experiment 1, but reversing the order of the last two antecedents.

Following Primus (1999), we assume that the linearization of thematic roles in canonical order gives rise to a clear prominence ranking among the three potential antecedents in ditransitive contexts (with an inanimate patient): Agent > Recipient > Patient. Note that, while animacy has been viewed as a prominence-lending feature in previous research (e.g., Aissen, 1999, 2003), we assume that animacy is an epiphenomenon of agentivity and its feature specification (García García

et al., 2018; Schumacher, 2018). If it were to exert an independent influence on prominence, animacy would, in any case, align with the ranking shown above.⁹

Experiment 1a

Hypotheses, predictions, participant numbers, and a data analysis plan (including exclusion criteria for participants) were registered in advance on aspredicted.org (see supplemental materials in the Data Availability section).

Materials

Items consisted of two sentences each and the second sentence included an unambiguous pronoun, which was disambiguated on the basis of gender agreement or plausibility information. The first sentence (S1) was a ditransitive construction, that is, a three-argument construction containing an agent, a recipient and a patient.¹⁰ Agent and recipient were always animate role nouns and the patient was always inanimate.¹¹ The thematic roles of the recipient argument include recipient-possessor, goal-possessor, addressee-listener, and addressee-viewer. Benefactive constructions were also included.¹² The second sentence (S2) began with a pronoun referring unambiguously to one of the three NPs and described a state or event relating to that referent. Two three-level factors are fully crossed in order to create nine conditions. The factor Pronoun manipulated the pronoun type that was presented: the personal pronoun *er*; the demonstrative pronoun *der*; or the demonstrative pronoun *dieser*; in half the items, the pronouns were masculine, otherwise feminine. The factor Referent manipulated which NP the pronoun refers to (agent, recipient, or patient). This was achieved via gender and plausibility cues. In the Agent conditions, the pronoun matched in gender with the agent and patient but not the recipient; the event/state described in S2 was only compatible with an animate agent, thus ruling out the patient as a possible referent. In the Recipient conditions, the pronoun matched only the recipient in gender, and the patient was also ruled out on grounds of plausibility. In the Patient conditions, the gender of the pronoun matched all three NPs, but the event/state described in S2 was only compatible with an inanimate entity, thus ruling out the agent and recipient as possible referents.

The event or state described in S2 was always different for the Patient conditions in comparison to the Agent and Recipient conditions to ensure that it was compatible with an inanimate referent. In order to retain a high degree of comparability between the Agent and Recipient conditions, S2 was kept as similar as possible between these two conditions (in 13 items it is identical), with the same event or state being described. However, it was also crucial to maximize the discourse coherence between the two sentences so that the participants' ratings were not adversely affected by unnatural-sounding stimuli. For this reason, where necessary S2 contained an adverb that differed between Agent and Recipient conditions, or the event/state was slightly different, ensuring that the scenario sounded natural and coherent in each of the conditions. An example of the nine conditions is shown in Table 1. The full set of materials can be found in the supplemental materials linked to in the Data Availability section.¹³

Table 1. Materials for Experiment 1a (nine conditions)

	<i>Agent conditions</i> (pronoun refers to NP1)	<i>Recipient conditions</i> (pronoun refers to NP2)	<i>Patient conditions</i> (pronoun refers to NP3)
Context sentence	Die Eigentümerin vermietete <i>the.F.SG.NOM owner.F rent.PST-3SG</i> dem Anwohner die Parkfläche. <i>the.M.SG.DAT resident.M the.F.SG.ACC parking-space.F</i>	Der Eigentümer vermietete <i>the.M.SG.NOM owner.M rent.PST-3SG</i> der Anwohnerin die Parkfläche. <i>the.F.SG.DAT resident.F the.F.SG.ACC parking-space.F</i>	Die Eigentümerin vermietete <i>the.F.SG.NOM owner.F rent.PST-3SG</i> der Anwohnerin die Parkfläche. <i>the.F.SG.DAT resident.F the.F.SG.ACC parking-space.F</i>
ER/SIE	Sie war froh über den <i>she be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Sie war froh über den <i>she be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Sie war glücklicherweise schattig. <i>she be.PST-3SG fortunately shady</i>
DER/DIE	Die war froh über den <i>she.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Die war froh über den <i>she.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Die war glücklicherweise schattig. <i>she.DEM be.PST-3SG fortunately shady</i>
DIESER/ DIESE	Diese war froh über den <i>she.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Diese war froh über den <i>she.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Diese war glücklicherweise schattig. <i>she.DEM be.PST-3SG fortunately shady</i>
Translation	"The (female) owner rented the parking space to the (male) resident. She was delighted about the long-term contract."	"The (male) owner rented the parking space to the (female) resident. She was delighted about the long-term contract."	"The (female) owner rented the parking space to the (female) resident. Fortunately it was a shady spot."

Procedure

Thirty-six items, which had been pretested for plausibility¹⁴ were distributed across nine lists in a Latin square design; participants saw each of the 36 items in one condition only, and saw four items per condition. The 36 items were interspersed with 36 filler scenarios that were one or two sentences long (a list of fillers can be found in the supplemental materials linked to in the Data Availability section). Fourteen fillers contained no pronoun, and the remainder contained a variety of pronoun types. Fourteen fillers were deliberately made unnatural or implausible, in order to encourage participants to use the full range of the rating scale and to test whether participants were paying attention to the task.¹⁵ The presentation order of the items was randomized for each participant. Participants completed the questionnaire remotely using the Qualtrics survey platform (Qualtrics, Provo, UT). The task was an acceptability rating task. Participants were asked to read each text carefully and rate how good each scenario sounds from a scale of 1 *sehr seltsam* (“very strange”) to 7 *perfekt* (“perfect”). They indicated their response by clicking on stars below each item.

Participants

Based on a power calculation¹⁶, 60 participants were recruited via the Prolific platform (www.prolific.ac). Two participants were excluded based on their responses to the test fillers. The remaining 58 participants were all native speakers of German (35 male, 22 female, 1 nonbinary; no reported language-related disorders) and had a mean age of 31 years (range 18–66). All participants gave written informed consent.

Predictions

Agent conditions

The personal pronoun *er* should be the most suitable pronoun for referring to the Agent. In all three outlined scenarios, both *der* and *dieser* should elicit significantly lower scores than *er* because demonstratives tend to avoid reference to prominent referents.

Recipient conditions

The recipient is intermediate between the agent and the patient with respect to role prominence. If *er* is flexible in its reference, we expect high ratings in this condition (though possibly not as high as in the Agent conditions, if *er* truly prefers prominent antecedents). If *der* avoids reference to all but the least prominent candidate, then ratings for *der* should be equally low in this condition as in the Agent condition. If, on the other hand, *der* simply avoids reference to the most prominent candidate, then it should receive higher ratings in this condition than in the Agent condition. *Dieser* should pattern with *der* (Lange, 2016; Özden, 2016). However, low ratings for *dieser* would also be compatible with Zifonun et al.’s (1997) claim that *dieser* is reserved for referring to the last-mentioned antecedent.

Patient conditions

The patient has the lowest role prominence (Primus, 1999); it is also last-mentioned, and inanimate. Therefore, both *der* and *dieser* should receive high scores in this condition, which would be attributable either to role prominence or a last-mention effect for *dieser* (Zifonun et al., 1997). If *er* is truly flexible in its reference it should also receive high ratings here; however, if *er* prefers a higher prominence antecedent the scores for *er* will be significantly lower here than scores for *der* and *dieser*.

Data analysis

Ratings (1–7) for each item (including fillers) from each participant were converted to z-scores, as recommended in Schütze and Sprouse (2014), to account for variation in participants' use of the scale. To establish that the task was carried out correctly, mean z-scores for the unnatural fillers were compared with those from the normal fillers, using a linear mixed-effects model with a fixed effect of filler type (normal; unnatural) and random intercepts for participant and item. Additionally, each participant's mean z-score for the two filler types were compared to check if any participants should be excluded (exclusion criteria was having a higher mean z-score for the unnatural fillers than for the normal fillers). Two participants were excluded on the basis of their ratings for the unnatural fillers leaving 58 participants for the analysis.

Participants' z-scores for the experimental items were analyzed using a series of linear mixed-effects models. As the predictions relate primarily to the reference forms for each NP separately, and because there are slight differences in the materials between the three Referent conditions, a separate model was run for each level of Referent to ensure maximum comparability. The model contained the fixed factor Pronoun (*er*; *der*; *dieser*) and random intercepts for participant and item; the inclusion of random slopes for Pronoun were determined by assessment of singularity and by using the rePCA function in the package RePsychLing (Baayen et al., 2015).

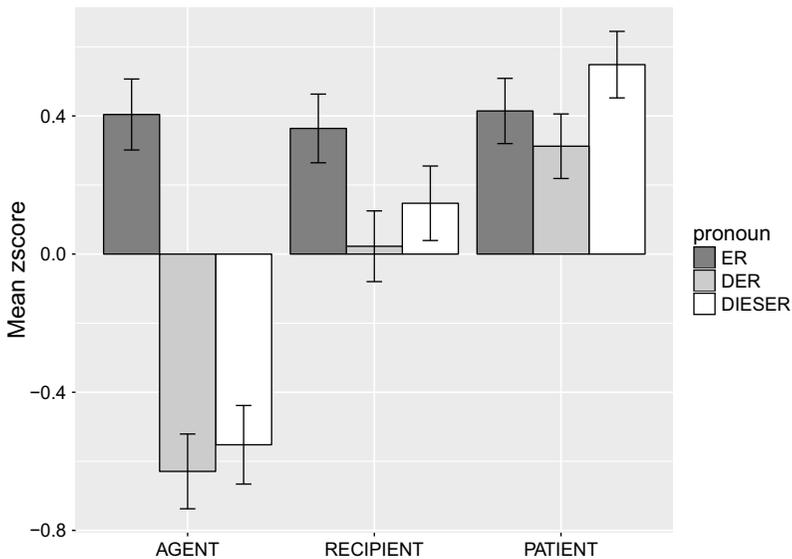
Further to the main analysis, ratings were also explored across the three Referent conditions. First, to explore the flexibility of the personal pronoun *er*, ratings for *er* across the three Referent conditions were compared. A linear mixed-effects model containing the fixed factor Referent was run on the ratings for *er*. Second, to compare the demonstrative forms *der* and *dieser* in the lower prominence Referent conditions, a model containing the fixed factors Pronoun (*der*; *dieser*) and Referent (Recipient; Patient) was run on the *der* and *dieser* ratings. For both models, random intercepts for participant and item were included and slopes were determined as described for the primary analysis above.

Results

All data, scripts, and model outputs are available in the supplemental materials linked to in the Data Availability section. Unnatural fillers received significantly lower ratings than normal fillers ($\beta = -1.54$, $SE = 0.14$, $t = -11.22$, $p < 0.001$). For the experimental items, mean ratings (raw scores and z-scores) per condition are shown in Table 2. Z-scores are plotted in Figure 1.

Table 2. Mean raw ratings and mean z-scores per Pronoun and Referent for Experiment 1a. Standard deviation (SD) are shown in parentheses

Referent	Pronoun	Mean raw ratings	Mean z-score
Agent	ER/SIE	5.70 (1.49)	.40 (.74)
	DER/DIE	3.68 (1.84)	-.63 (.83)
	DIESER/DIESE	3.80 (2.02)	-.55 (.86)
Recipient	ER/SIE	5.59 (1.55)	.36 (.72)
	DER/DIE	4.94 (1.75)	.02 (.76)
	DIESER/DIESE	5.13 (1.75)	.15 (.81)
Patient	ER/SIE	5.69 (1.53)	.41 (.70)
	DER/DIE	5.51 (1.53)	.31 (.69)
	DIESER/DIESE	5.97 (1.39)	.55 (.68)

**Figure 1.** Mean z-scores by condition set for Experiment 1a. Error bars show 95% confidence intervals, adjusted for within-subject designs as per Morey (2008).

Primary analysis

A likelihood ratio test of an overall model containing a Referent \times Pronoun interaction against the same model without the interaction showed a significant difference between models ($\chi^2(4) = 152.72$, $p < 0.0001$), justifying the separate inspection of the data at each level of Referent.¹⁷

Model outputs for each level of Referent are shown in Table 3. In the Agent conditions, ratings for *er* were significantly better than for *der* and *dieser*. This was also

Table 3. Fixed-effect model outputs by condition set, treatment coded (simple effects) for Experiment 1a. Baseline = ER

		<i>Estimate</i>	<i>Standard error</i>	<i>t-value</i>	<i>p-value</i>
Agent	Intercept	0.406	.076	5.372	<.001
	ER versus DER	-1.039	.067	-15.441	<.001
	ER versus DIESER	-0.945	.067	-14.042	<.001
Recipient	Intercept	0.373	.073	5.105	<.001
	ER versus DER	-0.342	.064	-5.365	<.001
	ER versus DIESER	-0.227	.064	-3.555	<.001
Patient	Intercept	0.409	.065	6.302	<.001
	ER versus DER	-0.103	.058	-1.784	.075
	ER versus DIESER	0.132	.058	2.281	.023

Table 4. Fixed-effect model output of demonstratives (*der* and *dieser*) across Referent levels Recipient and Patient, treatment coding for Experiment 1a. Baseline = DER and Recipient

	<i>Estimate</i>	<i>Standard error</i>	<i>t-value</i>	<i>p-value</i>
Intercept	.034	.063	0.538	.592
DER versus DIESER	.114	.070	1.635	.111
Recipient versus Patient	.270	.064	4.207	<.001
DIESER × Patient	.126	.091	1.389	.165

true for the Recipient conditions. In the Patient conditions, ratings for *er* were marginally better than for *der*. Ratings for *dieser*, however, were significantly better than for *er*.

Secondary analysis (preregistered)

While the primary analysis reflects the maximal comparability between the Pronoun conditions within one level of the Referent factor, it is also interesting to look at the pattern of responses across the three Referent conditions. In order to explore the flexibility of reference for the personal pronoun *er*, a model containing the fixed factor Referent was computed for the *er* data. This model showed no differences in ratings for *er* across the three antecedent conditions (all *ts* < 2). In order to explore the pattern of both demonstrative pronouns across the lower prominence antecedents, a model containing the fixed factors Pronoun (*der*; *dieser*) and Referent (Recipient; Patient) was computed. The output is shown in Table 4. This analysis showed that there was no significant difference between *der* and *dieser* in the Recipient condition. Ratings for *der* in the Patient condition are significantly better than in the Recipient condition; this is also true for *dieser*.

Table 5. Fixed-effect model output of demonstratives (DER and DIESER) across condition sets, forward-contrast coding for Experiment 1a. Baseline = DIESER

	<i>Estimate</i>	<i>Standard error</i>	<i>t-value</i>	<i>p-value</i>
Intercept	.052	.048	1.098	.277
Pronoun DER	-.149	.040	-3.760	<.001
Recipient – Agent	.693	.069	10.095	<.001
Patient – Recipient	.397	.069	5.783	<.001
Pronoun*Referent (Recipient – Agent)	-.037	.097	-0.376	.707
Pronoun*Referent (Patient – Recipient)	-.121	.097	-1.246	.213

Exploratory analysis (not-preregistered)

Finally, two more models were calculated to address specific questions arising from the results that were not covered by the preregistered analysis. First, to directly test whether *dieser* received the best ratings of all pronouns in the Patient conditions, the Patient model was rerun with *dieser* as the baseline. This model showed that ratings for *dieser* in the Patient condition were significantly better than ratings for *er* ($t = -2.127$) and for *der* ($t = -3.487$). Second, a model was calculated to further explore the results across all three referent conditions for the demonstrative pronouns, because the preregistered analysis does not test specifically whether ratings for the demonstratives significantly improved sequentially across the three Referent conditions. This question was addressed using a model with forward-contrast coding for the Referent conditions, which compares each level with the subsequent level. The output of this model is shown in Table 5. This model shows that overall ratings for *dieser* were better than ratings for *der*. Furthermore, ratings for both pronouns improved significantly from Agent to Recipient and from Recipient to Patient. The degree of improvement did not differ between *der* and *dieser*.

Experiment 1b

Before discussing the outcome of Experiment 1a, we present here the results of a follow-up experiment (Experiment 1b). In Experiment 1a, pronouns were made to refer unambiguously to one of the three potential antecedents through a combination of gender cues and plausibility cues. While reference to the two animate entities, the agent and recipient, was always disambiguated through gender, in the Patient conditions the gender of the pronoun matched all three NPs, but the event/state described in S2 was only compatible with an inanimate entity. This should have ruled out the agent and recipient as possible referents. However, it leaves open the possibility that the pronoun in this condition is nonetheless interpreted as referring to the agent or patient. If so, the event/state in S2, would be seen as incongruous (because it is not compatible with an animate referent). On this basis, participants could have given bad ratings in this condition for seemingly incongruent scenarios. While we think that this is unlikely, we cannot rule out the possibility that this took place, and the ratings therefore come under question.

We therefore conducted a follow-up experiment in which the gender cues fully disambiguate the reference in all three conditions. Example sentences are given in Table 6. The full set of materials can be found in the supplemental materials linked to in the Data Availability section.

Methods

All methodological and analysis details for Experiment 1b are identical to Experiment 1a, with the exception of the amended items as described above and a new set of participants. Participant details can be found in the supplemental materials linked to in the Data Availability section.

Results

The results from Experiment 1b are illustrated in Figure 2. A full summary of results for this experiment can be found in the supplemental materials linked to in the Data Availability section. Results for Experiment 1b were identical to the results for Experiment 1a, with the following exceptions: there was no significant difference between ratings for the three pronouns in the Recipient conditions; in the Patient conditions, there was no significant difference between *dieser* and *er*; the secondary and exploratory analysis showed that the improvement in scores for *der* between the Recipient and Patient conditions was only marginal.

Experiment 1 discussion

The results showed that for the Agent conditions, the personal pronoun *er* was rated significantly better than both the demonstrative pronouns. This aligns with the prediction that the personal pronoun is more suitable than the demonstrative pronouns for referring to a prominent antecedent and is compatible with all three scenarios outlined for demonstratives. In the Recipient conditions there is no difference between the three pronouns (Experiment 1b), but, as the exploratory analyses show, ratings for the demonstratives in Recipient conditions significantly improved compared to the Agent conditions while the ratings for the personal pronoun did not change. This result shows that using a demonstrative to refer to a medium prominence antecedent was not considered as bad as referring to the most prominent antecedent. But it is considered less good than referring to the least prominent antecedent, at least for *dieser*. This result is informative for two reasons. First, the description that demonstratives simply avoid the most prominent antecedent appears to be inadequate in light of our results. Second, the results show that the demonstrative pronouns do not simply seek an antecedent that is the least prominent and strongly reject all others. Rather, it suggests that sensitivity to the prominence of an antecedent is a gradable phenomenon. This is discussed in more detail in the general discussion. Finally, we turn to the prediction based on Zifonun et al that *dieser* is reserved for referring to the last-mentioned antecedent. Based on the current results, this appears to be too strong a formulation. *Dieser* was sensitive to prominence in a graded way across the three Referent conditions. This does not accord with the characterization of a form being reserved for the last-mentioned

Table 6. Materials for Experiment 1b (nine conditions)

	<i>Agent conditions</i> (pronoun refers to NP1)	<i>Recipient conditions</i> (pronoun refers to NP2)	<i>Patient conditions</i> (pronoun refers to NP3)
Context sentence	Der Eigentümer vermietete <i>the.M.SG.NOM owner.M rent.PST-3SG</i> der Anwohnerin die Parkfläche. <i>the.F.SG.DAT resident.F the.F.SG.ACC parking-space.F</i>	Die Eigentümerin vermietete <i>the.F.SG.NOM owner.F rent.PST-3SG</i> dem Anwohner die Parkfläche. <i>the.M.SG.DAT resident.M the.F.SG.ACC parking-space.F</i>	Der Eigentümer vermietete <i>the.M.SG.NOM owner.M rent.PST-3SG</i> dem Anwohner die Parkfläche. <i>the.M.SG.DAT resident.M the.F.SG.ACC parking-space.F</i>
ER/SIE	Er war froh über den <i>he be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Er war froh über den <i>he be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Sie war glücklicherweise schattig. <i>she be.PST-3SG fortunately shady</i>
DER/DIE	Der war froh über den <i>he.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Der war froh über den <i>he.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Die war glücklicherweise schattig. <i>she.DEM be.PST-3SG fortunately shady</i>
DIESER/ DIESE	Dieser war froh über den <i>he.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Dieser war froh über den <i>he.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Diese war glücklicherweise schattig. <i>she.DEM be.PST-3SG fortunately shady</i>
Translation	<i>"The (male) owner rented the parking space to the (female) resident. He was delighted about the long-term contract."</i>	<i>"The (female) owner rented the parking space to the (male) resident. He was delighted about the long-term contract."</i>	<i>"The (male) owner rented the parking space to the (male) resident. Fortunately it was a shady spot."</i>

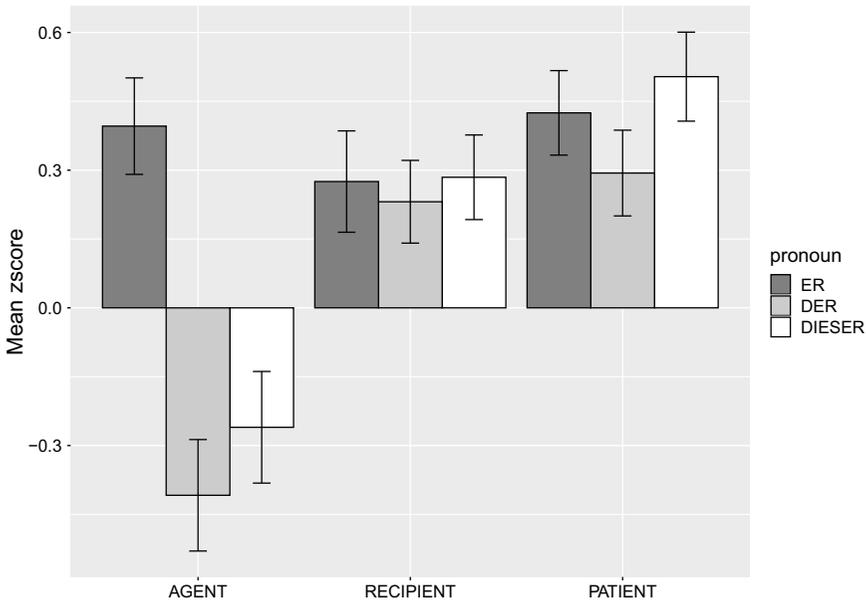


Figure 2. Mean z-scores by condition set for Experiment 1b. Error bars show 95% confidence intervals, adjusted for within-subject designs as per Morey (2008).

antecedent, but suggests that *dieser* is sensitive to distinctions in prominence when referring to antecedents that are not in final position. That being said, our results cannot tell us the degree to which the ratings were affected by the last-mention effect because the last-mentioned referent was also lowest in terms of its thematic role prominence (and was also inanimate). In order to further explore the importance of the last-mention position for *dieser*, we conducted Experiment 2.

It was noted by an anonymous reviewer that it is unclear whether the medial position is really more prominent than the final position (contra Primus, 1999); if so, then the demonstrative may have a preference for the medial rather than the final referent if prominence ranking is sensitive to position. This scenario appears to be ruled out for *dieser*, given the clear preference for the last-mentioned antecedent, although it may be a relevant factor for *der*, which did not show such a strong distinction between the recipient and patient. Antecedent order is explored further in Experiment 2.

Experiment 2

Hypotheses, predictions, participant numbers, and a data analysis plan (including exclusion criteria for participants) was registered in advance of carrying out the experiment on aspredicted.org (see the supplemental materials linked to in the Data Availability section). The purpose of this experiment was to examine the extent to which the final position contributes to resolution preferences of demonstratives, particularly *dieser*, against thematic role prominence. This was done by using

materials from Experiment 1a and changing the order of the NPs so that the least prominent antecedent (based on animacy and thematic role assignment, i.e., the patient) appeared in second position and the medium prominence antecedent (i.e., the recipient) appeared as last-mentioned.

It should be noted that swapping the order of the recipient and patient arguments results in a noncanonical order. The new scrambled order is, according to our intuitions, somewhat marked but not strongly ungrammatical. There is a great deal of discussion in the theoretical literature about the motivation for scrambled orders such as the one we are using for Experiment 2 (see Müller, 1999 for an overview). Furthermore, some authors have argued that there is an animacy constraint placing animate arguments before inanimate ones (e.g., Fanselow, 2003; Lenerz, 1977; Vogel & Steinbach, 1998). What is important to our study, however, is the relative prominence of the three arguments, that is whether changing the argument order to PATIENT – RECIPIENT impacts the relative prominence of the arguments. To our knowledge, there is no discussion in the literature of the relative prominence of the arguments as a result of the scrambled order. However, final position is typically seen as a focus position, so it is possible that moving the patient to the medial position, leaving the recipient in the final position, puts the recipient in focus (Fanselow, 2003 and references therein). If edge positions confer extra prominence (Himmelfmann & Primus, 2015), this would mean that the patient (now in NP2 position) remains the lowest ranked in terms of prominence. Thus, in Experiment 2, the (relative) prominence status of the arguments remains the same as in Experiment 1, but the linear order has changed. The lowest prominence argument is now in NP2 position.

An alternative scenario is that putting the recipient in focus confers a more topical status to the other two referents, and thus makes them more suitable than the recipient as referents for the pronoun (this is the basis of the so-called “anti-focus” effect, see Colonna et al., 2015; de la Fuente, 2015; Patterson et al., 2017).¹⁸ This is why it is critical to retain the other two pronouns in this experiment and not just test *dieser* alone. Any changes to the availability of particular antecedents based on the altered information structure should be reflected in responses for *er* and *der*, but these two pronouns should not be affected by a strong preference for the final position *per se*.

To avoid ambiguity between genitive and dative case when feminine NPs were placed in final position, the gender of the pronouns and antecedents had to be manipulated differently from Experiment 1, however, this was counterbalanced in other conditions so that there was still an even split between masculine and feminine pronouns.¹⁹ The same factors (Pronoun, Referent) were manipulated as in Experiment 1 to give nine conditions as before. Sample materials are shown in Table 7.

Procedure

The design and procedure were as per Experiment 1. An additional nine filler items, which started with masculine NPs and resembled the experimental items were added in order to counterbalance the experimental items that always start with a feminine NP.

Table 7. Sample materials in nine conditions for Experiment 2

	<i>Agent conditions</i> (pronoun refers to NP1)	<i>Patient conditions</i> (pronoun refers to NP2)	<i>Recipient conditions</i> (pronoun refers to NP3)
Context sentence	Die Eigentümerin vermietete <i>the.F.SG.NOM owner.F rent.PST-3SG</i> die Parkfläche dem Anwohner. <i>the.F.SG.ACC parking-space.F the.M.SG.DAT resident.M</i>	Die Eigentümerin vermietete <i>the.F.SG.NOM owner.F rent.PST-3SG</i> die Parkfläche dem Anwohner. <i>the.F.SG.ACC parking-space.F the.M.SG.DAT resident.M</i>	Die Eigentümerin vermietete <i>the.F.SG.NOM owner.F rent.PST-3SG</i> die Parkfläche dem Anwohner. <i>the.F.SG.ACC parking-space.F the.M.SG.DAT resident.M</i>
ER/SIE	Sie war froh über den <i>she be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Sie war glücklicherweise schattig. <i>she be.PST-3SG fortunately shady</i>	Er war froh über den <i>he be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>
DER/DIE	Die war froh über den <i>she.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Die war glücklicherweise schattig. <i>she.DEM be.PST-3SG fortunately shady</i>	Der war froh über den <i>he.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>
DIESER/DIESE	Diese war froh über den <i>she.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>	Diese war glücklicherweise schattig. <i>she.DEM be.PST-3SG fortunately shady</i>	Dieser war froh über den <i>he.DEM be.PST-3SG delighted over the.M.SG.ACC</i> langfristigen Vertrag. <i>long-term.M.SG.ACC contract.M</i>
Translation	"The (female) owner rented the parking space to the (male) resident. She was delighted about the long-term contract."	"The (female) owner rented the parking space to the (male) resident. Fortunately it was a shady spot."	"The (female) owner rented the parking space to the (male) resident. He was delighted about the long-term contract."

Participants

Sixty participants were recruited via the Prolific platform (www.prolific.ac). They all gave written informed consent. Two were excluded because they were not native German speakers. One further participant was excluded based on responses to the unnatural fillers. The remaining 57 participants (32 male, 25 female; no reported language-related disorders) had a mean age of 30 years (range 18–66).

Predictions

In comparison to Experiment 1, we expected ratings for Experiment 2 to be slightly lower due to the noncanonical constituent ordering. *Er* was shown to be flexible in its reference in Experiment 1 and should therefore score well in all three conditions of Experiment 2. Predictions for the demonstratives *der* and *dieser* are as follows.

Agent conditions

Der and *dieser* should elicit low scores in this condition because they tend to avoid reference to prominent referents, as per Experiment 1.

Patient conditions

According to the last-mention account (Zifonun et al., 1997), scores for *dieser* should be low (comparable to the scores for *dieser* in the Agent condition). Conversely, the prominence hierarchy account predicts that *der* and *dieser* are sensitive to the hierarchy from the thematic roles, which would lead to very high ratings for *der* and *dieser* here. Ratings should be as high or higher than ratings for *er*, and comparable to *der* and *dieser* ratings for the Patient condition in Experiment 1. If both constituent order and prominence are taken into account for the interpretation of *dieser*, higher ratings based on thematic role prominence should be somewhat tempered by the medial placement of the constituent. If *der* is more sensitive to prominence than constituent order, we expect high ratings in this condition, comparable to the *der* ratings for the patient in Experiment 1.

Recipient conditions

According to the last-mention account (Zifonun et al., 1997), *dieser* should receive high ratings in this condition, possibly higher than ratings for *er*, as per Experiment 1. According to the Prominence hierarchy account, *der* and *dieser* should receive lower ratings than *er* here but not as low as in the Agent condition. If both constituent order and prominence play a role, scores for *dieser* should fall somewhere between the level of recipient and patient scores from Experiment 1. If *der* is sensitive only to the prominence hierarchy it should get lower scores here than in the Patient conditions.

Data analysis

The data were analyzed as in Experiment 1.

Table 8. Mean raw scores and mean z-scores per pronoun and antecedent for Experiment 2. Standard deviation (SD) are shown in parentheses

Antecedent	Pronoun	Mean raw ratings	Mean z-score
AGENT	ER/SIE	4.19 (1.90)	-.09 (.85)
	DER/DIE	3.26 (1.77)	-.58 (.80)
	DIESER/DIESE	3.42 (1.84)	-.51 (.84)
PATIENT	ER/SIE	3.93 (1.76)	-.23 (.76)
	DER/DIE	3.72 (1.75)	-.37 (.73)
	DIESER/DIESE	3.98 (1.93)	-.22 (.82)
RECIPIENT	ER/SIE	4.12 (1.87)	-.14 (.87)
	DER/DIE	4.17 (1.93)	-.12 (.87)
	DIESER/DIESE	4.29 (2.01)	-.05 (.93)

Results

All data, scripts, and model outputs can be found in the supplemental materials linked to in the Data Availability section. One participant rated the catch fillers slightly better than the normal fillers; this participant was removed from the analysis. All other participants gave the catch fillers lower ratings than the normal fillers; group mean ratings were 3.11 (*SD* 1.81) for the catch fillers and 5.55 (*SD* 1.72) for the normal fillers (*z*-scores -0.68 and 0.60), which was significant in the LME model ($\beta = -1.28$, *SE* = 0.15, $t = -8.38$, $p < 0.001$). For the experimental items, mean ratings (raw scores and *z*-scores) per condition are shown in Table 8. *Z*-scores are plotted in Figure 3.

Primary analysis (preregistered)

A likelihood ratio test of an overall model containing a Referent \times Pronoun interaction against the same model without the interaction showed a significant difference between models ($\chi^2(4) = 40.81$, $p < 0.0001$), justifying the separate inspection of the data at each level of Referent.²⁰

Model outputs for each level of Referent are shown in Table 9. In the Agent condition, ratings for *er* were significantly better than for *der* and *dieser*. In the Patient condition, ratings for *er* were better than for *der* but not for *dieser*. In the Recipient condition, there were no significant differences between ratings.

Secondary analysis (preregistered)

In order to explore the flexibility of reference for the personal pronoun *er*, a model containing the fixed factor Referent was computed for the *er* data. This model showed that *er* was rated significantly lower in the Patient condition when compared to the Agent condition ($t = -2.035$). There was no difference between ratings in the Agent and Recipient conditions ($t = -0.787$). In order to explore the pattern of both demonstrative pronouns across the three Referent conditions, a model

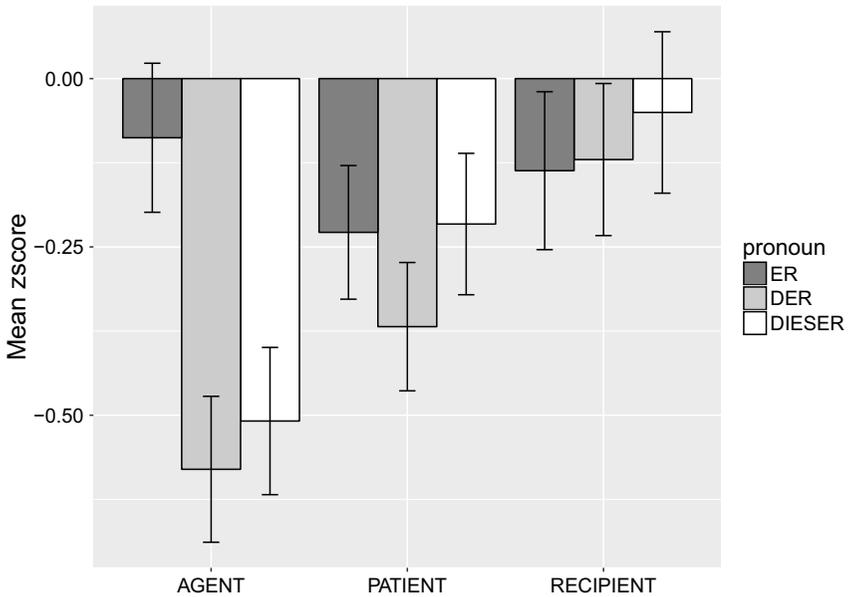


Figure 3. Mean z-scores by condition set for Experiment 2. Error bars show 95% confidence intervals, adjusted for within-subject designs as per Morey (2008).

containing the fixed factors Pronoun (*der*; *dieser*) and Referent (Agent; Patient; Recipient) was computed. The output is shown in Table 10. This analysis showed no difference in ratings between *der* and *dieser* in the Agent conditions. Both *der* and *dieser* received better ratings when referring to the Patient compared to the Agent, and when referring to the Recipient compared to the Agent.

Exploratory analysis (not-preregistered)

Finally, a model for the demonstratives with forward-contrast coding for the Referent conditions was computed as per Experiment 1. The output of this model is shown in Table 11. This model shows that overall ratings for *dieser* were better than ratings for *der*. Furthermore, ratings for both pronouns improved significantly from Agent to Patient and from Patient to Recipient. The degree of improvement did not differ between *der* and *dieser*.

Experiment 2 discussion

We predicted that the ratings for Experiment 2 would be slightly lower than for those in Experiment 1, due to the noncanonical ordering of the constituents. This was indeed the case, as can be seen from the z-scores which were all negative, meaning that the experimental items in Experiment 2 were given lower ratings than the fillers in the same experiment (which had a canonical argument order).

Table 9. Fixed-effect model outputs by condition set (Agent, Patient, and Recipient), treatment coded (simple effects) for Experiment 2. Baseline = ER

		<i>Estimate</i>	<i>Standard error</i>	<i>t-value</i>	<i>p-value</i>
AGENT	Intercept	-.092	.074	-1.233	.220
	ER versus DER	-.492	.070	-7.056	<.001
	ER versus DIESER	-.418	.070	-5.983	<.001
PATIENT	Intercept	-.233	.074	-3.127	.002
	ER versus DER	-.138	.062	-2.222	.027
	ER versus DIESER	.020	.062	0.319	.750
RECIPIENT	Intercept	-.145	.097	-1.501	.138
	ER versus DER	.016	.068	0.241	.810
	ER versus DIESER	.093	.068	1.363	.173

Table 10. Fixed-effect model output of demonstratives (DER and DIESER) across all antecedent levels, treatment coding for Experiment 2. Baselines = DER; AGENT

	<i>Estimate</i>	<i>Standard error</i>	<i>t-value</i>	<i>p-value</i>
Intercept	-.582	.078	-7.497	<.001
DER versus DIESER	.073	.070	1.045	.297
AGENT versus PATIENT	.215	.069	3.144	.002
AGENT versus RECIPIENT	.456	.068	6.673	<.001
DIESER: AGENT versus PATIENT	.087	.097	0.900	.368
DIESER: AGENT versus RECIPIENT	.0003	.097	0.003	.998

Table 11. Fixed-effect model output of demonstratives (DER and DIESER) across condition sets with forward-contrast coding for Experiment 2. Baseline = DER

	<i>Estimate</i>	<i>Standard error</i>	<i>t-value</i>	<i>p-value</i>
Intercept	-.358	.068	-5.234	<.001
Pronoun DIESER	.102	.040	2.577	.010
AGENT versus PATIENT	.215	.069	3.137	.002
PATIENT versus RECIPIENT	.241	.069	3.521	<.001
DIESER: AGENT versus PATIENT	.087	.097	0.898	.369
DIESER: PATIENT versus RECIPIENT	-.087	.097	-0.895	.371

Our interest lies in examining the difference between the conditions in the experiment.

We expected that the personal pronoun *er* would be as flexible as it was in Experiment 1, but this prediction was not completely borne out; *er* was given significantly lower ratings in the Patient condition. This finding was unexpected. It is possible that a combination of the low prominence of the patient role and the medial position in the sentence renders the patient in this experiment particularly non-prominent, making *er* a less suitable referential form. However, if this was the case we would expect the demonstrative pronouns to do better in this condition; in fact, *der* received even lower ratings than *er*, and *dieser* ratings were similar to those for *er*. This suggests perhaps that using any type of pronoun is sub-optimal for this referent, because participants were not expecting the second sentence to continue with that referent. We speculate that using the full NP to refer to the patient in this context may have been more felicitous. In this situation, it would be advantageous to have insight production and interpretation data in the style of Kehler and Rohde's experiments (e.g., Kehler & Rohde, 2013), to be able to separate the likelihood of continuing with a particular entity from the likelihood of using a pronoun.

Relatedly, the movement of the recipient to the final position may have given it a boost in prominence. For example, Himmelmann & Primus (2015) discuss edge placement as a potential prominence-lending cue. In this regard, initial and final arguments are more privileged than an argument in medial position. This may have two causes: either edge placement is a purely serial criterion that marks the borders of certain grammatical structure, or edge placement is motivated information structurally by representing the topic or focus. In our experiment, this placement may have increased the expectation that the second sentence would continue with the recipient, leading to higher ratings in the Recipient condition and conversely to lower ratings in the Patient condition.

An alternative explanation of the lower ratings for the Patient condition, also motivated by information structure, is that the inanimate patient is licensed to move in front of the recipient if it is a topic.²¹ If demonstratives strongly avoid topics then the lower ratings for demonstratives could be explained by an anti-topic preference. However, if the patient is a topic, we would not expect ratings for the personal pronoun also to be low in this condition.

The existence of several conflicting interpretations for the results of Experiment 2, which was designed principally to test predictions based on Zifonun et al.'s (1997) claim that *dieser* is reserved for referring to the last-mentioned referent, highlights a drawback in the design of this experiment. Untangling the information-structural influences underlying the behavior would require several further targeted experiments, which is outside the scope of the current study. This limitation does not prevent us from extracting some useful and interpretable findings.

The primary analysis showed that when referring to the Agent, the personal pronoun *er* was rated significantly better than both the demonstrative pronouns. This replicates our finding from Experiment 1 and confirms that the personal pronoun is more suitable than the demonstrative pronouns for referring to the most prominent antecedent. Results from the Patient and Recipient conditions give us some indication as to whether the demonstrative pronouns are sensitive to prominence from role hierarchies, to order of mention (last mention) or both factors. Reference

to the Recipient, which is in the final position but has medium role prominence (possibly even boosted by its edge position as discussed above), elicit the best ratings for both demonstrative pronouns. This points to a strong role of order of mention (possibly motivated by information structural factors) and suggests, importantly, that the demonstratives are very well suited to referring to a last-mentioned entity even when it is not the referent with the lowest role prominence. However, the ratings in the Patient condition are also informative here. Unlike the Agent conditions, there is no difference in ratings between *dieser* and *er* in the Patient conditions; such a difference would be expected if *dieser* was **only** suitable for last-mentioned referents. Additionally, the ratings for both *dieser* and *der* are significantly better in the Patient condition than in the Agent condition; again, this is unexpected on a strong last-mention account, where all other options should be strongly rejected. These results are discussed further in the general discussion.

General discussion

Many studies of pronoun resolution preferences are limited to contexts in which two candidate antecedents are presented, leaving open the question of how pronoun resolution plays out when more than two possible referents are available in the discourse context. This is particularly important to examine given that many pronoun resolution studies rely on the assumption of an ordered set of potential candidate antecedents. While demonstratives appear to prefer less prominent antecedents, it is unclear if they reject all but the least prominent antecedent, or if they reject only the most prominent antecedent. It is not possible to assess these possibilities using only two potential antecedents. In the current study, we addressed this limitation by examining pronoun interpretation preferences in ditransitive contexts in which there are three potential antecedents. We examined both personal and demonstrative pronouns in German because they are assumed to be sensitive to the prominence of potential antecedents, while relatively little is known about the interpretations of *dieser* in particular.

The main finding from our experiments was a graded sensitivity to referent prominence for both demonstrative pronouns. This was seen in the response patterns for the demonstratives, particularly *dieser*, in the lower prominence conditions (Recipient and Patient), which showed differences in both experiments, further confirmed in the exploratory analyses. This finding is important because it furthers our understanding about the preference of demonstrative pronouns for less prominent antecedents. From previous studies, it was not clear whether demonstratives avoid reference to all but the least prominent candidate, or simply avoid reference to the most prominent candidate. It seems that neither of these characterizations is complete, since they imply that the remaining candidates are considered to be equal. Our results show that in fact, prominence gives rise to more nuanced preferences among the three candidate antecedents. However, the outcome of Experiment 2 demonstrated that referring to the last-mentioned candidate is preferable for both demonstrative forms even when the last-mentioned candidate is not the least prominent candidate in terms of thematic role. Given that the literature on demonstratives in German has largely concentrated on prominence of grammatical or

thematic roles or on notions such as topichood rather than positional preferences, it was surprising to find such a strong influence of position in our results. It seems that the description in Zifonun et al. (1997) is partially correct. However, despite the preference for demonstratives to refer to referents in final position, the medially placed referents were rated better than the agent in both experiments, confirming that the demonstrative preferences cannot simply be described as a final position preference, but are also influenced by a referent's role prominence.

We noted in the introduction that three studies had found that *dieser* referred preferentially to the patient even when it was not in final position (Fuchs & Schumacher, 2020; Lange, 2016; Özden, 2016); on the surface this is a different pattern than was found in our data where final position **was** important. But in these studies, only two potential antecedents were tested, an agent and a patient, in both agent-before-patient and patient-before-agent order. In the latter order, the agent was in final position, which was never the case in the current study, so we cannot directly compare the two patterns. In our experiments referring to an agent with a demonstrative pronoun always received the lowest ratings. It remains to be seen whether using a demonstrative to refer to an agent in final position in ditransitive contexts would result in equally low ratings. But it should be noted that this would result in a very marked change to the information structure of the context sentence.

By using ditransitive contexts, we made more than two potential antecedents available in the context, which also allowed us to test whether there is a division of labor between *der* and *dieser* with respect to their interpretation preferences. This is especially important given the lack of a significant literature on *dieser*. The results from Experiment 1 show that both forms have a similar pattern with respect to the three antecedents, which is not suggestive of a strong division of labor. However, in Experiment 1b, the difference between recipient and patient was only marginally significant for *der*. This could indicate that *der* is less strongly influenced by a final position preference than *dieser*. Furthermore, in both experiments, *dieser* received better scores than *der* overall. This may be due to *dieser* being the more formal counterpart of *der* (Patil et al., 2020); given that the experiment was conducted as judgments on written material, this may have led to participants assigning higher ratings to *dieser*. The language modality may also play a role, with *der* being more suitable to spoken language. The interaction of register and modality on the occurrence and resolution of demonstrative forms is a topic for future research (Patil et al., 2020). Another possibility (albeit speculative) is that *dieser* has a stronger singling out function that is particularly suitable to contexts in which several antecedents are available.

As regards the division of labor between personal and demonstrative forms, which has been found in previous studies, we made the basic assumption that in German, personal and demonstrative pronouns are sensitive to the same set of factors affecting prominence, as was the case in Schumacher et al. (2015, 2016). While our study was not set up directly to test this, we can see that the final position may well be more important for demonstratives, in particular *dieser*, than for personal pronouns. However, future studies further probing the information-structural motivation for the importance of final position may sharpen our insight in this direction.

An important difference between this study and previous ones is the use of a rating task instead of an antecedent choice which is more commonly used in studying pronoun resolution preferences. We find the rating method to be very useful for assessing pronoun resolution preferences because it reveals not only the most preferred option in a particular condition, but also more subtle information such as relative acceptability of dispreferred alternatives that are not outright rejected. In a forced-choice or sentence completion task, a lot of data points are gathered about the most preferred option, while very little data are gathered about less preferred options. It is difficult to tell apart options that are completely rejected from those that are simply dispreferred, as the behavior (choosing the preferred answer option) is the same in both cases. For the assessment of the hypotheses we presented, it was important to have sufficient data about less preferred antecedents as well. That being said, the judgment task does not tap into a participants' preferred interpretations as directly as a forced-choice or sentence completion task. Nevertheless, we think that having more information about lesser preferred candidates is also a valid point of investigation, as readers and listeners are often confronted with interpretation situations in which the preferred candidate is not available.

As far as the relational notion of prominence of individual referents is concerned, the experiments support the assumption of an ordered set of potential candidate antecedents that give rise to graded preferences. This is particularly evident from the ratings for *dieser*, which strengthen the prominence-lending function of thematic role by turning down the agent as a potential candidate for coreference, while at the same time showing a gradient resolution pattern with respect to the two non-agent roles. The prominence framework offers a broader perspective on reference resolution by making available an ordered set of referential candidates. Rather than merely focusing on the most prominent entity from the set, the current data indicate that the full set is utilized during reference resolution. Overall, this suggests that gradience should be taken more seriously in research on reference resolution. Furthermore, since the prominence framework is geared toward a general organizational principle of the language system, gradience should also be accounted for when prominence relations are considered at other levels of linguistic description (see, e.g., Roessig et al., 2019 for prosody; Kretzschmar et al., 2019 for an agentivity cline).

Finally, the data indicate a considerable contribution of order – or rather edge placement. However, we cannot draw strong conclusions about the information-structural triggers without further testing. The pattern is in line with Himmelmann and Primus's (2015) discussion of edge as a potential prominence-lending cue which would be a fruitful avenue for further research. In the current case, the sentence boundary serves as a privileged position (possibly contributing information structural cues triggered by the marked argument order in Experiment 2). The medial position in turn may function as a landing site for entities that are demoted in terms of their accessibility, which would explain why even the personal pronoun, which was otherwise very flexible, received lower ratings in this position for Experiment 2. It is also noteworthy that the personal pronoun showed a high degree of flexibility in its antecedent selection. This suggests that using personal pronouns to identify the features that lend prominence to an entity may not be the most promising avenue for investigation.

In sum, presenting more than two potential antecedents for German pronouns enables us to better characterize their interpretation preferences. The personal pronoun is highly flexible, being equally acceptable when referring to any of the three candidates in canonical conditions, while not always being the preferred pronoun. The interpretation of demonstrative pronouns follows a graded sensitivity to referent prominence, ranging from lower acceptability when referring to highly prominent antecedents to the highest acceptability for the least prominent antecedents. Importantly, the ratings for the medium prominence candidate differed from both the most prominent and the least prominent candidate, revealing a graded pattern. Models for reference resolution should thus capture the gradient nature underlying competing referential candidates.

Data Availability. Supplementary data and materials can be found on OSF: <https://osf.io/zrkvx/>.

Acknowledgments. The authors would like to thank Hilde Penner, Julia Veth, Celine Cuma, and Magdalena Repp for their help with item construction and experiment preparation.

Financial Support. The research for this paper has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 281511265 – SFB 1252 “Prominence in Language” in the project C07 “Forward and backward functions of discourse anaphora” at the University of Cologne, Department of German Language and Literature I, Linguistics.

Conflict of Interest. The authors report no conflict of interest.

Notes

1. Note that we do not contend that the use of a particular referring expression confers a particular prominence status – in our view, the prominence ranking is determined independently of referential expressions.
2. Centering Theory, developed in computational linguistics/artificial intelligence, is a framework for tracking prominence as a discourse unfolds, which informs the choice of referential expression. It has been adopted widely in psycholinguistics as a testable mechanism for pronoun resolution preferences based on discourse prominence.
3. Furthermore, In German, *der* and *dieser* refer to entity denoting referents and not to propositional content, contrary to *this* and *that* in English (as investigated in Çokal et al., 2018; see also Brown-Schmidt et al., 2005 for reference to “composite entities” such as a cup on the saucer).
4. The categorization of thematic roles varies considerably from a large set of roles defined over verb-specific properties to accounts with only two or three generalized semantic roles. The postulation of proto-roles goes back to Dowty (1991) who offers a feature-based account of agentivity with two proto-roles: proto-agents show volitional involvement, causation, sentience (and/or perception), and autonomous movement; proto-patients undergo change of state, are volitionally and causally affected, and stationary (relative to other characters). Primus (1999) revisits Dowty’s features and proposes three proto-roles – agent, recipient and patient – which are presented in more detail below.
5. For example, *Der Arzt will den Patienten untersuchen* (“The doctor [AGENT] wants to examine the patient [PATIENT]”).
6. This is not possible for the verb *geben* in German, but a prepositional object construction is available for many other verbs and is not restricted to transfer events (Adler, 2011).
7. In, for example, *The grandpa buys the granddaughter a scooter*, the benefactive (*the granddaughter*) undergoes a change in possession and/or benefits from the action caused by the proto-agent (*the grandpa*) and at the same time there is a possessor-possessed dependency between the benefactive and the proto-patient (*scooter*).
8. Indeed, an EEG experiment (Uzun, 2015b) using similar materials resulted in different profiles for the different pronoun types (an N400 for *der* and a late positivity for *dieser*), suggesting different underlying behaviour for the different pronoun types.

9. It should be noted, however, that it is not entirely clear that the final position is the least important, from an information structural perspective (we thank an anonymous reviewer for highlighting this issue). The final position may alternatively be seen as a focus position and thereby be more important than medial position even in canonical contexts. Furthermore, Himmelmann and Primus (2015) consider linguistic edges as potential prominence-lending cues and thereby the final (and initial) position is more important than the medial position. For the purposes of Experiment 1, the canonicity combined with the thematic role outweigh a strong focus reading, but we return to this issue for Experiment 2 where the information structure may play a more important role.

10. Agent – Recipient – Patient is the canonical argument order in German.

11. Ditransitive verbs that sound natural with three animate entities are less common, and we thus opted for more natural event representations in S1. In addition, the patient argument is already the lowest ranked argument according to the thematic role hierarchy and the animacy contrast works in the same direction (animate entities being more agentive than inanimate entities). As noted above, we assume that animacy is an epiphenomenon of agentivity and its feature specification (García García et al., 2018; Schumacher, 2018).

12. Adler (2011) identified five classes of ditransitive verbs in German: “Geben (give)”, “Verkaufen (sell)”, “Schicken (send)”, “Stehlen (steal)”, and “Werfen (throw)”. The “werfen” verbs have very different syntactic properties from the others so were not included in the study. We included a selection of verbs from the other four classes, as well as benefactives.

13. Note that the items included verbs such as *stehlen* (“to steal”), whose dative argument might not be traditionally analyzed as having a recipient role, since someone experiences a loss rather than receives something. However, under the proto-roles analysis (Dowty, 1991; Primus, 1999) that we follow, the dative argument for such verbs is subsumed under the proto-recipient role.

14. The plausibility pretest is described in the supplemental materials linked to in the Data Availability section.

15. An example from the implausible fillers: “Obwohl der Jäger das Geheimnis kannte, schwieg er. Denn er ist so Gesprächig.” (Although the hunter knew the secret, he kept silent. Because he is so talkative.)

16. The power calculation for Experiment 1 can be found in the supplemental materials linked to in the Data Availability section.

17. Note that this test was requested at review and is therefore not included in the pre-registered analysis.

18. We thank an anonymous reviewer for this suggestion.

19. The final argument in example (i) is case ambiguous between the dative (ii) and the genitive (iii). Such an ambiguity can be avoided by using a masculine NP, which is case disambiguated (dative: *dem Dozenten*; genitive: *des Dozenten* (the lecturer)).

(i) *Der Student stahl den Laptop der Dozentin.*

the student (nom.masc.) stole the laptop (acc.masc.) the lecturer (dat/gen.fem.)

(ii) *The student stole the laptop from the lecturer.*

(iii) *The student stole the laptop of the lecturer.*

20. Note that this test was requested at the review and is therefore not included in the preregistered analysis.

21. We thank an anonymous reviewer for this suggestion.

References

- Abraham, W. (2002). Pronomina im Diskurs: Deutsche Personal- und Demonstrativpronomina unter ‘Zentrierungsperspektive.’ *Grammatische Überlegungen zu einer Teiltheorie der Textkohärenz. Sprachwissenschaft*, 27, 447–491.
- Adler, J. (2011). *Dative alternations in German: The argument realization options of transfer verbs* [PhD Thesis]. Hebrew University.
- Ahrenholz, B. (2007). *Verweise mit Demonstrativa im gesprochenen Deutsch: Grammatik, Zweitspracherwerb und Deutsch als Fremdsprache*. De Gruyter.
- Aissen, J. (1999). Markedness and Subject Choice in Optimality Theory. *Natural Language & Linguistic Theory*, 17(4), 673–711. <https://doi.org/10.1023/A:1006335629372>
- Aissen, J. (2003). Differential Object Marking: Iconicity vs. Economy. *Natural Language & Linguistic Theory*, 21(3), 435–483. JSTOR.

- Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., & Clifton Jr, C.** (2002). Null vs. Overt pronouns and the topic–focus articulation in Spanish. *Rivista Di Linguistica*, **14**, 151–169.
- Ariel, M.** (1990). *Assessing noun-phrase antecedents*. Routledge.
- Arnold, J. E.** (2001). The Effect of Thematic Roles on Pronoun Use and Frequency of Reference Continuation. *Discourse Processes*, **31**(2), 137–162. https://doi.org/10.1207/S15326950DP3102_02
- Arnold, J. E.** (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, **4**, 187–203.
- Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C.** (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, **76**(1), B13–B26. [http://dx.doi.org/10.1016/S0010-0277\(00\)00073-1](http://dx.doi.org/10.1016/S0010-0277(00)00073-1)
- Baayen, H., Bates, D., Kliegl, R., & Vasishth, S.** (2015). *RePsychLing: Data sets from Psychology and Linguistics experiments*.
- Bader, M., & Portele, Y.** (2019). The interpretation of German personal pronouns and d-pronouns. *Zeitschrift Für Sprachwissenschaft*, **38**(2), 155–190. <https://doi.org/10.1515/zfs-2019-2002>
- Barbosa, P., Duarte, M. E. L., & Kato, M. A.** (2005). Null Subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, **4**, 11–52.
- Bisle-Müller, H.** (1991). *Artikelwörter im Deutschen: Semantische und pragmatische Aspekte ihrer Verwendung*. Niemeyer.
- Bosch, P., & Hinterwimmer, S.** (2016). Anaphoric reference by demonstrative pronouns in German. In *Empirical Perspectives on Anaphora Resolution* (pp. 193–212). De Gruyter.
- Bosch, P., & Umbach, C.** (2007). Reference determination for demonstrative pronouns. In D. Bittner & N. Gagarina (Eds.), *Intersentential Pronominal Reference in Child and Adult Language: Proceedings of the Conference on Intersentential Pronominal Reference in Child and Adult Language* (Vol. 48). ZAS.
- Bosch, P., Katz, G., & Umbach, C.** (2007). The non-subject bias of German demonstrative pronouns. In *Anaphors in Text: Cognitive, formal and applied approaches to anaphoric reference* (pp. 145–164).
- Bosch, P., Rozario, T., & Zhao, Y.** (2003). Demonstrative pronouns and personal pronouns: German der versus er. *Proceedings of the EAFL 2003*. Workshop on the computational treatment of anaphora, Budapest.
- Bouma, G., & Hopp, H.** (2006). Effects of word order and grammatical function on pronoun resolution in German. In R. Artstein & M. Poesio (Eds.), *Ambiguity in Anaphora Workshop Proceedings* (pp. 5–13). University of Essex.
- Bouma, G., & Hopp, H.** (2007). Coreference preferences for personal pronouns in German. In D. Bittner & N. Gagarina (Eds.), *Intersentential pronominal reference in child and adult language* (Vol. 48, pp. 53–74). ZAS.
- Brown-Schmidt, S., Byron, D. K., & Tanenhaus, M. K.** (2005). Beyond salience: Interpretation of personal and demonstrative pronouns. *Journal of Memory and Language*, **53**(2), 292–313. <https://doi.org/10.1016/j.jml.2005.03.003>
- Carminati, M. N.** (2002). *The processing of Italian subject pronouns* [PhD Thesis]. University of Massachusetts.
- Çokal, D., Sturt, P., & Ferreira, F.** (2018). Processing of It and This in Written Narrative Discourse. *Discourse Processes*, **55**(3), 272–289. <https://doi.org/10.1080/0163853X.2016.1236231>
- Colonna, S., Schimke, S., & Hemforth, B.** (2012). Information structure effects on anaphora resolution in German and French: A crosslinguistic study of pronoun resolution. *Linguistics*, **50**(5), 991–1013.
- Colonna, S., Schimke, S., & Hemforth, B.** (2015). Different effects of focus in intra- and inter-sentential pronoun resolution in German. *Language, Cognition and Neuroscience*, **30**(10), 1306–1325. <https://doi.org/10.1080/23273798.2015.1066510>
- Cowles, H. W., Walenski, M., & Kluender, R.** (2007). Linguistic and cognitive prominence in anaphor resolution: Topic, contrastive focus and pronouns. *Topoi*, **26**(1), 3–18. <https://doi.org/10.1007/s11245-006-9004-6>
- de la Fuente, I.** (2015). *Putting pronoun resolution in context: The role of syntax, semantics, and pragmatics in pronoun interpretation* [PhD Thesis]. Université Sorbonne-Paris Cité and Université Paris Diderot – Paris 7.
- Dimitriadis, A.** (1996). When pro-drop language don't: Overt pronominal subjects and pragmatic inference. In L. M. Dobrin, K. Singer, & L. McNair (Eds.), *Proceedings of the CLS* (Vol. 32, pp. 33–47).
- Dowty, D.** (1991). Thematic Proto-Roles and Argument Selection. *Language*, **67**(3), 547–619.
- Ehlich, K.** (1983). Deixis und Anapher. In G. Rauh (Ed.), *Essays on deixis* (pp. 79–97). Narr.

- Ellert, M.** (2010). *Ambiguous pronoun resolution in L1 and L2 German and Dutch* [PhD Thesis]. Radboud University Nijmegen.
- Ellert, M., & Holler, A.** (2011). Semantic and Structural Constraints on the Resolution of Ambiguous Personal Pronouns—A Psycholinguistic Study. In I. Hendrickx, S. L. Devi, A. Branco, & R. Mitkov (Eds.), *Anaphora Processing and Applications* (pp. 157–170). Springer.
- Fanselow, G.** (2003). Free Constituent Order: A Minimalist Interface Account. *Folia Linguistica*, 37(1–2), 191–231. <https://doi.org/10.1515/flin.2003.37.1-2.191>
- Filiaci, F.** (2010). *Anaphoric preferences of null and overt subjects in Italian and Spanish: A cross-linguistic comparison* [PhD Thesis]. University of Edinburgh.
- Filiaci, F., Sorace, A., & Carreiras, M.** (2014). Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language, Cognition and Neuroscience*, 29(7), 825–843. <https://doi.org/10.1080/01690965.2013.801502>
- Fossard, M., Garnham, A., & Cowles, H. W.** (2012). Between anaphora and deixis . . . The resolution of the demonstrative noun phrase ‘that N’. *Language and Cognitive Processes*, 27(9), 1385–1404. <https://doi.org/10.1080/01690965.2011.606668>
- Fuchs, M., & Schumacher, P. B.** (2020). Referential shift potential of demonstrative pronouns—Evidence from text continuation. In A. Naess, A. Margetts, & Y. Treis (Eds.), *Demonstratives in Discourse* (pp. 185–213). Language Science Press.
- García García, M., Primus, B., & Himmelmann, N. P.** (2018). Shifting from animacy to agentivity. *Theoretical Linguistics*, 44(1–2), 25. <https://doi.org/10.1515/tl-2018-0002>
- Grosz, B. J., Joshi, A. K., & Weinstein, S.** (1995). Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*, 21(2), 203–225.
- Gundel, J. K., Hedberg, N., & Zacharski, R.** (1993). Cognitive Status and the Form of Referring Expressions in Discourse. *Language*, 69(2), 274–307.
- Haeseryn, W., Romijn, K., Geerts, G., De Rooij, J., & Van den Toorn, M. C.** (Eds.). (1997). *Algemene Nederlandse spraakkunst [General Dutch grammar]* (Vol. 1). M. Nijhoff.
- Hemforth, B., Konieczny, L., Scheepers, C., Colonna, S., Schimke, S., Baumann, P., & Pynte, J.** (2010). Language specific preferences in anaphor resolution: Exposure or gricean maxims? *32nd Annual Conference of the Cognitive Science Society*, 2218–2223. <https://hal.archives-ouvertes.fr/hal-01015048>
- Himmelmann, N. P., & Primus, B.** (2015). Prominence Beyond Prosody—A First Approximation. In A. De Dominicis (Ed.), *PS-prominences: Prominences in Linguistics* (pp. 38–58). DISUCOM Press.
- Hinterwimmer, S., & Bosch, P.** (2016). Demonstrative Pronouns and Perspective. In P. Grosz & P. Patel-Grosz (Eds.), *The Impact of Pronominal Form on Interpretation* (pp. 189–220). De Gruyter.
- Järvikivi, J., van Gompel, R. P. G., Hyönä, J., & Bertram, R.** (2005). Ambiguous Pronoun Resolution: Contrasting the First-Mention and Subject-Preference Accounts. *Psychological Science*, 16(4), 260–264. <https://doi.org/10.1111/j.0956-7976.2005.01525.x>
- Kaiser, E.** (2003). *The quest for a referent: A crosslinguistic look at reference resolution* [PhD Thesis]. University of Pennsylvania.
- Kaiser, E.** (2005). When salience isn't enough: Pronouns, demonstratives and the quest for an antecedent. In R. Laury (Ed.), *Minimal reference in Finnic: The use and interpretation of pronouns and zero in Finnish and Estonian discourse* (pp. 135–162). Suomalaisen Kirjallisuuden Seura.
- Kaiser, E.** (2011a). Salience and contrast effects in reference resolution: The interpretation of Dutch pronouns and demonstratives. *Language and Cognitive Processes*, 26(10), 1587–1624. <https://doi.org/10.1080/01690965.2010.522915>
- Kaiser, E.** (2011b). On the relation between coherence relations and anaphoric demonstratives in German. In I. Reich, E. Horch, & D. Pauly (Eds.), *Proceedings of Sinn und Bedeutung* (Vol. 15). Saarland University Press.
- Kaiser, E., & Trueswell, J. C.** (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2), 113–147. <https://doi.org/10.1016/j.cognition.2004.01.002>
- Kaiser, E., & Trueswell, J. C.** (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, 23(5), 709–748. <https://doi.org/10.1080/01690960701771220>
- Keating, G. D., Jegerski, J., & VanPatten, B.** (2016). Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition*, 19(1), 36–49. <https://doi.org/10.1017/S1366728914000418>

- Kehler, A., & Rohde, H.** (2013). A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics*, *39*(1–2), 1–37. <https://doi.org/10.1515/tl-2013-0001>
- Kehler, A., & Rohde, H.** (2017). Evaluating an Expectation-Driven Question-Under-Discussion Model of Discourse Interpretation. *Discourse Processes*, *54*(3), 219–238. <https://doi.org/10.1080/0163853X.2016.1169069>
- Kehler, A., Kertz, L., Rohde, H., & Elman, J. L.** (2008). Coherence and Coreference Revisited. *Journal of Semantics*, *25*(1), 1–44. <https://doi.org/10.1093/jos/ffm018>
- Kretzschmar, F., Graf, T., Philipp, M., & Primus, B.** (2019). An experimental investigation of agent prototypicality and agent prominence in German. In A. Gattnar, R. Hörnig, M. Störzer, & S. Featherston (Eds.), *Proceedings of Linguistic Evidence 2018: Experimental data drives linguistic theory* (pp. 101–123). Universität Tübingen.
- Lange, A.** (2016). *Interpretation von Demonstrativpronomen im Kontext von Dativ-Experienzer-Konstruktionen* [Bachelor thesis]. University of Cologne.
- Lenzer, J.** (1977). *Zur Abfolge nominaler Satzglieder im Deutschen*. Narr.
- Morey, R.** (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Quantitative Methods for Psychology*, *4*(2), 61–64. <https://doi.org/10.20982/tqmp.04.2.p061>
- Müller, G.** (1999). Optimality, markedness, and word order in German. *Linguistics*, *37*(5), 777. <https://doi.org/10.1515/ling.37.5.777>
- Özden, F.** (2016). *Interpretationspräferenzen von „dieser“: Eine empirische Studie* [Bachelor thesis]. University of Cologne.
- Patil, U., Bosch, P., & Hinterwimmer, S.** (2020). Constraints on German diese demonstratives: Language formality and subject-avoidance. *Glossa: A Journal of General Linguistics*, *5*(1), 14.
- Patterson, C., & Schumacher, P. B.** (2020). The timing of prominence information during the resolution of German personal and demonstrative pronouns. *Dialogue & Discourse*, *11*(1), 1–39.
- Patterson, C., Esaulova, Y., & Felser, C.** (2017). The impact of focus on pronoun resolution in native and non-native sentence comprehension. *Second Language Research*, *33*(4), 403–429. <https://doi.org/10.1177/0267658317697786>
- Portele, Y., & Bader, M.** (2016). Accessibility and Referential Choice: Personal Pronouns and D-pronouns in Written German. *Discours. Revue de Linguistique, Psycholinguistique et Informatique. A Journal of Linguistics, Psycholinguistics and Computational Linguistics*, *18*, 1–39.
- Primus, B.** (1999). *Cases and thematic roles – Ergative, accusative and active*. Niemeyer.
- Roessig, S., Mücke, D., & Grice, M.** (2019). The dynamics of intonation: Categorical and continuous variation in an attractor-based model. *PLoS ONE*, *14*(5), e0216859. <https://doi.org/10.1371/journal.pone.0216859>
- Rohde, H.** (2019). Pronouns. In C. Cummins & N. Katsos (Eds.), *The Oxford Handbook of Experimental Semantics and Pragmatics*. Oxford University Press.
- Rohde, H., & Kehler, A.** (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*(8), 912–927. <https://doi.org/10.1080/01690965.2013.854918>
- Rohde, H., Kehler, A., & Elman, J. L.** (2006). Event structure and discourse coherence biases in pronoun interpretation. In R. Sun (Ed.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates.
- Schumacher, P. B.** (2018). On type composition and agentivity. *Theoretical Linguistics*, *44*(1–2), 81. <https://doi.org/10.1515/tl-2018-0007>
- Schumacher, P. B., Backhaus, J., & Dangl, M.** (2015). Backward- and Forward-Looking Potential of Anaphors. *Frontiers in Psychology*, *6*, 1746. <https://doi.org/10.3389/fpsyg.2015.01746>
- Schumacher, P. B., Dangl, M., & Uzun, E.** (2016). Thematic role as prominence cue during pronoun resolution in German. In *Empirical Perspectives on Anaphora Resolution* (pp. 213–240). De Gruyter.
- Schumacher, P. B., Roberts, L., & Järvikivi, J.** (2017). Agentivity drives real-time pronoun resolution: Evidence from German er and der. *Lingua*, *185*, 25–41. <http://dx.doi.org/10.1016/j.lingua.2016.07.004>
- Schütze, C. T., & Sprouse, J.** (2014). Judgment data. In R. J. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 27–50). Cambridge University Press. <https://doi.org/10.1017/CBO9781139013734.004>
- Stevenson, R. J., Crawley, R. A., & Kleinman, D.** (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*(4), 519–548. <https://doi.org/10.1080/01690969408402130>

- Turan, Ü. D.** (1995). *Null vs. Overt subjects in Turkish discourse: A centering analysis* [PhD Thesis]. University of Pennsylvania.
- Uzun, E.** (2015a). *Ambiguous pronoun resolution in ditransitive contexts* [Forschungsvortiefung/Hausarbeit].
- Uzun, E.** (2015b). *Eine explorative EEG-Studie zur Verarbeitung von ambigen Personal- und Demonstrativpronomen in ditransitiven Kontexten* [Master thesis]. Johannes Gutenberg-Universität Mainz.
- Vogel, R., & Steinbach, M.** (1998). The Dative—An Oblique Case. *Linguistische Berichte*, **173**, 65–90.
- von Heusinger, K.** (2006). Salience and anaphoric definite noun phrases. *Acta Linguistica Hafniensia*, **38**, 33–53.
- von Heusinger, K., & Schumacher, P. B.** (2019). Discourse prominence: Definition and application. *Journal of Pragmatics*, **154**, 117–127. <https://doi.org/10.1016/j.pragma.2019.07.025>
- Weinrich, H.** (1993). *Textgrammatik der deutschen Sprache*. Dudenverlag.
- Wilson, F.** (2009). *Processing at the syntax–discourse interface in second language acquisition* [PhD Thesis]. University of Edinburgh.
- Zifonun, G., Hoffmann, L., Strecker, B., & Ballweg, J.** (1997). *Grammatik der Deutschen Sprache* (Vol. 1). De Gruyter.

Cite this article: Patterson, C. and Schumacher, P.B. (2021). Interpretation preferences in contexts with three antecedents: examining the role of prominence in German pronouns. *Applied Psycholinguistics* **42**, 1427–1461. <https://doi.org/10.1017/S0142716421000291>