



Hinterwimmer, Stefan & Patil, Umesh. 2022. The interpretative options of anaphoric complex demonstratives. *Glossa: a journal of general linguistics* 7(1), pp. 1–30. DOI: <https://doi.org/10.16995/glossa.5700>



Open Library of Humanities

The interpretative options of anaphoric complex demonstratives

Stefan Hinterwimmer, German studies, University of Wuppertal, DE, hinterwimmer@uni-wuppertal.de

Umesh Patil, IdSL I, University of Cologne, DE, umesh.patil@gmail.com

In this paper, we present experimental evidence from a ‘yes’/‘no’ judgement task and two acceptability rating studies (Experiments 1a-c) for the claim made in Hinterwimmer (2019) that sentences with two anaphorically interpreted complex demonstratives are less acceptable than sentences with two anaphorically interpreted definite descriptions and sentences where one of the two previously introduced referents is picked up by a complex demonstrative, while the other one is picked up by a definite description. The results of Experiment 1a and 1b are in principle compatible with the account argued for in Hinterwimmer (2019), according to which the (potentially abstract) demonstrations presupposed by demonstratives may not have overlapping trajectories. However, sentences with two anaphorically interpreted complex demonstratives are not judged as unacceptable as would be expected if they involved a presupposition violation. Therefore, we propose an alternative, economy-based pragmatic account that builds on Ahn (2019) and Nowak (2019). The question of whether the observed pattern is more compatible with the account proposed by Hinterwimmer (2019) or the alternative pragmatic account is directly addressed in a further acceptability rating study (Experiment 1c). The design of that study is similar to that of Experiment 1b, but it includes as fillers both sentences clearly violating a presupposition and sentences violating a pragmatic constraint. Since the ratings for sentences containing two anaphorically interpreted complex demonstratives are closer to the ratings for sentences violating a pragmatic constraint than for sentences violating a presupposition, we conclude that the alternative pragmatic account is preferable to the account by Hinterwimmer (2019).

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by the Open Library of Humanities. © 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

OPEN ACCESS



1 Introduction

In this paper, we present experimental evidence from a ‘yes’/‘no’ judgement task (Experiment 1a) and two acceptability rating studies (Experiment 1b-c) for the following claim made in Hinterwimmer (2019) on the basis of purely introspective judgements: Sentences with two anaphorically interpreted complex demonstratives are less acceptable than sentences with two anaphorically interpreted definite descriptions and sentences where one of two referents introduced by indefinites in the previous sentence is picked up by a complex demonstrative, while the other one is picked up by a definite description. In Hinterwimmer (2019), this contrast is accounted for by combining Roberts’ (2002) analysis according to which complex demonstratives not only on their deictic, but also on their anaphoric and bound uses presuppose (possibly abstract) demonstrations with the following assumption: The trajectories of these demonstrations may not overlap. In a configuration where two demonstratives presuppose abstract demonstrations linking back to antecedents contained in a previous sentence, this presupposition is necessarily violated. While the results of Experiments 1a and 1b are in principle compatible with the account of Hinterwimmer (2019), sentences with two anaphorically interpreted complex demonstratives are not judged as unacceptable as would be expected if they involved a presupposition violation. Therefore, we propose an alternative, purely pragmatic account for why sentences with co-occurring anaphorically interpreted complex demonstratives are perceived as deviant (building on King 2001; Elbourne 2008; Nowak 2014; 2019; Ahn 2019). Since on this alternative account, sentences exemplifying the pattern under consideration do not violate a presupposition, but only an economy-based pragmatic principle, the prediction is that they should only be rated as slightly deviant. On the account proposed by Hinterwimmer (2019), in contrast, they are expected to be rated as strongly deviant. The competing predictions of the two accounts are directly tested in a further acceptability rating study (Experiment 1c). The design of that study is similar to that of Experiment 1b, but it includes as fillers both sentences clearly violating a presupposition and sentences violating a pragmatic constraint.¹ Since the ratings for sentences containing two anaphorically interpreted complex demonstratives are closer to the ratings for sentences violating a pragmatic constraint than for sentences violating a presupposition, we conclude that the alternative pragmatic account is preferable to the account of Hinterwimmer (2019).

The paper is structured as follows. In Section 2.1 the analysis of complex demonstratives proposed by Hinterwimmer (2019) is summarized, and in Section 2.2 three alternative pragmatic analyses are introduced that all account for the deviance of sentences with two anaphoric complex demonstratives. Two of them make different predictions than those reported in Hinterwimmer (2019) regarding sentences where an anaphoric complex demonstrative is combined with an

¹ We are grateful to the editors for suggesting the basic design of Experiment 1c to us.

anaphoric definite description. The other one makes essentially the same predictions as those reported in Hinterwimmer (2019), but predicts the effect size to be smaller than the analysis argued for in Hinterwimmer (2019). In Section 3, the three experimental studies testing the predictions of the various accounts are discussed. Section 4 is the conclusion.

2 A presupposition-based account of the interpretative options of complex demonstratives

2.1 The analysis proposed in Hinterwimmer (2019)

In Hinterwimmer (2019) an unexpected contrast between anaphoric uses of complex demonstratives and definite descriptions is observed: Sentences in which two discourse referents that were introduced by indefinites in the previous sentence are both picked up by complex demonstratives (as in (1b)) are quite deviant. Sentences in which both are picked up by definite descriptions (as in (1a)) or one is picked up by a definite description and the other one by a complex demonstrative (as in (1c) and (1d)), in contrast, are perfectly acceptable.

- (1)
- a. Last night, a dog chased a cat in front of my house. Fortunately, [the cat] was pretty fast, while [the dog] was rather slow.
 - b. Last night, a dog chased a cat in front of my house. ^{??}Fortunately, [that cat] was pretty fast, while [that dog] was rather slow.
 - c. Last night, a dog chased a cat in front of my house. Fortunately, [that cat] was pretty fast, while [the dog] was rather slow.
 - d. Last night, a dog chased a cat in front of my house. Fortunately, [the cat] was pretty fast, while [that dog] was rather slow.

There is obviously neither a general constraint that prohibits the co-occurrence of two complex demonstratives in adjacent clauses nor a constraint against the co-occurrence of two complex demonstratives with contrasting nominal predicates in two adjacent sentences: The sentences in (2a-b), for example, are perfectly fine when they are uttered by a speaker who points into the direction of a dog to the right of her while uttering the first complex demonstrative, and into the direction of a dog (in the case of 2a) or a cat (in the case of 2b) to the left of her.

- (2)
- a. [That dog] is smart, but [that dog] is rather stupid.
 - b. [That dog] is smart, but [that cat] is rather stupid.

The infelicity of (1b) as opposed to (1a) and (1c-d) therefore seems to be related to the anaphoricity of the complex demonstratives it contains. This is not readily explained on the accounts of King (2001), Elbourne (2008), Nowak (2014; 2019); but see footnote 2 for a qualification concerning Elbourne (2008) and Section 3.1 for a pragmatic analysis based on the underlying assumption shared by these analyses). Therefore, Hinterwimmer (2019) argues for an analysis of complex

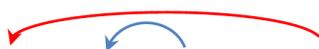
demonstratives that builds on Roberts (2002).² Roberts (2002) assumes that (simplex as well as) complex demonstratives always presuppose concrete or abstract demonstrations. On deictic uses, those demonstrations are usually concrete pointing gestures in physical space. On anaphoric or bound uses, in contrast, what is demonstrated are not individuals or objects that are present in the utterance situation, but rather previously uttered DPs. Hinterwimmer (2019) adds to these assumptions a constraint that prevents demonstrations from having overlapping trajectories.

The basic assumptions can be summarized as follows: First, whenever a sentence is uttered, not only the proposition it denotes becomes part of the Context Set, i.e. the set of worlds compatible with the interlocutor's mutually shared background assumptions (Stalnaker 1978). Rather, discourse referents for the sentences themselves and their constituents are introduced as well. Second, discourse referents for sentences contain information concerning the linear sequence of their constituents. Third, demonstrations in discourse occur in the 'discourse space' set up by those linear sequences. Crucially, the trajectories of demonstrations in discourse connect the position of a demonstrative in a currently uttered sentence with the position of its antecedent in the previous sentence. From these assumptions it follows that the trajectories of demonstrations in discourse necessarily overlap whenever there is more than one demonstrative anaphorically related to a DP in a previous sentence (see the schematic graphical representations in (3b) and (4b)).

To see this, consider the text segments in (1b), repeated here as (3a) and (4a): In each case, the demonstrations in discourse presupposed by the complex demonstratives in the second sentence necessarily have overlapping trajectories, irrespective of the order in which the antecedents introduced in the previous sentence are picked up by the demonstratives.

- (3) a. Last night, a dog chased a cat in front of my house. ??Fortunately, [that cat] was pretty fast, while [that dog] was rather slow.

b. ... a dog ... a cat ... that cat ... that dog ...



² As has been pointed out to us by an anonymous reviewer, the analysis argued for by Elbourne (2008) could in principle be developed along similar lines as the one of Roberts (2002) in order to account for the data discussed in this section. In order to account for cases of deferred ostension and descriptive indexicals (Nunberg 1993), Elbourne (2008) assumes the individuals denoted by complex demonstratives to stand in some relation (to be determined by the context) to an index, where that index is proximal to the speaker in the case of *this*, and distal in the case of *that*. Crucially, for donkey anaphoric complex demonstratives like those in (1b-d), Elbourne assumes that the index is the occurrence of the indefinite antecedent. Since this idea is rather similar to the idea underlying Roberts' (2002) account, the crucial aspects of the analysis argued for in Hinterwimmer (2019) could also have been implemented in Elbourne's (2008) account. Since the resulting analyses would be very similar, however, and would therefore face the same problems (to be discussed below), we will not discuss this option further.

- (4) a. Last night, a dog chased a cat in front of my house. ^{??}Fortunately, [that dog] was pretty slow, while [that cat] was pretty fast.

b. ... a dog ... a cat ... that dog ... that cat ...

The denotation assumed in Hinterwimmer (2019) for *that* is given in (6).³ Informally, it can be paraphrased as in (5) (Note that *C* is the context where the respective sentence is uttered):

- (5) *that* takes a predicate *P* of type $\langle e, \langle s, t \rangle \rangle$ and a situation *s* as arguments and maps them to the unique *x* such that
- x* satisfies *P* in *s*,
 - there is a (concrete or abstract) demonstration δ_1 in the domain of *C* whose agent is the speaker and whose demonstratum is a *y* on which *x* is interpretatively dependent,
 - there is no (concrete or abstract) demonstration δ_2 in the domain of *C* whose agent is the speaker and whose demonstratum is a *z* distinct from *y* such that the trajectories of δ_1 and δ_2 overlap.

- (6) $[[\text{that}]]^{C,w,g} = \lambda P_{\langle e, \langle s, t \rangle \rangle} . \lambda s . \iota \{x: P(x)(s) \wedge \exists y \in \text{DOM}_C \exists \delta_1 \in \text{DOM}_C \neg [\exists \delta_2 \in \text{DOM}_C \exists z \in \text{DOM}_C [\text{Agent}(\delta_1)(\text{Author}(C)) \wedge \text{Demonstratum}(\delta_1)(y) \wedge \text{interpretatively_dependent_on}(x)(y) \wedge \text{Agent}(\delta_2)(\text{Author}(C)) \wedge \text{Demonstratum}(\delta_2)(z) \wedge z \neq y \wedge \text{overlap}(\text{trajectory}(\delta_1))(\text{trajectory}(\delta_2))]]\}$.

A1. If *y* is a concrete individual in physical space, *x* is interpretatively dependent on *y* iff *x* is identical to *y*.

A2. If *y* is a DP that is contained in a previously uttered sequence, *x* is interpretatively dependent on *y* iff *x* is either identical to the discourse referent introduced by *y* or to (the value of) a variable bound by *y*.

B1. If the demonstratum of a demonstration δ is a concrete individual *y* in physical space and the speaker is the agent of δ , the trajectory of δ connects the location of the speaker in the utterance situation with the location of *y* in the utterance situation.

B2. If the demonstratum of a demonstration δ is a DP that is contained in a previously uttered sequence and the speaker is the agent of δ , the trajectory of δ connects the location of the demonstrative determiner presupposing in the sequence of uttered words with the location of *y* in the previously uttered sequence.

With these assumptions in place, let us now have a detailed look at what goes wrong in cases like (3a) and (4a). Concerning the first complex demonstrative in the second sentence of (3a)/(4a), there are no problems. In both cases it can be interpreted as denoting the unique individual *x* that

³ Since the difference between *this* and *that* plays no role for the data discussed in this paper, we will set *this* aside.

- (a) has the property of being a cat/a dog in s , where s is bound by a covert existential quantifier over situations, and
- (b) for which there is a DP y and a demonstration d_1 such that the agent of d_1 is the speaker and the demonstratum of d_1 is y and x is identical to the discourse referent introduced by y , where y is the indefinite *a cat/a dog* contained in the respective opening sentence, and
- (c) for which there is no DP z and demonstration d_2 such that the agent of d_2 is the speaker and the demonstratum of d_2 is z and x is identical to the discourse referent introduced by z , and the trajectories of d_1 and d_2 overlap.

Concerning the second complex demonstrative, in contrast, a presupposition failure is unavoidable since in both cases there is a demonstration, the trajectory of which overlaps with the trajectory of the demonstration required for that demonstrative – namely the demonstration that was required for the interpretation of the first complex demonstrative. No such problem occurs in cases like (1c) and (1d), where only one of two discourse referents introduced in the preceding sentence is picked up by a demonstrative. Consequently, the trajectory of the demonstration presupposed by that demonstrative does not overlap with the trajectory of another demonstration by the speaker.

Let us finally return to the sentences in (2a-b), repeated here as (7a-b). Let us assume that both sentences are uttered in a situation where two dogs (in the case of (7a)) or a dog and a cat are present and where the use of each complex demonstrative is accompanied by a pointing gesture into the direction of one of the two dogs (in the case of (7a)) or the dog and the cat, respectively. In neither case is there a presupposition failure, since the trajectories of the demonstrations accompanying the utterances of the respective demonstratives never have to be construed as overlapping, due to their occurrence in three-dimensional space: Even in a case where the speaker is standing in the middle of a room, pointing at the dog that is sitting in the right corner in front of her with her left hand and at the dog or cat that is sitting in the left corner in front of her with her right hand, the trajectory of the first demonstration can be construed as slightly below or above the trajectory of the other one, thus avoiding overlap.

- (7) a. [That dog] is smart, but [that dog] is rather stupid.
- b. [That dog] is smart, but [that cat] is rather stupid.

While the analysis discussed in this section captures the observed pattern, we would like to point out some problems it faces, which, together with the results of Experiments 1a and 1b to be reported in Section 2.2, lead us to propose an alternative pragmatic account to be discussed in Section 2.2.⁴ First, it is rather stipulative insofar as the presupposition for the demonstrations associated with

⁴ The pragmatic analyses discussed in Section 2.2 were only developed after we obtained the results of Experiments 1a and 1b, which were designed to test the predictions of Hinterwimmer (2019) exclusively. Experiment 1c was specifically designed to test the competing predictions of Hinterwimmer (2019) and the alternative pragmatic accounts, following a suggestion of the editors. To increase the readability of the paper, however, we present the competing accounts first before discussing all three experimental studies together.

complex demonstratives to have non-overlapping trajectories is not derived from any more general principles. Second, it builds on a number of non-standard concepts such as abstract demonstrations in discourse space and discourse referents for sentences containing information about the linear sequence of their constituents whose precise formal status is unclear (but see Hinterwimmer 2019 for discussion). Third, as has been pointed out to us by an anonymous reviewer, it runs into the following empirical problem: Suppose the following sentence is uttered by a speaker standing in front of a conveyor belt which transports dogs, each dog stopping for a moment when it gets to the speaker, while pointing to two different dogs: ‘That dog... [PAUSE] is faster than THAT dog’. Since the dogs appear in the same spot in front of the speaker, the gestures would have overlapping, and indeed identical, trajectories. Nevertheless, the utterance is absolutely fine.

2.2 An alternative pragmatic account

The pragmatic account discussed in this section builds on the idea that demonstrative determiners are similar to the definite determiner insofar as they presuppose uniqueness (King 2001; Wolter 2006; Elbourne 2008; Nowak 2014; 2019; Ahn 2019). In contrast to the definite determiner, however, there need not be a unique individual satisfying the predicate denoted by its overt NP-complement, but rather a unique individual satisfying both that predicate and an additional predicate (King 2001; Elbourne 2008; Nowak 2014; 2019; Ahn 2019). The latter predicate is either provided by a relative clause (Nowak 2019) – assuming with Ross (1967), Stockwell et al. (1973), Bach & Cooper (1978), Larson (1982) and del Gobbo (2003) that even restrictive relative clauses do not have to form a constituent with the noun they modify, but can be adjoined to the respective DP (see footnote 5 for further discussion) – or it initially corresponds to a free variable over predicates denoted by a covert NP. That predicate variable can now be resolved to the property of being identical to the individual the speaker is pointing at in the utterance situation, giving rise to a deictic, directly referential interpretation (Kaplan 1989), or to some contextually salient property.

Let us now follow Nowak (2019), who goes one step further by imposing a particular restriction on the value R of the free predicate value: It has to restrict the predicate P denoted by the overt NP, i.e. the intersection of P and R has to be a subset of P . This restriction is motivated by the observation that complex demonstratives such as *that author of Waverley* are extremely awkward, in contrast to definite descriptions such as *the author of Waverley*. The proposed constraint, which is technically implemented as a presupposition, readily accounts for this observation: Since the predicate *author of Waverley* is true of exactly one individual, there can be no value for the free predicate value that restricts that predicate.⁵ As pointed out by

⁵ Nowak (2019) observes that the semantically equivalent complex demonstrative *that guy who wrote Waverley* is fine. He accounts for the contrast between *that author of Waverley* and *that guy who wrote Waverley* by assuming as follows: In the latter case, the demonstrative determiner takes the relative clause *who wrote Waverley* as its second argument. Since the (singleton) set of individuals who wrote Waverley is a subset of the set denoted by the noun *guy*, the presupposition is fulfilled that the intersection of the two predicates is a subset of the set corresponding to the first predicate.

Nowak (2019), the proposed analysis is in line with an intuition shared by many linguists and philosophers working on demonstratives: When using a complex demonstrative *that F* instead of a definite description *the F* the speaker at least implicitly contrasts the individual or object x she is referring to with other individuals or objects in F by invoking an additional property setting x apart from all the other elements of F (see, e.g. Bisle-Müller 1991 and Reimer 1991 for relevant discussion). This, of course, only works if F is not a singleton set to begin with.

A slightly modified version of the denotation proposed by Nowak (2019) for *that* is given in (8a), with the denotation of the definite determiner given for comparison in (8b).

- (8) a. $[[\text{that}]] = \lambda P_{\langle e, \langle s, t \rangle \rangle} \lambda Q_{\langle e, \langle s, t \rangle \rangle} . \lambda s: (\{x : P(x)(s)\} \cap \{x : Q(x)(s)\} \subset \{x : P(x)(s)\} .$
 $\iota\{x: P(x)(s) \wedge Q(x)(s)\}$
- b. $[[\text{the}]] = \lambda P_{\langle e, \langle s, t \rangle \rangle} . \lambda s . \iota\{x: P(x)(s)\}$

Let us now return to the pattern discussed in Section 2.1. Following Elbourne (2001; 2005; 2013), we assume that anaphoric readings of definite descriptions come about via relativizing the uniqueness presupposition of the definite determiner to minimal extensions of the minimal situations in which their antecedents were introduced. Minimal situations are situations that contain only those individuals, properties and relations that are required to make some proposition true (Kratzer 1989). A minimal situation s of a dog chasing a cat is thus a situation that contains nothing beyond an individual that has the property of being a dog, an individual that has the property of being a cat and the two individuals standing in the relation of the first chasing the second. A minimal extension $s1$ of s such that the unique cat in $s1$ is faster than the unique dog in $s1$ is therefore guaranteed to contain the cat and the dog that were introduced in $s1$ (since otherwise they would not be minimal extensions of s). Simplifying strongly and concentrating only on what is important for our purposes, the two sentences in (1a), repeated here as (9a), therefore receive a reading that can be paraphrased as in (9b), assuming that the situation arguments of the two definite descriptions are bound by the covert existential quantifiers that also bind the situation arguments of the respective verbs.

- (9) a. Last night, a dog chased a cat in front of my house. Fortunately, [the cat] was pretty fast, while [the dog] was rather slow.
- b. There is a past situation s that is a minimal situation of a dog chasing a cat in front of the speaker's house and there is a past situation $s1$ that is a minimal extension of s such that the unique cat in $s1$ is faster than the unique dog in $s1$.

Note that in virtue of the two individuals in the second sentence being referred to by DPs that only differ with respect to the nominal predicates, they are automatically implicitly contrasted with each other, i.e. the unique individual that has the property of being a cat in $s1$ is contrasted with the unique individual that has the property of being a dog in $s1$.

Let us turn to (1b) next, repeated here as (10a), where the second sentence contains two anaphorically interpreted complex demonstratives. Now, as soon as the situation argument of the respective demonstrative determiner is bound by the covert existential operator quantifying over minimal extensions of the situation introduced by the first sentence, the restriction presupposition is automatically violated, no matter which value is chosen for the second predicate: Both the set of dogs and the set of cats in $s1$ are singletons, so they cannot be further restricted. Consequently, another value has to be chosen for the respective situation argument. Let us assume that the world of evaluation is chosen by default (following Kratzer 1989, we assume that worlds are maximal situations). Technically, this can be implemented by assuming the presence of a covert pronoun introducing a free variable over situations that is then resolved to the world of evaluation. The next step consists in finding a value for the second predicate argument of the respective demonstrative determiner. Since neither of the two complex demonstratives contains a relative clause, the respective demonstrative determiner is combined with a covert NP denoting a free variable over predicates first. In the absence of pointing gestures, both variables have to be resolved to some predicate provided by the linguistic context. For the first complex demonstrative, a natural choice is the property of having been chased by a dog in front of the speaker's house on the previous night, while for the second one it is the property of having chased a cat in front of the speaker's house on the previous night. The two sentences are thus interpreted as paraphrased as in (10b).

- (10) a. Last night, a dog chased a cat in front of my house. ^{??}Fortunately, [that cat] was pretty fast, while [that dog] was rather slow.
- b. There is a past situation s that is a minimal situation of a dog chasing a cat in front of the speaker's house and there is a past situation $s1$ that is a minimal extension of s such that the unique cat in $w0$ that was chased by a dog in front of the speaker's house on the previous night is faster in $s1$ than the unique dog in $w0$ that has chased a cat in front of the speaker's house on the previous night.

The interpretation paraphrased in (10b) is perfectly coherent. So, what is the problem with the second sentence in (10a)? In principle, choosing the demonstrative determiner instead of the definite determiner should be the dispreferred option whenever the intended interpretation could just as easily have been arrived at by combining the definite determiner with the respective NP-complement, i.e. whenever the linguistic context guarantees the existence of exactly one individual satisfying the predicate denoted by that NP (see Ahn 2019 for related discussion). The reason for this is that using a demonstrative instead of a definite determiner in such cases adds an unnecessary complication since a value for the free predicate variable has to be determined by the addressee(s). Now, the interpretation of (9a) paraphrased in (9b) is clearly equivalent to the interpretation of (10a) paraphrased in (10b). The only difference between the two is that

(10b) has been arrived at in a more complicated way that involves determining the values of two predicate variables. Seen this way, it is not surprising that (10a) is awkward.

This cannot be the end of the story, however, since the same logic can be applied to (1c) and (1d) – repeated here as (11a) and (11b), respectively –, which both combine a complex demonstrative with a definite description in the second sentence.

- (11) a. Last night, a dog chased a cat in front of my house. Fortunately, [that cat] was pretty fast, while [the dog] was rather slow.
 b. Last night, a dog chased a cat in front of my house. Fortunately, [the cat] was pretty fast, while [that dog] was rather slow.

In each of the two cases, choosing the demonstrative instead of the definite determiner once should make the respective sentence at least mildly deviant, with the awkwardness of (1a)/(10a) resulting from the combined effect of two violations of the economy principle that is violated only once in both (1c)/(11a) and (1d)/(11b). An analysis along these lines (which we will henceforth dub **PA 1**) is an attractive, rather straightforward option. It is not in line with the intuitions reported in Hinterwimmer (2019), however, according to which (1c)/(11a) and (1d)/(11b) are not deviant at all. In order to be consistent with this intuition (and with the general observation that complex demonstratives can be used anaphorically), we need to modify **PA1**.

Recall from above that there is a widespread intuition that when uttering a complex demonstrative, the speaker implicitly compares the referent of that demonstrative with all other entities that satisfy the predicate denoted by the NP-complement of the demonstrative determiner. In the case of the first complex demonstrative in (1b)/(10a) the particular cat denoted by that demonstrative is implicitly contrasted with all other cats in the world of evaluation, while in the case of the second demonstrative the particular dog denoted by that demonstrative is implicitly contrasted with all other dogs in the world of evaluation. At the same time, since the two individuals are referred to by DPs that only differ with respect to the nominal predicates, they are automatically implicitly contrasted with each other as well, just as in the case of the two definite descriptions in (1a)/(9a). We assume that the co-presence of two implicit contrasts makes the second sentence in (1b) unnecessarily complex and is therefore responsible for its being dispreferred to the other options exemplified by (1a)/(9a), (1c)/(11a) and (1d)/(11b), respectively: In the case of (1a)/(9a), only one implicit contrast is invoked – namely the contrast between the unique cat in *s1* and the unique dog in *s1*. In the case of both (1c)/(11a) and (1d)/(11b), on the other hand, the only implicit contrast that is invoked is the contrast between the individual denoted by the respective complex demonstrative and the individuals in the set denoted by the respective nominal predicate. Crucially, no implicit contrast between the two

individuals denoted by the complex demonstratives is invoked, since the respective DPs do not only differ with respect to the nominal predicates, but also with respect to the determiners.

Note that both the modified economy-based pragmatic account under discussion (which we will henceforth dub **PA2**) and **PA1** are compatible with the observation that sentences such as (2b) (repeated here as (12)), which contain two complex demonstratives with distinct NPs and therefore implicitly invoke two contrasts at the same time, are perfectly acceptable: In such cases, there simply is no semantically equivalent alternative available since neither of the two demonstrative determiners could have been replaced by a definite determiner without a change in interpretation. After all, definite descriptions simply do not allow for deictic interpretations where reference is (at least partially) determined by an accompanying pointing gesture (or some equivalent way of indicating the speaker's referential intention).

(12) [That dog] is smart, but [that cat] is rather stupid.

PA 2 essentially makes the same predictions as the analysis of Hinterwimmer (2019) discussed in Section 2.1: Sentences with two anaphorically interpreted complex demonstratives are deviant. Sentences with two anaphorically interpreted definite descriptions and sentences where an anaphoric definite description is combined with an anaphoric complex demonstrative, in contrast, are perfectly acceptable. Still, the empirical predictions of the two accounts are not completely equivalent: According to Hinterwimmer (2019), the presence of two anaphorically interpreted complex demonstratives necessarily triggers a presupposition violation. This should have a rather strong effect on acceptability. According to **PA 2**, in contrast, sentences with two complex demonstratives do not violate any lexically encoded constraints. Rather, they should simply be dispreferred as compared to less complex, but semantically equivalent alternatives. Consequently, they are expected to be perceived as mildly deviant, i.e. while there should be a clear contrast between sentences with two anaphorically interpreted complex demonstratives, on the one hand, and the three other types of sentences under discussion, on the other, the contrast should still be relatively small.

Finally, we would like to briefly mention a third conceivable pragmatic account (henceforth: **PA 3**) of the deviance of sentences with two anaphoric complex demonstratives.⁶ **PA 3** builds on the observation that demonstrative pronouns of the *dieser/diese/dieses* paradigm (DemPros), which are homophonous to the demonstrative determiner in German, have a strong preference for picking up the last-mentioned referent (see Patil, Bosch & Hinterwimmer 2020 and the references cited therein for recent discussion). This tendency could be assumed to apply to complex demonstratives as well, even in English, where DemPros can only refer to inanimate

⁶ An account along these lines has been suggested to us by an anonymous reviewer.

objects. The deviance of sentences with two anaphoric complex demonstratives would then result from both complex demonstratives ‘trying’ to pick up the last-mentioned referent of the previous sentence, which is of course impossible. **PA 3** makes a further prediction, however, that is different from the predictions made by Hinterwimmer (2019)/**PA 2** (and **PA 1** as well): Cases like (11b), where the complex demonstrative cannot be interpreted as picking up the last-mentioned referent of the preceding sentence, but only as picking up the first-mentioned one because of the respective nominal predicates, should be just as deviant as sentences with two complex demonstratives. After all, the tendency for demonstratives to pick up the last-mentioned referent is violated in both cases.

Let us briefly summarize the predictions of the analyses discussed in this section and in Section 2.1. Both Hinterwimmer (2019) and **PA 2** predict sentences with two anaphoric complex demonstratives (such as (1b), repeated here as (13b)) to be deviant and sentences where an anaphoric complex demonstrative is combined with an anaphoric definite description (such as (1c) and (1d), repeated here as (13c) and (13d), respectively) to be fully acceptable, irrespective of whether the anaphoric complex demonstrative picks up the first or the second of the two discourse referents introduced in the previous sentence. They make different predictions regarding the effect size, however: Hinterwimmer (2019) predicts sentences with two complex demonstratives to be strongly deviant, since they violate a presupposition. **PA 2**, in contrast, predicts them to be only mildly deviant, since they only violate an economy-based pragmatic constraint.

- (13) a. Last night, a dog chased a cat in front of my house. Fortunately, [the cat] was pretty fast, while [the dog] was rather slow.
 b. Last night, a dog chased a cat in front of my house. ^{??}Fortunately, [that cat] was pretty fast, while [that dog] was rather slow.
 c. Last night, a dog chased a cat in front of my house. Fortunately, [that cat] was pretty fast, while [the dog] was rather slow.
 d. Last night, a dog chased a cat in front of my house. Fortunately, [the cat] was pretty fast, while [that dog] was rather slow.

PA1 predicts not only sentences with two anaphoric complex demonstratives to be (mildly) deviant (such as (1b)/(13b)), but also sentences where an anaphoric complex demonstrative is combined with an anaphoric definite description (such as (1c)/(13c) and (1d)/(13d)), irrespective of which discourse referent is picked up by the complex demonstrative. Additionally, the deviance of sentences with two anaphoric complex demonstratives is predicted to be the combined deviance of the sentences where an anaphoric complex demonstrative is combined with an anaphoric definite description.

Finally, **PA3** predicts the two following types of sentences to be deviant: Sentences with two anaphoric complex demonstratives (such as (1b)/(13b)) and sentences where the last-mentioned

referent of the previous sentences is picked up by a definite description, while the first-mentioned one is picked up by a complex demonstrative (such as (1d)/(13d)).

3. The experimental studies

3.1 Overview

In this section, we will discuss three experimental studies testing the predictions of the accounts that we presented in Section 2: a ‘yes’/‘no’ judgement task (Experiment 1a) and two acceptability rating studies (Experiments 1b and 1c). Experiments 1a and 1b were initially conducted to test the predictions of Hinterwimmer (2019) exclusively,⁷ but, as we will see, their results speak to the predictions of **PA1** and **PA3** as well, insofar as they are more compatible with the analysis of Hinterwimmer (2019) than with either **PA1** or **PA3**. Concerning **PA2**, recall that its predictions only differ from those of Hinterwimmer (2019) with respect to the effect size. Consequently, the results of Experiment 1a and 1b do not allow us to distinguish between Hinterwimmer (2019) and **PA2**. Experiment 1c was therefore specifically designed to test the competing predictions of **PA2** and Hinterwimmer (2019).

3.2. Experiments 1a and 1b

Experiment 1a is a ‘yes’/‘no’ acceptability judgement task and Experiment 1b a 1–7 Likert scale acceptability judgement task. Participants had to judge text segments (see (14) below) in which two referents introduced in the first sentence by indefinites were either picked up by two complex demonstratives (Dem-Dem), a complex demonstrative followed by a definite description (Dem-Def), a definite description followed by a complex demonstrative (Def-Dem), or two definite descriptions (Def-Def). We first describe the two experiments and then discuss the results together.

3.2.1 Experiment 1a

Method

Participants

Thirty-nine self-reported native speakers of German (27 females, mean age = 23 years, age range = 20–35) were recruited from the University of Cologne for course credits.

Design and Materials

We used 28 experimental items in 4 conditions as in (14). These were interspersed with 48 filler items of comparable length and complexity. There were 5 types of fillers that covered a broad range of acceptability – from completely acceptable to very low in acceptability.

⁷ As already said in Footnote 4 above, **PA1-3** were only developed after we obtained the results of Experiments 1a and 1b, which were initially designed to test the predictions of Hinterwimmer (2019) exclusively.

- (14) Im Konzert gestern Abend hat sich eine Musikerin mit einem Stimmnachbarn gestritten.
In the concert last night, a musician (female) argued with a colleague (male).
 [Dem-Dem] Nach dem Konzert hat **diese Musikerin** mit **diesem Stimmnachbarn** kein Wort mehr gesprochen.
After the concert, this musician did not speak a word to this colleague
 [Dem-Def] Nach dem Konzert hat **diese Musikerin** mit **dem Stimmnachbarn** kein Wort mehr gesprochen.
After the concert, this musician did not speak a word to the colleague
 [Def-Dem] Nach dem Konzert hat **die Musikerin** mit **diesem Stimmnachbarn** kein Wort mehr gesprochen.
After the concert, the musician did not speak a word to this colleague
 [Def-Def] Nach dem Konzert hat **die Musikerin** mit **dem Stimmnachbarn** kein Wort mehr gesprochen.
After the concert, the musician did not speak a word to the neighbor.

Procedure

The experiment involved a ‘yes’/‘no’ judgement task. Participants were instructed that the texts were beginnings of short stories produced by advanced German learners, where each example text was produced by a different learner. The participants’ task was to judge whether the student had reached native-like proficiency in German (by responding “yes, they have” or “no, they have not”). The reason why we asked participants to judge language proficiency instead of acceptability was that judging nativity is a more well-defined or natural task than judging acceptability, which could be influenced by factors such as prescriptive knowledge of the grammar and meta-level reasoning such as the plausibility of the text.

Data analysis

All data processing and analyses reported in the paper were carried out in R (R Core Team 2020). We used the Bayesian framework for data analysis. Carrying out data analysis in the Bayesian framework has many advantages over the frequentist framework. For us the two main advantages were quantifying the uncertainty about the effects through 95% credible intervals (95% CrI) around the estimates and the ease of fitting complex models. The CrI specifies the range of plausible parameter values given the data. These values are considered “credible” enough for making theoretical inferences about the presence of an effect in light of the data. Moreover, if we want to fit a complex model, for example a model with maximal random effects structure (Barr et al. 2013), it almost always converges in the Bayesian framework which is not the case with frequentist models that are fit using tools such as *lme4* (see Nicenboim & Vasishth 2016 for a review of Bayesian data analytical methods and an example analysis of

a typical psycholinguistic experiment). Most of the models reported in the paper (5 out of 7) involved making an inference based on the interaction of at least two factors and those models did not converge when we used frequentist tools. All models reported in the paper were fit with a maximal random effects structure and they did converge, albeit with more time and computational power.

Recall that Hinterwimmer (2019) (and **PA 2**) predicts that *Dem-Dem* should be worse than the other three conditions, while there should be no contrast between the other three conditions. **PA 1**, in contrast, predicts that, first, both *Dem-Def* and *Def-Dem* should be worse than *Def-Def* and, second, that the deviance of *Dem-Dem* should equal the combined effects of *Dem-Def* and *Def-Dem*. Finally, **PA 3** predicts that *Dem-Def*, where the complex demonstrative picks up the first-mentioned referent instead of the last-mentioned one, should pattern with *Dem-Dem*. In Experiment 1 we originally intended to compare the effect of the *Dem-Dem* condition with the other three conditions to test the prediction of Hinterwimmer (2019). In light of the differing predictions of **PA1** and **PA3**, however we added a comparison of the effect of having a complex demonstrative in positions DP1 and DP2. Since the two comparisons cannot technically be carried out in a single statistical model, we tested one model for each comparison. The model involving the first comparison (*comparison-1*) was fitted with the “treatment contrast” and *Dem-Dem* as the reference level against which the rating of the other three conditions was compared. The model involving the second comparison (*comparison-2*) was fitted with two treatment contrasts. The first contrast, *DP1-complex*, had *Def-Def* and *Def-Dem* as the reference level, and *Dem-Def* and *Dem-Dem* were the treatment level. The second contrast, *DP2-complex*, had *Def-Def* and *Dem-Def* as the reference level and the other two as the treatment level. *Comparison-2* also included interaction of these two factors.

In Experiment 1a, since the responses were binomial (*yes* or *no*), we analysed them using mixed-effects logistic regression through the R package *brms* (Bürkner 2017) which provides an interface to Stan, a probabilistic programming language for specifying Bayesian statistical models (Carpenter et al. 2017). To avoid the extreme probability values (0 and 1), we used weakly informative priors, $N(0, 2.5)$, instead of the default priors of *brms* for logistic regression (Student-*t* with $df = 3$, $\mu = 0$ and $\sigma = 2.5$). Each model included four sampling chains that ran for 5000 iterations with a warm-up period of 2000 iterations. For each effect we report its mean and 95% CrI under the posterior distribution. As described above, we use CrI to make inference about the presence of an effect. If the 95% CrI for an effect does not include zero, we consider that there is compelling evidence for that effect or the effect is strong or reliable enough to make theoretical inference. We also report the posterior probability of an effect being greater than zero or less than zero depending on the sign of the estimated parameter mean. The posterior probability is calculated by using the posterior sample for a parameter

generated by the statistical model and it is the proportion of the sample less than or greater than zero.

Results

The mean proportions of responses are plotted in **Figure 1** and the analysis of the data is summarized in **Table 1**. In comparison-1, the three conditions that contained either two definite descriptions or one complex demonstrative and one definite description were rated more native-like than the condition with two complex demonstratives (*Dem-Dem*). In comparison-2, we found that there was no main effect of having a complex demonstrative as DP1 or as DP2 – the presence of having just one complex demonstrative did not have any effect on the acceptability of sentences. However, we observed an interaction between these two factors (see the effect *DP1-complex:DP2-complex* in **Table 1**) which denotes that having just one complex demonstrative as either DP1 or DP2 did not affect the “nativeness” of the sentence but having both DPs as complex demonstratives did reduce its “nativeness”. The results of Experiment 1a are thus more compatible with Hinterwimmer (2019) (and **PA 1**) than with either **PA 2** or **PA 3**.

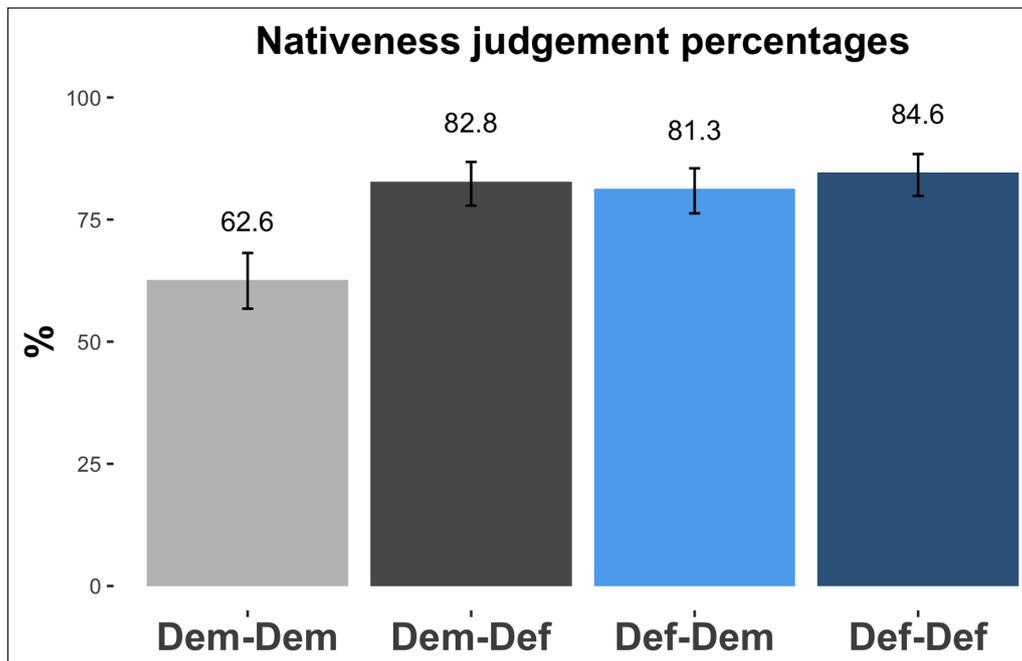


Figure 1: Proportions of trials rated as native in Experiment 1a (the confidence intervals for proportions are calculated using *prop.test()* function in R).

	Effect	Estimate	95% CrI	Post. prob.
Comparison-1	Intercept	0.86	[0.14, 1.64]	0.99
	Dem-Def	1.66	[0.8, 2.67]	1.00
	Def-Dem	1.24	[0.47, 2.05]	1.00
	Def-Def	1.87	[0.78, 3.09]	1.00
Comparison-2	Intercept	2.48	[1.73, 3.36]	1.00
	DP1-complex	-0.05	[-0.96, 0.9]	0.55
	DP2-complex	-0.37	[-1.11, 0.33]	0.85
	DP1-complex:DP2-complex	-1.07	[-2.14, -0.02]	0.98

Table 1: Results from statistical tests for Experiment 1a – estimates of the model for various effects, corresponding 95% credible intervals (95% CrI) and the posterior probabilities (Post. Prob.). Effects for which the credible interval excludes zero are shown in bold. *Comparison-1* evaluates the hypothesis of Hinterwimmer (2019) (and **PA 2**), whereas *comparison-2* evaluates **PA 1** and **PA 3**. In *comparison-1*, the estimates for *Dem-Def*, *Def-Dem* and *Def-Def* show if these conditions are rated better in comparison to *Dem-Dem*. In *comparison-2*, the estimates for *DP1-complex* and *DP2-complex* show if there was any effect of having a complex demonstrative in positions DP1 and DP2, and the estimate for *DP1-complex:DP2-complex* shows the effect of interaction between these two factors.

3.2.2 Experiment 1b

Method

Participants

Forty-five self-reported native speakers of German were recruited through Prolific (<https://prolific.ac/>) (28 females and 1 unreported gender, mean age = 30.5 years, age range = 21–49). For the final analysis data from 3 participants was ruled out because they submitted low effort responses. Each participant was paid £2.

Design and Materials

We used the same design as in Experiment 1a. The experimental and filler items were also the same.

Procedure

The experiment involved an acceptability judgement task. Participants were presented with two sentence short stories as in (14) and had to judge how the second sentence sounds as a continuation of the first and then rate it on the 1–7 Likert scale (1 = completely unacceptable, 7

= completely acceptable). The experiment was programmed online on PCIBex Farm, a platform for running online experiments (Zehr & Schwarz 2018).

Data analysis

As in Experiment 1a, we analysed the data with the same two comparisons and same contrasts. But since in Experiment 1b the responses were on the ordinal scale (ordered values between 1 and 7) we used the mixed-effects ordinal regression in the Bayesian framework following the tutorial on ordinal regression using *brms* (Bürkner & Vuorre 2019). We used the default prior provided by *brms*. The model was fit with a maximal random effects structure. We again make inference using the CrI.

Results

The acceptability ratings are plotted in **Figure 2** and the analysis of the data is summarized in **Table 2**. As in Experiment 1a in comparison-1, the three conditions that contained either two definite descriptions or one complex demonstrative and one definite description were rated better than the condition with two complex demonstratives (Dem-Dem). However, in comparison-2, unlike in Experiment 1a, we found that there was a main effect of having a complex demonstrative as DP1 or as DP2 – the presence of a complex demonstrative reduced the acceptability of sentences. We also observed an interaction between these two factors (see the effect *DP1-complex:DP2-complex* in **Table 2**) which denotes that the reduced acceptability of sentences with two complex demonstratives (Dem-Dem) was over-additive.

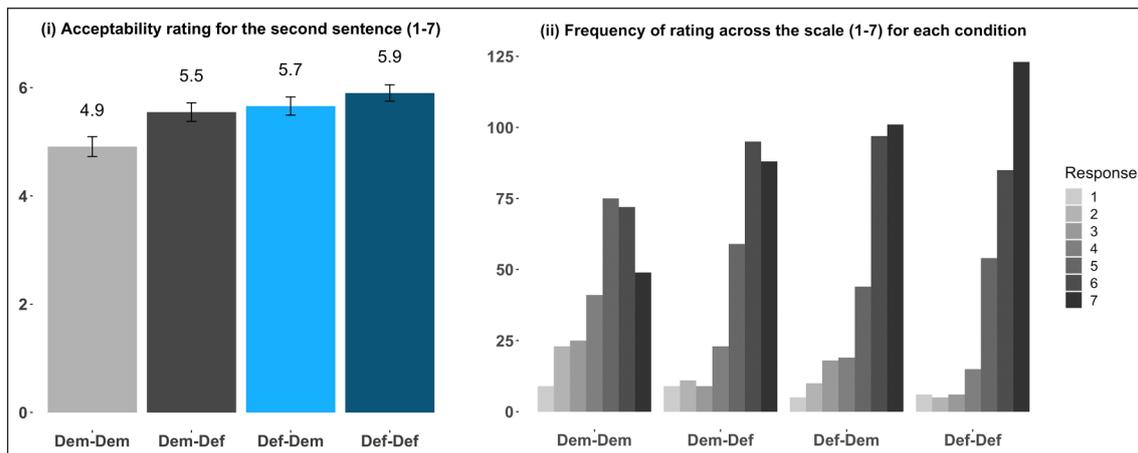


Figure 2: Acceptability rating (1–7 Likert scale: 1 = unacceptable, 7 = acceptable) for Experiment 1b. Panel (i) shows the mean acceptability across conditions, and panel (ii) shows variation in rating across the 1–7 Likert scale for each condition. Note that the mean responses are shown only for the visualization purpose; data analysis was carried out by treating the response as ordinal variable as in the frequency plot.

	Effect	Estimate	95% CrI	Post. prob.
Comparison 1	Dem-Def	0.56	[0.32, 0.8]	1.00
	Def-Dem	0.69	[0.45, 0.93]	1.00
	Def-Def	0.92	[0.64, 1.2]	1.00
Comparison 2	DP1-complex	-0.36	[-0.55, -0.16]	1.00
	DP2-complex	-0.23	[-0.44, -0.01]	0.98
	DP1-complex:DP2-complex	-0.32	[-0.65, -0.01]	0.98

Table 2: Results from statistical tests for Experiment 1b – estimates of the model for various effects, corresponding 95% credible intervals (95% CrI) and the posterior probabilities (Post. Prob.). Effects for which the credible interval excludes zero are shown in bold. *Comparison-1* evaluates the hypothesis of Hinterwimmer (2019) (and **PA2**), whereas *comparison-2* evaluates **PA 1** and **PA 3**. In *comparison-1*, the estimates for *Dem-Def*, *Def-Dem* and *Def-Def* show if these conditions are rated better in comparison to *Dem-Dem*. In *comparison-2*, the estimates for *DP1-complex* and *DP2-complex* show if there was any effect of having a complex demonstrative in positions DP1 and DP2, and the estimate for *DP1-complex:DP2-complex* shows the effect of interaction between these two factors.

Discussion

Taken together, the results from Experiment 1a-b are more compatible with both Hinterwimmer (2019) and **PA 2** than with either **PA 1** or **PA 3**: There is clear evidence that sentences with two anaphoric complex demonstratives are perceived as less acceptable than sentences with two anaphoric definite descriptions and sentences where a complex demonstrative is combined with a definite description. At the same time, the lower acceptability of sentences with two anaphoric complex demonstratives is neither reducible to the combined effect of having one of the two referents picked up by a complex demonstrative instead of a definite description (as predicted by **PA 1**) nor to a violation of the preference for complex demonstratives to pick up the last-mentioned referent (as predicted by **PA 3**), which would result in *Dem-Dem* patterning with *Dem-Def*. Still, for Experiment 1b we found that the presence of an anaphoric complex demonstrative as opposed to a definite description reduced the acceptability of sentences. This indicates that although it does not account for the reduced acceptability of sentences with two anaphoric complex demonstratives, the idea might still be on the right track that complex demonstratives are inherently marked and thus dispreferred whenever they could have been replaced by a definite description without any change in truth conditions (see Ahn 2019).

Concerning the question of whether the analysis proposed by Hinterwimmer (2019) or **PA 2** is ultimately preferable, the experimental results do not provide an answer. As already said in Section 2.2 above, the two analyses essentially make the same predictions regarding the tested pattern. The only difference is that sentences in the *Dem-Dem* condition are predicted

to be strongly deviant by Hinterwimmer (2019) since they necessarily trigger a presupposition violation, while they are predicted to be only mildly deviant by **PA 2**, since they are simply less optimal than the three semantically equivalent, but less complex alternatives that only involve one layer of implicit contrast. The observation that sentences with two complex demonstratives received mean acceptability ratings of almost 5 on the 1–7 Likert scale in Experiment 1b and were judged as native-like in more than 60% of the trials in Experiment 1a therefore provides anecdotal evidence in favour of **PA 2**. Additionally, as already said in Section 2.2 above, **PA 2** is arguably preferable to the account argued for in Hinterwimmer (2019) insofar as it largely builds on assumptions that have been argued for independently and requires less stipulations. It is therefore worthwhile to test the differing predictions of Hinterwimmer (2019) and **PA 2** directly in order to see whether there is clear empirical evidence favouring **PA 2**.

3.3 Experiment 1c

The main goal of this experiment was to make a direct comparison between the size of effects observed in Experiment 1a and 1b, on the one hand, and the size of effects usually observed in presupposition violations and discourse pragmatic violations, on the other. As already said in Section 2.2, the only difference in empirical predictions between Hinterwimmer (2019) and **PA 2** is that the former predicts sentences in the *Dem-Dem* condition to be strongly deviant since they violate a presupposition, while the latter predicts them to be only mildly deviant since they only violate an economy-based pragmatic constraint: They are less optimal than the three semantically equivalent, but less complex alternatives that only involve one layer of implicit contrast, namely *Dem-Def*, *Def-Dem* and *Def-Def*. Consequently, according to Hinterwimmer (2019), the acceptability ratings of sentences in the *Dem-Dem* condition should be similar to the acceptability ratings for sentences clearly violating a presupposition. According to **PA 2**, in contrast, they should be similar to the acceptability ratings of sentences that just violate a discourse pragmatic constraint. In order to test this prediction, we compared ratings for our experimental items (as in (14)) with two types of pragmatic violations (as in (15–16)) and two types of presupposition violations (as in (17–18)).

- (15) Pragmatic violation type 1:
 Mark hat gestern in der Stadt Paulina getroffen.
Mark met Paulina in town yesterday.
- a. **[Good]** Er hat sie erst gar nicht erkannt.
He didn't even recognize her at first
 - b. **[Bad]** Mark hat sie erst gar nicht erkannt.
Mark didn't even recognize her at first.

- (16) Pragmatic violation type 2:
 Lynn sprach einen Reisenden auf den nächsten Zug an, als eine Durchsage
 Verspätungen auf allen Strecken verkündete.
Lynn asked a traveler about the next train when delays on all routes were announced.
- a. **[Good]** Oh Mann, jetzt würde sie bestimmt ihren Anschlusszug
 verpassen.
Oh man, now she would definitely miss her connecting train
- b. **[Bad]** Oh Mann, jetzt würde er bestimmt seinen Anschlusszug
 verpassen.
Oh man, now he would definitely miss his connecting train.
- (17) Presupposition violation type 1:
 Gestern Abend vor meinem Haus haben zwei Hunde so ausgelassen miteinander
 gespielt, dass sie fast eine ältere Frau umgeworfen hätten.
*Last night in front of my house, two dogs played with each other so exuberantly that they
 almost knocked over an elderly woman.*
- a. **[Good]** Die Hunde waren schon ziemlich alt.
The dogs were quite old already.
- b. **[Bad]** Der Hund war schon ziemlich alt.
The dog was quite old already.
- (18) Presupposition violation type 2:
 Peter und Maria haben beide interessante Berufe.
Peter and Maria both have interesting jobs.
- a. **[Good]** Peter ist Linguist und Maria ist auch Geisteswissenschaftlerin.
Peter is a linguist and Maria is a humanities scholar, too.
- b. **[Bad]** Peter ist Linguist und Maria ist auch Tierärztin.
Peter is a linguist and Maria is a veterinarian, too.

Let us go through the four types of violations. Pragmatic violation type 1, exemplified by (15b), can be characterized as follows. It is generally assumed that subjects are more prominent than objects and that the prominence of a discourse referent is correlated with the preference for it to be picked up by a pronoun (see, e. g., Crawley & Stevenson 1990; Gordon et al. 1993). Sentences such as (15b), where the object of the preceding sentence is picked up by a pronoun, while the subject is picked up by repeating the proper name referring to it, are therefore clearly less optimal than sentences such as (15a), where both the subject and the object of the preceding sentence is picked up by a pronoun. The difference in rating between sentences such as (15a) and sentences such as (15b) should thus reflect the decrease in acceptability resulting from the violation of a discourse pragmatic constraint.

Concerning Pragmatic violation type 2, it likewise is related to the prominence status of the discourse referents introduced in the opening sentence. In both conditions the second sentence has to be interpreted as a thought of one of the two discourse referents rendered as *Free Indirect Discourse* (FID). FID is a special form of speech or thought representation in which all context-sensitive and perspective-dependent expressions with the exception of pronouns and tenses are interpreted with respect to the context and the perspective of the discourse referent functioning as implicit speaker or thinker, while pronouns and tenses are interpreted with respect to the narrator's context (Schlenker 2004; Sharvit 2008; Eckardt 2014). In our test items we always ensured that the second sentence has to be interpreted as FID by including interjections such as *Oh Mann* ('oh, man') in combination with expressions such as *jetzt* ('now') and *bestimmt* ('certainly'), which can only be interpreted sensibly with respect to one of the two discourse referent's perspective. Crucially, sentences like (16a) have to be interpreted as expressing the thought of the discourse referent which was referred to by a proper name functioning as the subject of the opening sentence, while sentences such as (16b) have to be interpreted as the thought of the discourse referent which was introduced by an indefinite DP functioning as the object of the opening sentence. The latter discourse referent is thus less prominent than the former in terms of grammatical function as well as type of referential expression and consequently less easily available as implicit thinker. As shown by Hinterwimmer & Meuser (2019), this leads to decreased acceptability in comparison with sentences such as (16a), which can be interpreted as expressing the thought of the more prominent discourse referent. The difference in rating between sentences such as (16a) and sentences such as (16b) should thus also reflect the decrease in acceptability resulting from the violation of a discourse pragmatic constraint.

Let us now turn to the two presupposition violations. In Presupposition violation type 1, exemplified by (17b), the uniqueness presupposition of the definite article (resulting from its combination with a singular NP) is violated, since two discourse referents satisfying the predicate denoted by the NP-complement of the definite determiner have been introduced by the preceding sentence. In (17a), in contrast, where the definite article is combined with a plural NP, the definite DP can be interpreted as referring to the maximal plurality of individuals satisfying the respective predicate (Link 1983). Consequently, no presupposition is violated. The difference in rating between sentences such as (17a) and sentences such as (17b) should thus reflect the decrease in acceptability resulting from the violation of a presupposition.

Finally, in Presupposition violation type 2, exemplified by (18), the presupposition of the focus-sensitive additive particle *auch* ('also', 'too') is violated. Simplifying somewhat, *auch* presupposes there to be a contextually salient proposition that entails a proposition that is (a) a focus alternative (Rooth 1996) and (b) logically independent of the proposition denoted by the sentence containing *auch* (Asher & Lascarides 1998; Beaver & Zeevat 2007; Tonhauser et al. 2013). Since in (18a-b) the subject, *Maria*, is naturally interpreted as focal, the focus alternatives are propositions of the form *x is a humanities scholar* in the case of (18a) and *x is a veterinarian* in the case of (18b), with *x* being

an alternative to Maria. Now, in the case of (18a) the presupposition of *auch* is satisfied since the first conjunct, *Peter ist Linguist* ('Peter is a linguist') entails the proposition *Peter is a humanities scholar*, which is a focus alternative of *Maria_F is a humanities scholar* (the subscript *F* indicates focus marking). In the case of (18b), in contrast, the presupposition of *auch* is violated since the first conjunct does not entail a proposition that is a focus alternative of *Maria_F is a veterinarian*. The difference in rating between sentences such as (18a) and sentences such as (18b) should thus also reflect the decrease in acceptability resulting from the violation of a presupposition.

If Hinterwimmer (2019) is correct and sentences in the *Dem-Dem* condition violate a presupposition, the difference in rating between *Dem-Dem* and the other three conditions should be similar to the differences in rating between the good and the bad sentences in the two presupposition violation conditions. If **PA 2** is correct, in contrast, and sentences in the *Dem-Dem* condition violate a pragmatic constraint, the difference in rating between *Dem-Dem* and the other three conditions should be similar to the differences in rating between the 'good' and the 'bad' sentences in the two pragmatic violation conditions.

Method

Participants

Fifty-two self-reported native speakers of German (47 females and 1 unreported gender, mean age = 21 years, age range = 18–40) were recruited from the University of Cologne for course credits. For the final analysis, data from 8 participants were excluded because they submitted low effort responses.

Design and Materials

We used the same design and experimental items as in Experiment 1a and 1b. Instead of using fillers we added 10 items of each of the four violations described above (15–18). As a result, we had 68 items in total (28 for the main experiment and 40 for the four violations experiments).

Procedure

The experiment involved an acceptability judgement task. Participants were presented with two sentence short stories as in (15–18) and had to judge how the second sentence sounds as a continuation of the first and then rate it on the 1–7 Likert scale (1 = completely unacceptable, 7 = completely acceptable). The experiment was programmed online on PCIbex Farm, a platform for running online experiments (Zehr & Schwarz 2018).

Data analysis

We followed the same data analysis procedures as in Experiment 1b. For the four violations experiments, we contrasted the ratings for 'bad' conditions with the ratings for 'good' conditions.

Results

The acceptability ratings are plotted in **Figure 3** and the analysis of the data is summarized in **Table 3**. As in Experiment 1a and 1b, in comparison-1, the three conditions that contained either two definite descriptions or one complex demonstrative and one definite description were rated better than the condition with two complex demonstratives (*Dem-Dem*). In comparison-2, as in Experiment 1a and unlike in Experiment 1b, we found that the main effects of having a complex demonstrative as DP1 or as DP2 – the presence of having just one complex demonstrative – were not strong enough (based on the posterior probability values of 0.90 and 0.84). However, we observed an interaction between these two factors (see the effect *DP1-complex:DP2-complex* in **Table 3**) which denotes that having just one complex demonstrative as either DP1 or DP2 did not affect the acceptability of the sentence but having both DPs as complex demonstratives did reduce the acceptability. In the case of the four violations experiments, all violations were reliably rated worse than the control conditions. The differences in rating between *Dem-Dem* and the other three conditions (0.5, 0.6 and 0.7) were clearly more similar to the differences in rating between the ‘good’ and the ‘bad’ sentences in the two pragmatic violation conditions (0.7 and 1.0, respectively) than the differences in rating between the ‘good’ and the ‘bad’ sentences in the two presupposition violation conditions (2.6 and 2.7, respectively).

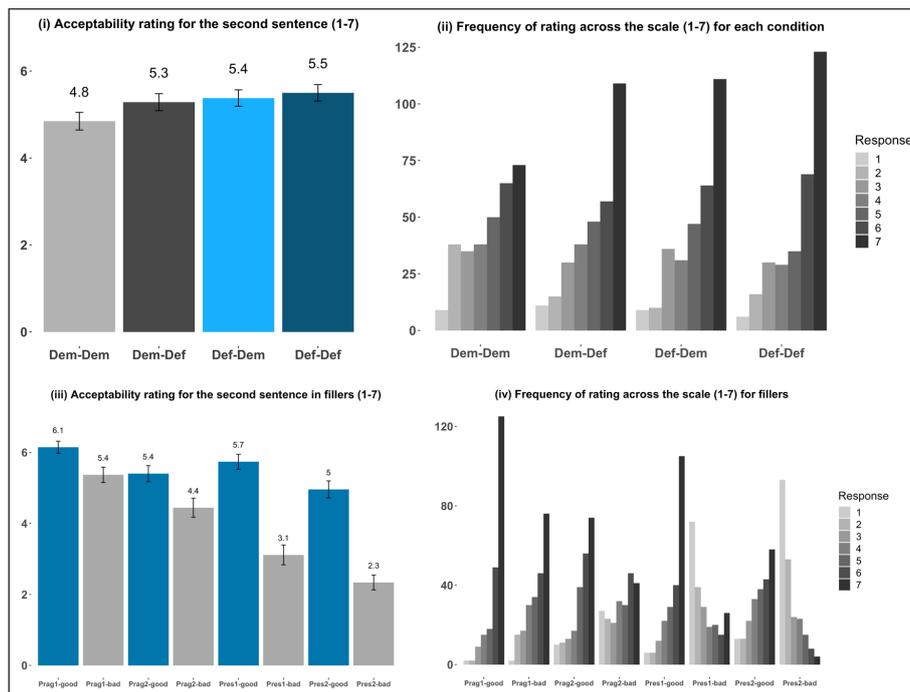


Figure 3: Acceptability rating (1–7 Likert scale: 1 = unacceptable, 7 = acceptable) for Experiment 1c. Panel (i) shows the mean acceptability across conditions, and panel (ii) shows variation in rating across the 1–7 Likert scale for each condition. Note that the mean responses are shown only for the visualization purpose; data analysis was carried out by treating the response as ordinal variable as in the frequency plots.

	Effect	Estimate	95% CrI	Post. prob.
Comparison 1	Dem-Def	0.39	[0.14, 0.64]	1.00
	Def-Dem	0.44	[0.20, 0.68]	1.00
	Def-Def	0.52	[0.27, 0.78]	1.00
Comparison 2	DP1-complex	-0.15	[-0.37, 0.08]	0.90
	DP2-complex	-0.10	[-0.29, 0.10]	0.84
	DP1-complex:DP2-complex	-0.27	[-0.57, 0.02]	0.97
Pragmatic 1	Violation	-0.83	[-1.29, -0.40]	1.00
Pragmatic 2	Violation	-0.78	[-1.21, -0.35]	1.00
Presupposition 1	Violation	-2.02	[-2.73, -1.36]	1.00
Presupposition 2	Violation	-2.04	[-2.56, -1.54]	1.00

Table 3: Results from statistical tests for Experiment 1a – estimates of the model for various effects, corresponding 95% credible intervals (95% CrI) and the posterior probabilities (Post. Prob.). Effects for which the credible interval excludes zero are shown in bold. *Comparison-1* evaluates the hypothesis of Hinterwimmer (2019) (and **PA 2**), whereas *comparison-2* evaluates **PA 1** and **PA 3**. In *comparison-1*, the estimates for *Dem-Def*, *Def-Dem* and *Def-Def* show if these conditions are rated better in comparison to *Dem-Dem*. In *comparison-2*, the estimates for *DP1-complex* and *DP2-complex* show if there was any effect of having a complex demonstrative in positions DP1 and DP2, and the estimate for *DP1-complex:DP2-complex* shows the effect of interaction between these two factors. Estimates in rows named *Pragmatic 1*, *Pragmatic 2*, *Presupposition 1*, and *Presupposition 2* denote the violation effect in the corresponding violation experiment.

Discussion

The results of Experiment 1c, first, confirm the results of Experiments 1a-b and are therefore less compatible with **PA 1** and **PA 3** than with Hinterwimmer (2019) and **PA 2**. Second, and most importantly, they are less compatible with Hinterwimmer (2019) than with **PA2** since the differences in rating between *Dem-Dem* and the other three conditions were much more similar to the differences in rating between the ‘good’ and the ‘bad’ sentences in the two pragmatic violation conditions than in the two presupposition violation conditions. In fact, the difference in rating between *Dem-Dem* and the other three conditions was even smaller than the difference in rating between the ‘good’ and the ‘bad’ sentences in the second pragmatic violation condition, which involves the attribution of FID to the less prominent of two contextually available discourse referents. The presence of two anaphoric complex demonstratives in a sentence thus only seems to be perceived as mildly deviant, comparable to the effect of picking up the more prominent one of two discourse referents by a repeated proper name and the less prominent one by a pronoun, as in the first pragmatic violation condition.

4 Conclusion

In this paper, we have provided experimental evidence from a ‘yes’/‘no’ judgement task and two acceptability rating studies for the claim made in Hinterwimmer (2019) that sentences with two anaphorically interpreted complex demonstratives are less acceptable than sentences with two anaphorically interpreted definite descriptions and sentences where one of two previously introduced referents is picked up by a complex demonstrative, while the other one is picked up by a definite description. The results of the first two experiments (Experiments 1a-b) were in principle compatible with the account proposed by Hinterwimmer (2019), according to which the (potentially abstract) demonstrations presupposed by demonstratives may not have overlapping trajectories. However, sentences with two anaphorically interpreted complex demonstratives were not judged as unacceptable as would be expected if they involved a presupposition violation. Therefore, we have proposed an alternative, economy-based pragmatic account that builds on Ahn (2019) and Nowak (2019). This account is based on the assumption that when uttering a complex demonstrative, the speaker implicitly contrasts the referent of that demonstrative with all other entities that satisfy the predicate denoted by the NP-complement of the demonstrative determiner. At the same time, in sentences with two complex demonstratives picking up individuals that have been introduced by indefinites in the preceding sentence, the two individuals that are referred to by the complex demonstratives only differ with respect to the nominal predicates. Consequently, they are automatically implicitly contrasted with each other as well, resulting in two layers of contrast. The co-presence of two contrasts makes sentences with two anaphoric complex demonstratives unnecessarily complex, however. Therefore, they are dispreferred as compared to sentences with two anaphoric definite descriptions or sentences where an anaphoric complex demonstrative is combined with a definite description, which all involve just one layer of contrast.

The question of whether the observed pattern is more compatible with the account proposed by Hinterwimmer (2019) or the alternative pragmatic account was directly addressed in a further acceptability rating study (Experiment 1c). The design of that study was similar to that of Experiment 1b, but it included as fillers both sentences clearly violating a presupposition and sentences violating a pragmatic constraint. Since the ratings for sentences containing two anaphorically interpreted complex demonstratives were closer to the ratings for sentences violating a pragmatic constraint than for sentences violating a presupposition, we conclude that the alternative pragmatic account is preferable to the account of Hinterwimmer (2019). The phenomena discussed in this paper therefore have important ramifications for our understanding of the interpretative options of complex demonstratives as opposed to definite descriptions: On the one hand, they provide indirect evidence that not just on their deictic, but also on their anaphoric uses, there remain important differences between complex demonstratives and definite descriptions. On the other hand, in order to account for those differences, it is not necessary

to invoke novel concepts such as abstract demonstrations in discourse and their trajectories. Rather, they can be accounted for by combining an analysis of complex demonstratives and definite descriptions along the lines of Elbourne (2008), Nowak (2019) and Ahn (2019) with independently motivated pragmatic economy principles.

Acknowledgements

We would like to thank our student assistants and lab members — Felix Jüstel, Janne Schmandt, Magdalena Schmitz and Sara Meuser — for their help with carrying out the experiments and participation in discussions. The paper improved substantially during the review process — many thanks to the reviewers and to the editors of the special issue, Radek Šimík and Carla Bombi, for their insightful comments and their support. We also thank the Glossa team for facilitating the (open access) publication process.

Funding Information

The work on this paper was funded by the DFG as part of the project *Bound-variable-like interpretations of demonstrative pronouns, complex demonstratives and definite descriptions* (Project-ID – 248411105) and as part of the project *Discourse referents as perspectival centers* within the SFB 1252 *Prominence in Language* (Project-ID 281511265).

Competing Interests

The authors have no competing interests to declare.

References

- Ahn, Dorothy. 2019. *THAT thesis: A competition mechanism for anaphoric expressions*. Cambridge, MA: Harvard University dissertation.
- Asher, Nicolas & Lascarides, Alex. 1998. The semantics and pragmatics of presupposition. *Journal of Semantics* 15. 239–300. DOI: <https://doi.org/10.1093/jos/15.3.239>
- Bach, Emmon & Cooper, Robin. 1978. The NP-S analysis of relative clauses and compositional semantics. *Linguistics and Philosophy* 2(1). 145–150. DOI: <https://doi.org/10.1007/BF00365132>
- Barr, Dale J. & Levy, Roger & Scheepers, Christoph & Tily, Harry J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. DOI: <https://doi.org/10.1016/j.jml.2012.11.001>
- Beaver, David & Zeevat, Henk. 2007. Accommodation. In Ramchand, Gillian & Reiss, Charles (eds.), *The Oxford Handbook of Linguistic Interfaces*, 503–538. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780199247455.013.0017>
- Bisle-Müller, Hansjörg. 1991. *Artikelwörter im Deutschen. Semantische und pragmatische Aspekte ihrer Verwendung*. Tübingen: Niemeyer. DOI: <https://doi.org/10.1515/9783111677590>
- Bürkner, Paul-Christian. 2017. brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software* 80(1). 1–28. DOI: <https://doi.org/10.18637/jss.v080.i01>
- Bürkner, Paul-Christian & Vuorre, Matti. 2019. Ordinal Regression Models in Psychology: A Tutorial. *Advances in Methods and Practices in Psychological Science* 2(1). 77-101. DOI: <https://doi.org/10.1177/2515245918823199>

- Carpenter, Bob & Gelman, Andrew & Hoffman, Matthew D. & Lee, Daniel & Goodrich, Ben & Betancourt, Michael & Brubaker, Marcus & Guo, Jiqiang & Li, Peter & Riddell, Allen. 2017. Stan: A Probabilistic Programming Language. *Journal of Statistical Software* 76(1). 1–32. DOI: <https://doi.org/10.18637/jss.v076.i01>
- Crawley, Rosalind A. & Stevenson, Rosemary J. 1990. Reference in single sentences and in texts. *Journal of Psycholinguistic Research* 19(3). 191–210. DOI: <https://doi.org/10.1007/BF01077416>
- del Gobbo, Francesca. 2003. *Appositives at the interface*. Irvine, CA: University of California at Irvine dissertation.
- Eckardt, Regine. 2014. *The Semantics of Free Indirect Discourse. How Texts Allow Us to Mind-Read and Eavesdrop*. Leiden: Brill. DOI: <https://doi.org/10.1163/9789004266735>
- Elbourne, Paul. 2001. E-type anaphora as NP-deletion. *Natural Language Semantics* 9(3). 241–288. DOI: <https://doi.org/10.1023/A:1014290323028>
- Elbourne, Paul. 2005. *Situations and individuals*. Cambridge, MA: MIT Press.
- Elbourne, Paul. 2008. Demonstratives as individual concepts. *Linguistics and Philosophy* 31(4). 409–466. DOI: <https://doi.org/10.1007/s10988-008-9043-0>
- Elbourne, Paul. 2013. *Definite descriptions*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199660193.001.0001>
- Gordon, Peter & Grosz, Barbara J. & Gilliom, Laura A. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17(3). 311–347. DOI: https://doi.org/10.1207/s15516709cog1703_1
- Hinterwimmer, Stefan. 2019. How to point at discourse referents. On anaphoric uses of complex demonstratives. In Espinal, M. Teresa & Castroviejo, Elena & Leonetti, Manuel & McNally, Louise & Real-Puigdollers, Cristina (eds.), *Proceedings of Sinn und Bedeutung* 23(1). 495–514. DOI: <https://doi.org/10.18148/sub/2019.v23i1.546>
- Hinterwimmer, Stefan & Meuser, Sara. 2019. Erlebte Rede und Protagonistenprominenz. In Engelberg, Stefan & Fortmann, Christian & Rapp, Irene (eds.), *Rede- und Gedankenwiedergabe in narrativen Strukturen – Ambiguitäten und Varianz*, *Linguistische Berichte, Sonderheft 27*. 177–200. Hamburg: Helmut Buske Verlag.
- Kaplan, David. 1989. Demonstratives. In Almog, Joseph & Perry, John & Wettstein, Howard (eds.), *Themes from Kaplan*, 481–563. Oxford: Oxford University Press.
- King, Jeffrey. 2001. *Complex demonstratives*. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/1990.001.0001>
- Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12(5). 607–653. DOI: <https://doi.org/10.1007/BF00627775>
- Larson, Richard. 1982. A note on the interpretation of adjoined relative clauses. *Linguistics and Philosophy* 5. 473–482. DOI: <https://doi.org/10.1007/BF00355583>
- Link, Godehard. 1983. The logical analysis of plurals and mass terms. A lattice-theoretical approach. In Bäuerle, Rainer & Schwarze, Christoph & von Stechow, Arnim (eds.), *Meaning, use and interpretation of language*, 302–323. Berlin: de Gruyter. DOI: <https://doi.org/10.1515/9783110852820.302>

- Nicenboim, Bruno & Vasishth, Shravan. 2016. Statistical methods for linguistic research. Foundational ideas – Part II. *Language and Linguistics Compass* 10(11). 591–613. DOI: <https://doi.org/10.1111/lnc3.12207>
- Nowak, Ethan. 2014. Demonstratives without rigidity and ambiguity. *Linguistics and Philosophy* 37(5). 409–436. DOI: <https://doi.org/10.1007/s10988-014-9159-3>
- Nowak, Ethan. 2019. Complex demonstratives, hidden arguments, and presupposition. *Synthese* 198(1). 2865–2900. DOI: <https://doi.org/10.1007/s11229-019-02250-5>
- Nunberg, Geoffrey. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16(1). 1–43. DOI: <https://doi.org/10.1007/BF00984721>
- Patil, Umesh & Bosch, Peter & Hinterwimmer, Stefan. 2020. Constraints on German *diese* demonstratives. language formality and subject-avoidance. *Glossa* 5(1). DOI: <https://doi.org/10.5334/gjgl.962>
- R Core Team. 2020. R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org>
- Reimer, Marga. 1991. Demonstratives, demonstrations, and demonstrata. *Philosophical Studies* 63(2). 187–202. DOI: <https://doi.org/10.1007/BF00381687>
- Roberts, Craige. 2002. Demonstratives as definites. In van Deemter, Kees & Kibble, Rodger (eds.), *Information sharing. Reference and presupposition in language generation and interpretation*, 89–196. Stanford, CA: CSLI Press.
- Rooth, Mats. 1996. Focus. In Lappin, Shalom (ed.), *The Handbook of Contemporary Semantic Theory*, 271–297. London: Basil Blackwell. DOI: <https://doi.org/10.1111/b.9780631207498.1997.00013.x>
- Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge, MA. Massachusetts Institute of Technology dissertation.
- Schlenker, Philippe. 2004. Context of thought and context of utterance. A note on free indirect discourse and the historical present. *Mind & Language* 19(3). 279–304. DOI: <https://doi.org/10.1111/j.1468-0017.2004.00259.x>
- Sharvit, Yael. 2008. The Puzzle of Free Indirect Discourse. *Linguistics and Philosophy* 31(3). 353–395. DOI: <https://doi.org/10.1007/s10988-008-9039-9>
- Stockwell, Robert P. & Schachter, Paul & Partee, Barbara H. 1973. *The major syntactic structures of English*. New York: Holt, Rinehart and Winston.
- Tonhauser, Judith & Beaver, David & Roberts, Craige & Simons, Mandy. 2013. Towards a taxonomy of projective content. *Language* 89(1). 66–109. DOI: <https://doi.org/10.1353/lan.2013.0001>
- Wolter, Lynsey K. 2006. *That's that. The semantics and pragmatics of demonstrative noun phrases*. Santa Cruz, CA. University of California at Santa Cruz dissertation.
- Zehr, Jérémy & Schwarz, Florian. 2018. PennController for Internet Based Experiments (IBEX). DOI: <https://doi.org/10.17605/OSF.IO/MD832>

