

Interpreting polarity-ambiguous propositional anaphors with negative antecedents: Some experimental results¹

Berry CLAUS — *Humboldt-Universität zu Berlin*

Felix FRÜHAUF — *University of Konstanz*

Manfred KRIFKA — *Leibniz-Centre General Linguistics*

Abstract. The starting point of the present paper is the assumption that negative sentences introduce two propositional discourse referents, one for the negative proposition and one for the negated, positive proposition. Both propositional discourse referents can be picked up by propositional anaphors, resulting in potential ambiguity (e.g. Ernie: *Cookie Monster didn't eat the cookie*. Bert: *Kermit believes that*_[CM didn't eat the cookie / CM ate the cookie]). We report an explorative experimental study on the interpretation of propositional anaphors that are polarity-ambiguous between a resolution with the negative and the positive propositional discourse referent. We employed two different methods, a direct task (forced choice) and a more indirect task (acceptability rating), which yielded mixed results. Taken together, the findings of our study point to a preference for resolving polarity-ambiguous propositional anaphors with the negative propositional discourse referent and they demonstrate the necessity of methodological variety.

Keywords: propositional anaphors, negation, anaphoric ambiguity, anaphora resolution.

1. Introduction

Pronouns and demonstratives such as English *it* and *that* can anaphorically refer to a proposition being introduced in the preceding discourse (e.g. Asher, 1986; Cornish, 1992; Geurts, 1998) by picking up a salient discourse referent that is anchored to that proposition, i.e. a *propositional discourse referent*. For example, A's utterance in (1) introduces the proposition 'that Spencer stole the painting' and the demonstrative in B's response targets the propositional discourse referent for that proposition.

- (1) A: Spencer stole the painting.
B: Louise believes that.

This paper addresses the case of propositional anaphors with a negative clause in the preceding context. Krifka (2013) proposes that clauses with sentential negation introduce two propositional discourse referents, one propositional discourse referent for the negative proposition, henceforth NEGDR, and a second propositional discourse referent for the positive proposition in the scope of the negation operator, henceforth POSDR (see also Snider, 2017). This proposal accounts for the observation that propositional anaphors can pick up the NEGDR as well as the POSDR. Consider the dialogue in (2) with two different responses to A's assertion.

- (2) A: Spencer didn't steal the painting.
B: Katherine believes that_{NEGDR}, too.
B': But Louise believes that_{POSDR}.

¹ This research was partially supported by DFG grant KR 951/7-1 (priority program XPRAG.de).

A's negative assertion is assumed to introduce the NEGDR 'that it is not the case that Spencer stole the painting' and the POSDR 'that Spencer stole the painting'. In (2B), the propositional anaphor *that* picks up the NEGDR, resulting in the interpretation that Katherine believes that Spencer did not steal the painting. In contrast, the propositional anaphor *that* in (2B') picks up the POSDR, resulting in the interpretation that Louise believes that Spencer stole the painting.

Yet, anaphoric reference to POSDRs is restricted. For instance, it is not felicitous when the propositional anaphor is the internal argument of a factive predicate such as *know*. The demonstrative *that* in (3B) can felicitously pick up the NEGDR but not the POSDR.² There are further factors that may constrain the felicity of anaphorically picking up the NEGDR and the POSDR, such as tense and modality and the possibility of accommodation (see Meijer, 2016; Meijer and Repp, 2017). However, with non-factive predicates, such as the belief verb in (3B') and the reporting verb in (3B''), the propositional anaphor *that* can in principle be felicitously resolved with both, the NEGDR and the POSDR. B's response in (3B') can either be read as 'Howard believes that Spencer didn't steal the painting' (NEGDR resolution) or as 'Howard believes that Spencer stole the painting' (POSDR resolution). By the same token, B's response in (3B'') can either be read as 'Howard told me that Spencer didn't steal the painting' (NEGDR resolution) or as 'Howard told me that Spencer stole the painting' (POSDR resolution).

- (3) A: Spencer didn't steal the painting.
 B: Katherine knows that_{NEGDR/#POSDR}.
 B': Howard believes that_{NEGDR/POSDR}.
 B'': Howard told me that_{NEGDR/POSDR}.

Hence, following a negative antecedent clause, utterances with a propositional anaphor as the internal argument of a non-factive predicate are polarity-ambiguous; they can either receive a NEGDR reading or a POSDR reading. The intended reading can be brought out by adding additional material to the utterance, such as an additive particle (e.g. *Howard believes that, too*) or an adversative adverb (e.g. *But Howard told me that*). Also, it may be inferable from the context, common ground, or a continuation sentence (e.g. *Howard told me that. He thinks that Spencer is innocent/guilty*). However, in the absence of any overt material or implicit information that indicates the intended reading, the propositional anaphor is truly ambiguous between a NEGDR and a POSDR resolution. Just as with ambiguous nominal anaphors, there may be a default preference for resolving ambiguous propositional anaphors with the most salient discourse referent (e.g., Ariel, 1990; Gundel, Hedberg, and Zacharski, 1993).

For nominal discourse referents, several saliency-affecting factors have been identified, such as animacy, syntactic function, and order of mention (see Garnham, 2001, for an overview). For obvious reasons, most of these factors are not applicable to the issue of the relative saliency of the two propositional discourse referents being introduced by a negative utterance,

² Note, that according to some authors (e.g. Asher, 1993; Peterson, 1982), the *that* in (3B) does not refer to a proposition but to a different ontological type, i.e. a fact. However, here we will not dwell into this distinction but rather follow Snider (2017) in following Frege (1918) in the view that a fact is a proposition that is supposed to be true.

NEGDR and POSDR. A notable exception is constituent embedding, a factor that has been shown to affect the saliency of nominal discourse referents. The findings of an experimental study on complex DPs by Gordon, Hendrick, Ledoux, and Yang (1999) indicate that discourse referents from non-embedded complex constituents are more easily accessible than discourse referents from embedded constituents. Reading times for clauses containing a pronoun were found to be shorter when the pronoun referred to an entire complex DP than when it referred to a component of the complex DP. Experimental results of Frazier and Clifton (2005) indicate an accessibility difference between matrix clause material and complement clause material, suggesting that the effect of embedding on saliency is not restricted to the level of nominal discourse referents.

Applying the findings of an effect of embedding on saliency to the case at hand leads to the assumption that the NEGDR is more salient than the POSDR, as the former is introduced by a non-embedded constituent, whereas the latter is introduced by an embedded constituent. A different (or additional) reason for assuming a higher saliency of the NEGDR is that it is anchored to a proposition that is asserted in the preceding discourse whereas the POSDR is anchored to a negated proposition.

An alternative assumption emerges when considering that negative sentences are typically uttered in contexts in which the positive proposition is salient already. Accordingly, Krifka (2013) assumes the POSDR to be by default more salient than the NEGDR. In line with this assumption is a finding of psychological language-comprehension research that indicates that when processing a negative sentence, people first mentally represent the negated state of affairs, which corresponds to the positive proposition (Kaup, Yaxley, Madden, Zwaan, and Lüdtké, 2007).

In view of the arguments for the two reverse patterns of relative saliencies of the NEGDR and the POSDR, one could also entertain the assumption that the two propositional discourse referents do not differ in saliency. Though the POSDR may have an a-priori-saliency, it may still be as salient as the NEGDR rather than more salient taking into account that it is introduced by embedded material and is not asserted but negated.

Let's now turn back to the issue of the resolution of polarity-ambiguous propositional anaphors. The assumption of a higher saliency of the NEGDR leads to the prediction that polarity-ambiguous propositional anaphors are by default resolved with the NEGDR. In contrast, the assumption of a higher saliency of the POSDR predicts a default preference for the POSDR resolution. Finally, the assumption of equal saliencies does predict that there is no default preference for the resolution of polarity-ambiguous propositional anaphors.

The aim of the present, explorative study was to investigate how polarity-ambiguous propositional anaphors are resolved and to gain some insight whether there is a default preference. To this end, we conducted a series of three experiments.

2. Experiments

All three experiments were conducted in German. Between experiments, we varied the experimental method and/or aspects of the materials. In Experiment 1, we used a direct task to ex-

amine resolution preferences for polarity-ambiguous propositional anaphors. In Experiment 2, we employed a more indirect task to see whether the findings from Experiment 1 are generalizable to a more natural condition of language comprehension. Finally, Experiment 3 served as a control experiment with the aim of scrutinizing the justification of our conclusion drawn from the results of Experiment 2.

2.1. Experiment 1

In Experiment 1, we employed a two-alternative forced choice (2AFC) task to tap into participants' resolution of polarity-ambiguous propositional anaphors. The task of the participants was to indicate how they interpreted a sentence with an ambiguous propositional anaphor by choosing between two interpretation alternatives, one corresponding to the NEGDR resolution of the propositional anaphor and one corresponding to the POSDR resolution.

More specifically, the participants were presented with short dialogues consisting in an assertion and a response, as in (4). In all target items, the assertion had sentential negation. The response contained a propositional anaphor being realized by the German demonstrative pronoun *das* ('that'). The propositional anaphor constituted the internal argument of a non-factive verb, such that it could be resolved with either of the two propositional discourse referents introduced by the assertion, NEGDR or POSDR. To explore potential effects of information structure, we manipulated the word order of the response clause. The word order was either subject-verb-object (SVO; unmarked) or object-verb-subject (OVS; marked).

(4) Sample target item of Experiment 1

A: Tom hat die Tasche nicht gestohlen.
'Tom didn't steal the bag.'

B: a. Jenny glaubt das. (SVO)
'Jenny believes that.'

b. Das glaubt Jenny. (OVS)
'That is what Jenny believes.'

Interpretation alternatives for the 2AFC task

Jenny glaubt, dass Tom die Tasche nicht gestohlen hat. (NEGDR resolution)
'Jenny believes that Tom didn't steal the bag.'

Jenny glaubt, dass Tom die Tasche gestohlen hat. (POSDR resolution)
'Jenny believes that Tom stole the bag.'

2.1.1. Method

Participants Thirty-two students (18 to 40 years, $M = 26.03$; 22 female) from universities in Berlin and Potsdam participated in the experiment. All were native speakers of German. They gave informed consent for participation and received a monetary reimbursement.

Materials There were 24 target items and 36 filler items. Each item consisted in a dialogue

between two interlocutors, A and B, which comprised two turns: an assertion and a response to it. All assertions pertained to fictitious entities to avoid effects of prior knowledge. The response was a sentence that included an ambiguous propositional anaphor. In all items, it was realized by the German demonstrative pronoun *das* ('that'), which always had object function.

In all target items, the assertion had negative polarity (see the sample item in (4)). There were two versions of the response sentence, differing only in word order. In the SVO version, the order was subject-verb-object; in the OVS version, it was object-verb-subject. Across target items, six different non-factive verbs were used in the response sentences; each of the six verbs was used in four target items. Three of the six verbs were belief verbs: *glauben* ('to believe'), *denken* ('to think'), *erwarten* ('to expect'). The other three verbs were reporting verbs: *sagen* ('to say'), *erzählen* ('to tell'), *berichten* ('to report'). The subject of the response sentence differed from the subject of the assertion and of the two interlocutors. It was a single person, a group of people, or a communication outlet (e.g. radio, newspaper), referred to by a proper name or a role description. Hence, B's response to A's assertion conveyed that some third party believes or reported either of the two propositional discourse referents introduced by A's assertion, NEGDR or POSDR, depending on the resolution of the ambiguous propositional pronoun *das* ('that'). The interpretation alternatives, between which participants had to choose, were paraphrases of B's responses, in which the ambiguous propositional pronoun was replaced with the wording of A's assertion including the negation particle *nicht* ('not') (NEGDR resolution) or excluding the negation particle (POSDR resolution), respectively.

In half of the filler items, the word order in the response sentence was SVO; in the other half, it was OVS. Ten different verbs were used in the response sentences of the filler items: *glauben* ('to believe'; $n = 5$), *denken* ('to think'; $n = 5$), *erwarten* ('to expect'; $n = 5$), *sagen* ('to say'; $n = 5$), *erzählen* ('to tell'; $n = 5$), *berichten* ('to report'; $n = 5$), *befürchten* ('to fear'; $n = 2$), *behaupten* ('to claim'; $n = 2$), *lernen* ('to learn'; $n = 1$), *schreiben* ('to write'; $n = 1$). In all filler items, one of the two interpretation alternatives of the 2AFC task was a paraphrase of B's responses in which the propositional pronoun was replaced with the wording of A's assertion. In most of the 36 filler items ($n = 23$), A's assertion was a complex structure with embedding and the ambiguous propositional pronoun in B's response could either be resolved with the complex proposition or with the embedded proposition. For these filler items, the second interpretation alternative of the 2AFC task differed from the first one in that the ambiguous propositional pronoun was replaced with the wording of the embedded structure of A's assertion. The embedding expression was a propositional attitude ($n = 8$), a modal adverb or verb ($n = 7$), the focus particle *auch* ('too'; $n = 4$) or the focus particle *nur* ('only'; $n = 4$). The embedded structures of A's assertion in the eight items with a focus particle all contained the negation marker *nicht* ('not'). In approximately one half of the remaining 13 filler items, A's assertion contained the indefinite negative article *kein-* ('none'; $n = 7$); in the second interpretation alternative of the 2AFC task, the negative indefinite article was replaced with its positive counterpart. In addition, there were filler items with lexical negation in A's response ($n = 4$), i.e. an adjective with the negation morpheme *un-* (e.g., *ungesund* ['unhealthy']). In the second interpretation alternative of these filler items, the negative adjective was exchanged with its positive version (e.g., *gesund* ['healthy']). Finally, the remaining filler items ($n = 2$) contained an embedding predicate with implicit negation (e.g., *verboten* ['for-

bidden’]) and had a second interpretation alternative, in which the predicate was exchanged with its positive counterpart (e.g., *erlaubt* [‘permitted’]).

Design and Procedure Experiment 1 employed a one-factorial within-subject design with the factor WORD ORDER (svo/ovs). The target items were assigned to two sets of twelve items each and participants were randomly assigned to two groups. The two conditions were allotted to sets and participant groups according to a Latin-square design. Thus, each participant was presented with each target item only once, in one of the two conditions, and each participant received twelve target items per condition. Target and filler items were presented to the participants in four different pseudorandomized orders. The experiment was run as a web study. Each item was presented in three parts. Participants were instructed to read each part carefully. By performing a mouse click, they proceeded to the next part, which was presented below the preceding part. Each item started with the presentation of A’s assertion. This was followed by B’s response. Then, the two interpretation alternatives of the 2AFC task were presented. The participants were instructed to click on the alternative that corresponded to their interpretation of B’s response. The order of the two alternatives was varied across participants. The experimental session lasted approximately 30 minutes.

2.1.1. Results and discussion

Table 1 shows the proportion of choices of the two interpretation alternatives in the two WORD ORDER conditions of Experiment 1. The data were analyzed by using a generalized linear mixed model with a binomial logit function and with participants and items as random factors. The fixed factor was WORD ORDER with sum coding (+.5 for SVO and -.5 for OVS). The analysis revealed a significant effect of WORD ORDER ($b = .64$, $SE = .22$, $z = 2.94$, $p < .01$). The NEGDR interpretation alternative was chosen more often (and the POSDR interpretation alternative was chosen less often) in the OVS condition compared with the SVO condition. Separate χ^2 -tests for the data sets of the SVO condition and OVS condition indicated that in both WORD ORDER conditions, the NEGDR interpretation alternative was chosen significantly more often than the POSDR interpretation alternative (SVO: $\chi^2 = 77.04$, $df = 1$, $p < .001$; OVS: $\chi^2 = 140.17$, $df = 1$, $p < .001$).

WORD ORDER: SVO		WORD ORDER: OVS	
NEGDR: 73%	POSDR: 27%	NEGDR: 80%	POSDR: 20%

Table 1: Proportion of choices of the two interpretation alternatives (NEGDR vs. POSDR) in the two WORD ORDER conditions of Experiment 1.

Overall, the results of Experiment 1 support the hypothesis that propositional anaphors that are ambiguous between a NEGDR resolution and a POSDR resolution are preferably resolved with the NEGDR. However, the results also suggest that the preference for NEGDR can be modulated by word order. The preference for NEGDR was found to be stronger with OVS word-order than with SVO word-order. One possibility to account for the effect of word order is in terms of topicalization. With OVS word-order (but not with SVO word-order), the discourse referent of the sentence-initial propositional pronoun may be conceived of as the topic of B’s response. It has been proposed that embedded constituents and non-embedded constit-

uents differ in their topicalizability, in that only non-embedded constituents are topicalizable in a subsequent utterance (Eckert, 1998: 213). If true, then this implies a higher likelihood of resolving the sentence-initial pronoun with the non-embedded NEGDR than with the embedded POSDR. Another possibility to explain the effect of word order is in terms of a deictic interpretation of the pronoun *das*. With OVS word-order (but not with SVO word-order), *das* may be conceived of as deictic reference, depending on the assignment of prosody when reading the sentence. Whenever the supposed reference of the pronoun is deictic rather than anaphoric, it refers to A's speech act, i.e. neither to POSDR nor to NEGDR. The deictic, speech-act reference, interpretation of *das* can be reconciled with the NEGDR interpretation alternative but not with POSDR interpretation alternative, which should result in more choices of the NEGDR interpretation alternative.

To further explore potential modulatory effects on participants' interpretation choices, we conducted two separate analyses, one for the data of target items with a reporting verb in B's response and one for the data of target items with a belief verb. Table 2 displays the proportion of choices, separately for the two verb types. For the reporting-verb data, there was no significant effect of WORD ORDER ($b = .62$, $SE = .45$, $z = 1.38$, $p = .17$); overall, the NEGDR interpretation alternative was chosen significantly more often than the POSDR interpretation alternative ($\chi^2 = 273.38$, $df = 1$, $p < .001$). In contrast, the analysis of the belief-verb data revealed a significant effect of WORD ORDER on participants' choices of the NEGDR and POSDR interpretation alternatives ($b = .76$, $SE = .26$, $z = 2.86$, $p < .01$). In the SVO condition, the choice pattern for the belief-verb data did not differ from equal distribution ($\chi^2 = 1.33$, $df = 1$, $p = .25$); only in the OVS condition, there was a significant preference for the NEGDR interpretation alternative ($\chi^2 = 21.33$, $df = 1$, $p < .001$).

	WORD ORDER: SVO	WORD ORDER: OVS
Reporting verbs	NEGDR: 91% POSDR: 9%	NEGDR: 94% POSDR: 6%
Belief verbs	NEGDR: 54% POSDR: 46%	NEGDR: 67% POSDR: 33%

Table 2: Proportion of choices of the two interpretation alternatives (NEGDR vs. POSDR) in the two WORD ORDER conditions of Experiment 1, separately for reporting and belief verbs

The results of the explorative post-hoc analyses for the two verb types suggest that the preference for the NEGDR resolution was stronger for the reporting verbs than for the belief verbs. This difference might be due to a discourse-functional difference between the responses with reporting verbs and belief verbs in the present dialogues. For the responses with reporting verbs, there may have been a pragmatic bias to conceive of them as providing evidence for or against the preceding utterance. For the responses with belief verbs, however, it can be assumed that there was no such bias (or at least a weaker one), considering that a belief does not implicate adequate evidence or good reasons. Taking the pragmatic bias with reporting-verb responses for granted, then with a NEGDR resolution of the propositional anaphor, B's response could be regarded as an affirmation of A's assertion, resulting in a smooth dialogue. However, with a POSDR resolution, B's response could be regarded as a rejection of A's assertion. Disagreements with another speaker are assumed to be unfavorable and marked (Krifka, 2013; see also Roelofsen and Farkas, 2015), which may have resulted in a particularly strong tendency to resolve the ambiguous propositional anaphor in reporting-verb responses with the NEGDR rather than with the POSDR, especially when considering

that typical markers of disagreement, i.e. special intonation contour or a contradiction-signaling expression, were not present. This may explain the very high proportion of NEGDR-interpretation choices for the responses with reporting verbs. The absence of the WORD ORDER effect for the data of the reporting verbs might reflect a ceiling effect. However, the interpretation of the different result patterns for the reporting verbs and belief verbs is problematic due to the between-items variation of verb type. There may be other or additional differences between the two item groups that contributed to the difference in choice pattern.

Taken together, the findings of Experiment 1 suggest a preference for the NEGDR resolution of ambiguous propositional pronouns, which can be modulated by word order and possibly by verb type up to an equal distribution of NEGDR and POSDR resolutions. Hence, Experiment 1 does not provide unequivocal evidence for a default resolution preference for polarity-ambiguous propositional anaphors. Yet, the findings stem from a direct method, in which the ambiguity of the sentences with the propositional anaphor as well as the two possible interpretations were made explicit. The goal of Experiment 2 was to investigate the resolution of polarity-ambiguous propositional anaphors under a more natural condition of language comprehension.

2.2. Experiment 2

In Experiment 2, the 2AFC task was replaced with a task in which participants' attention is not drawn to the ambiguity of the propositional anaphor and its two possible interpretations. We employed the same basic material as in Experiment 1 but added a continuation sentence to B's response. The content of the continuation sentence was either consistent with the NEGDR resolution of the propositional pronoun in B's preceding utterance, as in (5B2a), or consistent with the POSDR resolution, as in (5B2b). The task of the participants was to judge the acceptability of the continuation sentence with respect to the preceding context. Our rationale behind employing this experimental task was that we considered the acceptability ratings to be indicative of the interpretation of the ambiguous propositional anaphor. Resolving the anaphor with the NEGDR should result in high acceptability ratings for the NEGDR-consistent continuation sentences and in low ratings for the POSDR-consistent (i.e. NEGDR-inconsistent) continuation sentences. By the same token, resolving the anaphor with the POSDR should result in high acceptability ratings for the POSDR-consistent continuation sentences and in low ratings for the NEGDR-consistent (i.e. POSDR-inconsistent) continuation sentences.

(5) Sample target item of Experiment 2

A: Tom hat die Tasche nicht gestohlen.
'Tom didn't steal the bag.'

B1: a. Jenny glaubt das. (SVO)
'Jenny believes that.'
b. Das glaubt Jenny. (OVS)
'That is what Jenny believes.'

- B2: a. Denn sie ist von seiner Unschuld überzeugt. (NEGDR CONSISTENT)
'Because she is convinced of his innocence.'
- b. Denn sie traut ihm nicht über den Weg. (POS DR CONSISTENT)
'Because she doesn't trust him an inch.'

2.2.1. Method

Participants Participants were 32 students (18 to 40 years, $M = 28.47$; 25 female) from universities in Berlin and Potsdam. All were native speakers of German, gave informed consent to participate in the experiment, and received a monetary reimbursement. The data of one additional participant were replaced because she/he was not a native speaker of German.

Materials The materials comprised 24 target items and 36 filler items. The items were the same as those of Experiment 1, with the exception that B's response was continued with an additional sentence. For all target items, there were two versions of the continuation sentence (see the sample item in (5)). In the NEGDR CONSISTENT version, the content of the continuation sentence was consistent with the NEGDR resolution of the preceding ambiguous propositional anaphor and inconsistent with the POSDR resolution. In the POSDR CONSISTENT version, the content of the continuation sentence was consistent with the POSDR resolution of the ambiguous propositional anaphor and inconsistent with the NEGDR resolution. For the filler items, there was only one version of the continuation sentence. In half of the filler items, the content of the continuation sentence was consistent with the content of the first interpretation alternative of Experiment 1; in the other half, it was consistent with the content of the second interpretation alternative.

Design and Procedure Experiment 2 employed a 2x2 within-subject design with the factors WORD ORDER (SVO/OVS) and CONTINUATION TYPE (NEGDR CONSISTENT/POS DR CONSISTENT). Participants were randomly assigned to four groups, and the experimental items were assigned to four sets. The four conditions resulting from the 2x2 design were allotted to sets and participant groups according to the counterbalancing schema for complex within-subject designs suggested by Pollatsek and Well (1995: 793). The procedure was the same as in Experiment 1, with the exception that the 2AFC task was replaced with a rating task. After reading B's response including the continuation sentence, participants again had to perform a mouse click, which caused the appearance of a 5-point rating scale. The scale ranged from 1 (very bad) to 5 (very good). The participants' task was to judge the acceptability of the continuation sentence by taking into consideration the preceding context.

2.2.1. Results and discussion

Table 3 shows the median ratings in the four conditions of Experiment 2. The data were analyzed by using a cumulative link mixed model for ordinal data (R package ordinal) with participants and items as random effects and WORD ORDER and CONTINUATION TYPE as fixed effects. For both fixed factors, sum coding was used (+.5 for SVO and NEGDR-CONSISTENT; -.5 for OVS and POSDR-CONSISTENT).

CONTINUATION TYPE	WORD ORDER	
	SVO	OVS
NEGDR CONSISTENT	5	5
POSDR CONSISTENT	1	1

Table 3: Median ratings in the four conditions of Experiment 2

The only significant effect was that of CONTINUATION TYPE ($\beta = 3.75$, $SE = .20$, $z = 18.99$, $p < .001$). The ratings were significantly higher in the NEGDR-CONSISTENT condition than in the POSDR-CONSISTENT condition. The ratings were neither affected by WORD ORDER ($\beta = .14$, $SE = .15$, $z = .98$, $p = .33$) nor was there an interaction effect of CONTINUATION TYPE x WORD ORDER ($\beta = -.04$, $SE = .29$, $z = -.14$, $p = .89$).

The results of Experiment 2 point to a strong default preference for the NEGDR resolution of polarity-ambiguous propositional anaphors. The median of the ratings for the NEGDR-CONSISTENT continuations corresponds to the highest level on the rating scale whereas the median of the ratings for the POSDR-CONSISTENT continuations corresponds to the lowest level on the scale. This finding indicates that the propositional pronouns in the target items of Experiment 2 were overwhelmingly resolved with the NEGDR. Different from Experiment 1, there was no modulatory effect of word order.³ Rather, the results point to a strong preference for the NEGDR resolution, irrespective of word order. Thus, it seems that the different methods – direct in Experiment 1 vs. indirect in Experiment 2–yielded mixed results.

In order to facilitate the comparison of the (descriptive) results of Experiment 1 with those of Experiment 2, we recoded the rating data of Experiment 2 according to the presumably underlying interpretation of the ambiguous propositional anaphor. For the NEGDR-CONSISTENT continuations, we considered ratings of 5 and 4 (highest level and second-to-highest level) as indicative of a NEGDR-interpretation (NEGDR-indicative ratings) and ratings of 1 and 2 (lowest level and second-to-lowest level) as indicative of a POSDR-interpretation (POSDR-indicative ratings). Conversely, for the POSDR-CONSISTENT continuations, we considered ratings of 5 and 4 as indicative of a POSDR-interpretation (POSDR-indicative ratings) and ratings of 1 and 2 as indicative of a NEGDR-interpretation (NEGDR-indicative ratings). Ratings of 3 (mid-level) were generally considered to be nonindicative, irrespective of the CONTINUATION TYPE. The distribution of NEGDR-indicative ratings and POSDR-indicative ratings and of nonindicative ratings in the two WORD ORDER conditions is displayed in Table 4a. Table 4b shows the distribution of the three rating categories separately for the two verb types.

The comparison of the proportions of the assumed interpretations in Experiment 2, as being derived from the rating recoding, with the proportions of interpretation-alternative choices in Experiment 1 (see Table 4a and 4b), is in line with the conjecture that the two experiments yielded mixed results regarding the interpretation of polarity-ambiguous propositional

³ Different from Experiment 1, we did not conduct separate explorative analyses for the two verb types (reporting verbs vs. belief verbs) because of the low number of data points per participant and condition ($n = 3$) on which these analyses would have been based. Note, however, that the median ratings in the four experimental conditions of Experiment 2, were for both verb types equal to the overall median ratings as presented in Table 3.

anaphors. In Experiment 1, the NEGDR interpretation alternative was chosen more often (and the POSDR interpretation alternative was chosen less often) with OVS word-order than with SVO word-order. Furthermore, the preference for the NEGDR interpretation was considerably more pronounced for the subset of target items with reporting verbs rather than belief verbs as predicates. In contrast, the pattern of interpretations derived from the ratings of Experiment 2 did neither differ in the two word-order conditions (see Table 4a) nor was there a substantial difference between the two verb types (see Table 4b).

	WORD ORDER	
	SVO	OVS
NEGDR-indicative ratings	83% (Expt 1: 73%)	83% (Expt 1: 80%)
POSDR-indicative ratings	12% (Expt 1: 27%)	11% (Expt 1: 20%)
Nonindicative ratings	5%	6%

Table 4a: Proportion of NEGDR-indicative ratings, POSDR-indicative ratings, and nonindicative ratings in the two WORD ORDER conditions of Experiment 2; the percentages in parentheses show the proportion of choices of the corresponding interpretation alternative (NEGDR vs. POSDR) in Experiment 1.

	Reporting verbs	Belief verbs
	NEGDR-indicative ratings	85% (Expt 1: 92%)
POSDR-indicative ratings	10% (Expt 1: 8%)	12% (Expt 1: 39%)
Nonindicative ratings	5%	6%

Table 4b: Proportion of NEGDR-indicative ratings, POSDR-indicative ratings, and nonindicative ratings in Experiment 2, separately for the two verb types; the percentages in parentheses show the proportion of choices of the corresponding interpretation alternative (NEGDR vs. POSDR) in Experiment 1.

An obvious reason for the different findings is the methodological difference between the two experiments. First, the direct task of Experiment 1 may have prompted special response strategies. Second, the indirect task in Experiment 2 may not have been suitable to investigate the interpretation of polarity-ambiguous propositional anaphors. We will turn back to the first issue in the Conclusion. The second issue was addressed in Experiment 3.

2.3. Experiment 3

The purpose of Experiment 3 was to scrutinize the rationale behind the method that was employed in Experiment 2. It was designed to examine whether it is justified to assume that the acceptability data for NEGDR- vs. POSDR-consistent continuations reflect the interpretation of the propositional anaphor in the preceding sentence. To this end, we investigated how the acceptability of the continuation sentences is judged when the resolution of the propositional anaphor is constrained to either the NEGDR or the POSDR. The material was the same as in Experiment 2, except that the two versions of B's first response sentence did not differ in

word order (constantly SVO) but in whether they enforced a NEGDR-interpretation or a POSDR-interpretation of the propositional anaphor, via the additive particle *auch* 'too' or the adversative adverb *aber* 'but', respectively (see (6)). The task of the participants was the same as in Experiment 2. If it is true that the acceptability judgments reflect the resolution of the propositional anaphor, then the two interpretation-manipulation conditions should yield reverse results. In the NEGDR-enforcing condition, the acceptability ratings should be high for the NEGDR-consistent continuations and low for the POSDR-consistent continuations, whereas in the POSDR-enforcing condition, they should be high for the POSDR-consistent and low for the NEGDR-consistent continuations.

(6) **B's first response sentence in the two interpretation-manipulation versions of Experiment 3 for the sample target item in (5)**

B1: a. Jenny glaubt das auch. (NEGDR-enforcing)
'Jenny believes that, too.'

B1: b. Jenny glaubt das aber. (POSDR-enforcing)
'But Jenny believes that.'

2.3.1. Method

Participants Participants were 34 students (18 to 40 years, $M = 25.79$; 27 female) from universities in Berlin and Potsdam. All were native speakers of German, gave informed consent to participate in the experiment, and received a monetary reimbursement.

Materials The materials were the same as those of Experiment 2, with the following two modifications. First, in all 24 target items and 36 filler items, the word order of B's first response sentence was SVO. Second, for all target items, there were two versions of B's first response sentence. In the *AUCH*-[NEGDR] version, it contained the additive focus particle *auch* ('too') to enforce the NEGDR resolution of the propositional anaphor. In the *ABER*-[POSDR] version, B's first response sentence contained the adversative adverb *aber* ('but') to enforce the POSDR resolution of the propositional anaphor. For the filler items, there was only one version of B's first response sentence. In 20 filler items, it either contained the particle *auch* ('too') or the adverb *aber* ('but'); in the remaining 26 filler items, it contained neither of the two expressions.

Design and Procedure Experiment 3 employed a 2x2 within-subject design with the factors INTERPRETATION MANIPULATION (*AUCH*-[NEGDR]/*ABER*-[POSDR]) and CONTINUATION TYPE (NEGDR CONSISTENT/POSDR CONSISTENT). The four conditions were counterbalanced across four participant groups and four sets of items (cf. Pollatsek and Well, 1995). The procedure was the same as in Experiment 2.

2.3.1. Results and discussion

The median ratings in the four conditions of Experiment 3 are shown in Table 5. The data were submitted to a cumulative link mixed model analysis with participants and items as ran-

dom effects. The fixed factors were INTERPRETATION MANIPULATION and CONTINUATION TYPE with sum coding (+.5 for *AUCH*-[NEGDR] and NEGDR-CONSISTENT; -.5 for *ABER*-[POS DR] and POSDR-CONSISTENT). The analysis yielded no significant main effect of INTERPRETATION MANIPULATION ($\beta = -.01$, $SE = .14$, $z = -.03$, $p = .98$). There was a significant main effect of CONTINUATION TYPE ($\beta = .93$, $SE = .14$, $z = 6.57$, $p < .001$), which was qualified by a significant interaction between the two factors ($\beta = 5.68$, $SE = .33$, $z = 17.37$, $p < .001$). To examine the interaction, we conducted separate analyses for the data subsets of the two INTERPRETATION MANIPULATION conditions. In the *AUCH*-[NEGDR] data subset, the ratings were significantly higher in the NEGDR-CONSISTENT condition than in the POSDR-CONSISTENT condition ($\beta = 4.19$, $SE = .30$, $z = 14.07$, $p < .001$). In the *ABER*-[POS DR] data subset, the ratings were significantly higher in the POSDR-CONSISTENT condition than in the NEGDR-CONSISTENT condition ($\beta = -1.79$, $SE = .20$, $z = -8.92$, $p < .001$).

CONTINUATION TYPE	INTERPRETATION MANIPULATION	
	<i>AUCH</i> -[NEGDR]	<i>ABER</i> -[POS DR]
NEGDR CONSISTENT	5	2
POS DR CONSISTENT	1	5

Table 5: Median ratings in the four conditions of Experiment 3

The main finding of Experiment 3 is that the interpretation manipulation via *auch* ('too') vs. *aber* ('but') resulted in reverse acceptability-rating patterns. When a NEGDR-interpretation was enforced, the rating pattern was the same as the overall pattern of Experiment 2, with higher ratings in the NEGDR-consistent condition compared with the POSDR-consistent condition. This part of the results of Experiment 3 bolsters the conclusion that the rating pattern observed in Experiment 2 with ambiguous propositional anaphors reflects a strong preference for the NEGDR resolution. When a POSDR-interpretation was enforced, the ratings were higher in the POSDR-consistent version of the continuation sentence compared with the NEGDR-consistent version. This indicates that the rating pattern of Experiment 2 was not due to a general low acceptability of the POSDR-consistent version of the continuation sentences. As to the specific objective of Experiment 3, its findings do not corroborate the suspicion that the indirect method of Experiment 2 may not have been suitable to investigate resolution preferences for polarity-ambiguous propositional anaphors. Rather, the reverse result patterns for the two INTERPRETATION MANIPULATION conditions strengthen the notion that the ratings of the continuation sentences in Experiment 2 are indicative of the interpretation of the preceding ambiguous anaphor.

3. Conclusion

The goal of our experimental study was to explore how propositional anaphors that are polarity-ambiguous between a NEGDR interpretation and a POSDR interpretation are resolved. We considered three hypotheses, (1) a default preference for the NEGDR interpretation, under the assumption that the NEGDR is more salient than the POSDR; (2) a default preference for the POSDR interpretation, under the second, contrary assumption that the POSDR is more salient than the NEGDR; (3) no default preference, under the third assumption that the two propositional discourse referents do not differ in saliency.

In Experiment 1, we employed a 2AFC task, in which participants were asked to choose between a NEGDR-interpretation alternative and a PosDR-interpretation alternative. The results suggest a preference for the NEGDR interpretation of polarity-ambiguous propositional anaphors, which can be modulated by word order and possibly by verb type up to an equal distribution of NEGDR and PosDR interpretations with SVO word-order and belief verbs as predicate. One possible account of this finding is in terms of the no-default-preference hypothesis. According to this hypothesis, the two antecedent candidates, NEGDR and PosDR, do not differ in saliency. Consequently, other factors could take effect on the resolution of the ambiguous anaphor, such as characteristics of the linguistic material (e.g. word order and verb type). However, as was pointed out in the discussion of Experiment 1, the different results for the two verb types must be taken with a large grain of salt because the items in the two subsets differed not only in the predicate. Moreover, the modulating effect of word order could not be replicated in Experiment 2, in which we employed an indirect method to tap into the interpretation preferences of the participants.

In Experiment 2, the task of the participants was to rate the acceptability of a continuation sentence that was either consistent with the NEGDR resolution and inconsistent with the PosDR resolution of the polarity-ambiguous propositional anaphor in the preceding sentence or vice versa. The ratings were found to be affected solely by the type of the continuation sentence. The median of the ratings for the NEGDR-consistent continuation sentences corresponded to the highest level on the rating scale whereas the median of the rating for the PosDR-consistent continuation sentences corresponded to the lowest level, suggesting that the ambiguous anaphors in the preceding sentences were overwhelmingly resolved with the NEGDR. This conclusion was bolstered by the findings of Experiment 3.

In Experiment 3, we employed the same rating task as in Experiment 2. However, the interpretation of the polarity-ambiguous propositional anaphors was restricted by adding a constraining expression. The ratings were determined by an interaction of the interpretation restriction with the content of the continuation sentences. When the NEGDR-interpretation was enforced (via the additive particle *auch* ['too']), the rating pattern was the same as in Experiment 2. In contrast, when the PosDR-interpretation was enforced (via the adversative adverb *aber* ['but']), the ratings were significantly higher for PosDR-consistent continuation sentences compared with the NEGDR-consistent continuation sentences.⁴ This pattern of results strongly suggests that the ratings of the continuation sentences are indicative of the resolution of the preceding ambiguous anaphor. Correspondingly, the method employed in Experiment 2 can be considered as suitable for uncovering resolution preferences.

⁴ Note, that it is not the mere presence of the additive particle *auch* ('too') vs. adversative adverb *aber* ('but') which enforces the NEGDR- vs PosDR-resolution of polarity-ambiguous propositional anaphors. Rather, it is crucial which resolution alternative affords the additive vs. adversative function of the expressions. This is illustrated in (i) with negative versions of the sample item's response sentence, which show a pattern that is opposite to their positive counterparts: *Auch* enforces the PosDR-interpretation (see (iB)) and *aber* enforces the NEGDR-interpretation (see (iB')).

(i) A: Tom hat die Tasche nicht gestohlen. 'Tom didn't steal the bag.'
 B: Jenny glaubt das_{PosDR} auch nicht. 'Jenny doesn't believe that_{PosDR} either.'
 B': Aber Jenny glaubt das_{NEGDR} nicht. 'But Jenny doesn't believe that_{NEGDR}.'

With regard to the mixed findings of Experiment 1 and Experiment 2, we propose that they can be reconciled by taking into consideration, differences in the experimental situation that are inherent to the different methods in the two experiments. The forced choice task that we employed in Experiment 1 involved a rather artificial setting, in which the ambiguity of the sentences containing the propositional anaphor and the two possible interpretations were made explicit. By contrast, the rating task of Experiment 2 did not draw participants' attention to the ambiguity of the propositional anaphor and its two possible interpretations. In this regard, the experimental situation in Experiment 2 constituted a somewhat more natural condition of language comprehension than that of Experiment 1. We assume that the rating data that we obtained in Experiment 2 reflect the spontaneous resolution of the polarity-ambiguous anaphors. That is, we act on the assumption of a strong default preference for the NEGDR resolution under unprompted conditions⁵. It is likely, that the particularities of the direct task in Experiment 1 induced a tendency to override the default NEGDR-preference. More specifically, the presentation of the two interpretation alternatives may have prompted the participants to consciously or unconsciously reassess their interpretation of the sentence containing the propositional anaphor. This may have enhanced their sensitivity to characteristics of the linguistic material (such as word order and predicate type), which, in turn, resulted in favouring one of the two presented interpretation alternatives over the other.

As was already mentioned, we conclude from the findings of Experiment 2, that there is a default preference to resolve a polarity-ambiguous propositional anaphor with the NEGDR. One may object that the findings do not reflect a true default, which is generally in effect. Rather, the preference for the NEGDR might be due to particular features of our items. It might have been crucial that the material consisted in dialogues and that the target anaphor was contained in a reply to an assertion of another speaker. There may be a bias to favour responses that can be considered to provide evidence for the interlocutor's assertion or that convey that the interlocutor's belief is shared by a third party over responses that can be considered to provide evidence against the interlocutor's assertion or that convey that a third party has an opposite belief, resulting in a pragmatically driven preference for the NEGDR interpretation of the response. Though this pragmatic-bias account does not rule out that there is a default preference for the NEGDR, it does cast doubt on the soundness of that conclusion: Taking the pragmatic bias for granted, it should not matter whether and which default preference there is. That is, when there actually is no default preference there should still be a preference for the NEGDR interpretation due to the pragmatic bias, and when there actually is a default preference for the POSDR then this may be overridden by the pragmatic bias. Yet, there are two reasons to doubt that the assumed pragmatic bias can fully account for the findings of Experiment 2. First, it seems reasonable to assume that the pragmatic bias is stronger for reporting verbs than for belief verbs (see the discussion of Experiment 1). Thus, the preference for the NEGDR interpretation in Experiment 2 should have been more pronounced for the items with reporting verbs. Second, if the resolution preference for the NEGDR was solely due to a pragmatic bias, then the content of a POSDR-consistent continuation should have acted as a cue to the intended meaning of the preceding sentence. Thus, the finding that the

⁵ It remains to be seen whether this assumption can be experimentally supported with data from a genuine indirect method, e.g. reading-time or eye-tracking data. Based on the findings of Experiment 2, we expect shorter processing times and fewer regressions to the preceding context for the NEGDR-consistent continuations compared with the POSDR-consistent continuations.

PosDR-consistent continuation sentences overwhelmingly received low ratings is challenging for the pragmatic-bias account but it can be readily explained when assuming a default preference for the NEGDR due to its high saliency. However, it remains an empirical task to evaluate the validity and scope of the pragmatic-bias account.

More generally, future research is needed to disentangle default-preference effects from materials- or task induced effects. This necessitates the use of fine-grained online measures of anaphor processing, such as eye tracking and ERP data. A further relevant task for future research is to put the assumptions as to the relative saliencies of the NEGDR and the POSDR to a pure test by directly testing their accessibility with a plain measure, i.e. uncontaminated by concurrent task demands. However, this requires developing novel methodological approaches considering that established measures of accessibility that have been employed in studies on nominal discourse referents may be inapplicable to the investigation of the accessibility of propositional anaphors.

In conclusion, the present study explored the interpretation of polarity-ambiguous propositional anaphors, which has not been experimentally addressed before. The implications of our findings are two-fold. First, they (re)demonstrate the necessity of methodological variety in research on anaphor resolution. Second, they point to a preference for resolving polarity-ambiguous propositional anaphors with the NEGDR rather than with the POSDR, which may reflect a default preference stemming from a higher saliency of the NEGDR.

References

- Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.
- Asher, N. (1986). Belief in discourse representation theory. *Journal of Philosophical Logic* 15(2), 127–189.
- Asher, N. (1993). *Reference to abstract objects in English: a philosophical semantics for natural language metaphysics*. Dordrecht: Kluwer Academic Publishers.
- Cornish, F. (1992). So be it: The discourse–semantic roles of *so* and *it*. *Journal of Semantics* 9(2), 163–178.
- Eckert, M. (1998). *Discourse deixis and null anaphora in German*. PhD thesis. University of Edinburgh.
- Frazier, L. and C. Clifton (2005). The syntax-discourse divide: processing ellipsis. *Syntax* 8(2), 121–174.
- Frege, G. (1918). Der Gedanke. Eine logische Untersuchung. In *Beiträge zur Philosophie des deutschen Idealismus I*, pp. 58–77.
- Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Hove: Psychology Press.
- Geurts, B. (1998). Presuppositions and anaphors in attitude contexts. *Linguistics and Philosophy* 21(6), 545–601.
- Gundel, J. K., N. Hedberg, and R. Zacharski (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69(2), 274–307.
- Kaup, B., R. H. Yaxley, C. J. Madden, R. A. Zwaan, and J. Lüdtker. (2007). Experiential simulation of negated text information. *Quarterly Journal of Experimental Psychology* 60(7), 976–990.

- Krifka, M. (2013). Response particles as propositional anaphors. In T. Snider (Ed.), *Proceedings of Semantics and Linguistic Theory 23*, pp. 1–18. Ithaca, NY: CLC Publications.
- Meijer, A. M. (2016). Referring to proposition 'inside' a negated one: the modal particles *auch* and *ook* in contextual counterfactuals. In K. Bellamy, E. Karvovskaya, and G. Saad (Eds.), *Proceedings of 24th Conference of the Student Organisation of Linguistics in Europe (ConSole XXIV)*, pp. 19–44. Leiden: Leiden University Centre for Linguistics.
- Meijer, A. M. and S. Repp (2017). Modal subordination of propositional anaphora: On the role of tense and the modal particle *ook* in contextual counterfactuals in Dutch. In R. Truswell, C. Cummins, C. Heycock, B. Rabern, and H. Rohde (Eds.), *Proceedings of Sinn und Bedeutung 21*, pp. 881–898. Edinburgh: University of Edinburgh.
- Peterson, P. L. (1982). Anaphoric reference to facts, propositions, and events. *Linguistics and Philosophy* 5(2). 235–276.
- Pollatsek, A. and A. D. Well. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21(3), 785–794.
- Roelofsen, F. and D. Farkas. (2015). Polarity particle responses as a window onto the interpretation of questions and assertions. *Language* 91(2), 359–414.
- Snider, T. (2017). *Anaphoric reference to propositions*. PhD thesis. Cornell University.