

German Modal Particles as Discourse Signals

Hannah J. Seemann
Ruhr-Universität Bochum

HANNAH.SEEMANN@RUB.DE

Tatjana Scheffler
Ruhr-Universität Bochum

TATJANA.SCHEFFLER@RUB.DE

Editor: Amir Zeldes

Submitted 01/2024; Accepted 01/2025; Published online 01/2025

Abstract

This study investigates the German modal particles *ja* and *doch* in discourse relations. We conduct an acceptability study of modal particles in four discourse relations (CIRCUMSTANCE, CONDITION, EVIDENCE, JUSTIFY) to test predictions of (in)compatibilities derived from a corpus study by Döring and Repp (2019). As ratings for sentences representing the discourse relations CIRCUMSTANCE and CONDITION were significantly lower than for the two causal relations if presented with a modal particle, we confirm that modal particles and discourse relations interact. In a forced-choice study testing the particle *ja*'s effect on relation disambiguation, we show that *ja* supports a causal interpretation of an ambiguous context in the absence of explicit discourse markers. Our findings contribute to delineating the role of German modal particles in discourse, as we show that there is an interaction between discourse relations and modal particles, meaning that readers do not accept all modal particles in every discourse relation, and at least the modal particle *ja* can serve as a non-connective discourse signal for causal relations.

Keywords: Modal particles; discourse particles; discourse relations; discourse signals; *ja*; *doch*; German

1. Introduction

Speakers have a variety of linguistic means at their disposal to indicate their attitudes, knowledge, and internal states. In English, typical expressions are adverbs like *obviously* or *probably*, modal verbs like *might*, and certain verbs like *to doubt* or *to know* (Biber, 2006; Gray and Biber, 2015). In the German language, modal particles can carry some of these meanings (Zimmermann, 2011).

German modal particles, also sometimes referred to as ‘discourse particles’, are noninflected sentence modifiers that semantically scope over the whole sentence, but do not affect its truth conditions. In discourse, they can be used as epistemic markers to express assumptions about interlocutors’ shared knowledge, to express the author’s attitude towards a proposition, or to indicate a speaker’s (un)certainty (Hartmann, 1979; Zimmermann, 2011).

Each particle makes a specific contribution to a discourse. For example, *ja*¹ signals that the speaker assumes a piece of information to be either known, obvious in a given context, or easily

1. Since there is no clear English equivalent for the German modal particles, we did not attempt to translate them in the text or in the example glosses.

verifiable, see (1). This function imposes restrictions on the use of *ja*: It typically cannot be used if the information is not previously known to the hearer, as shown in (2) – but see Kratzer (1999) for a discussion of exceptions.

- (1) Käsekuchen ist **ja** mein Lieblingskuchen.
 cheesecake is JA my favorite.cake
 ‘As you and I assume/know, cheesecake is my favorite cake.’
- (2) #Du wusstest bisher nicht, dass Käsekuchen **ja** mein Lieblingskuchen ist.
 you knew until.now not that cheesecake JA my favorite.cake is
 Intended: ‘You did not know before that cheesecake is my favorite cake.’

Similar to *ja*, the particle *doch* signals information to be previously known, but additionally indicates contrast (Karagjosova, 2004). This contrast might be between what a speaker expects someone else to know and the knowledge they show, as in B’s utterance in (3), or between the current and a previous utterance, as indicated by utterance B’.

- (3) A: I made your favorite dessert: chocolate cake!
 B: Aber mein Lieblingsdessert ist **doch** Käsekuchen.
 but my favorite.dessert is DOCH cheesecake
 ‘But my favorite dessert is cheesecake (and you should know that).’
 B’: Aber du hattest **doch** gesagt, dass du Käsekuchen backst?
 but you had DOCH said that you cheesecake bake
 ‘But you said you would bake a cheesecake, didn’t you?’

Thus, modal particles interact with discourse by managing the integration of information and enabling speakers’ tendencies to indicate their own knowledge and beliefs. Previous studies have even stated that particles may themselves signal discourse relations, but their exact role in discourse is not yet clear (Helbig and Buscha, 2001; Döring, 2016).² Whether modal particles are discourse markers is discussed on both functional and formal terms in much detail in the collected volumes by Degand et al. (2013) and to a lesser extent in Fedriani and Sansò (2017). However, the discussions on German modal particles and discourse markers are limited to theoretical considerations, concluding that ‘discourse marker’ refers to a function, whereas ‘modal particle’ is a (morphosyntactic) word class. None of the mentioned studies analyse the German modal particles’ contribution in discourse as studies on connectives and other discourse signals do.

To address this gap, the present study investigates the interaction between modal particles and discourse relations from an experimental perspective. As previous work has shown that modal particles interact with discourse in the sense that they are only found in certain discourse relations that are specific for every particle (Döring, 2016), the aim of this study is to delineate the nature of this interaction further: Is the use of a particle restricted by the discourse relation? And does the particle itself play a role in which discourse relation is identified for a given context?

2. It should be noted that in the German research tradition, ‘discourse marker’ means something else than in the English-speaking (or Dutch-speaking) tradition. In the latter – and the field of discourse (structure) studies – conjunctions and other connectives indicating the discourse relations that hold between segments in a discourse are called ‘discourse markers’ (Fraser, 1999; Taboada, 2006). In German linguistics, the term ‘discourse marker’ refers mostly to conversation-structuring units, such as *ich meine* (‘I mean’) or English *well*. Historically, these units have alternatively been referred to as ‘discourse particles’ or ‘pragmatic markers’, terms which also have been used to refer to modal particles (Das, 2014; Blühdorn, 2017).

The following sections first summarize the theoretical background on German modal particles as well as discourse relations and their signalling in general, before we discuss previous work on modal particles in discourse. We then present an acceptability study (§6), testing the modal particles *ja* and *doch* in four different discourse relations. We show that readers are sensitive to the interaction of discourse relations and modal particles because they do not accept modal particles equally in the presented discourse relations. We follow this up with a forced-choice study (§7) which investigates whether the German modal particle *ja* functions as a discourse signal. The results show that *ja* is a causal discourse signal, as the presence of this particle in a sentence shifts the interpretation of an ambiguous discourse relation in favor of causal relations.

2. German Modal Particles

From a syntactic and semantic point of view, there is a vast amount of research on German modal particles – an overview can be found in Diewald (2009), Zimmermann (2011), and Müller (2017). Nonetheless, to offer an overview of modal particles in general and the two particles studied here in particular, this section summarizes their (mostly) agreed-upon characteristics.

Modal particles are an open word class (Thurmair, 2013). Diewald (2009) describes 15 particles as belonging to the ‘core class’ of modal particles, with an open set of additional modal particles in the periphery. While it is not precisely defined which particles belong to the periphery of the class, most authors discussing modal particles list at least the majority of the particles that Diewald includes in the core class: *aber, auch, bloß, denn, doch, eben, eigentlich, etwa, halt, ja, mal, nur, schon, vielleicht, wohl*, see e.g., Thurmair (1989); Hartmann (1998); Müller (2014); Abraham (2017). While different modal particles have different functions, they are all restricted to the middle field (*Mittelfeld*) of a sentence, indicating that they cannot be topicalized or focused. In addition, there are restrictions to the sentence types and speech acts in which the particle can occur: Some particles, like *denn*, are restricted to interrogative sentences, other modal particles are not restricted to one sentence type. Though modal particles are sometimes taken to be obligatory in certain contexts, they are typically not viewed as indicators of sentence types (for discussion on the particles’ distribution cf. Karagjosova, 2004; Kwon, 2005; Thurmair, 2013).

Since modal particles do not add any truth-conditional meaning, it is difficult to translate them into English without adding additional lexical content containing the particle’s expressive meaning, as seen in examples (1) and (3) (cf. Gutzmann, 2016). In German, modal particles have homonyms in other word classes, e.g., *ja* can also be an answer particle (‘yes’) and *doch* can also be a conjunction (‘but’). Moreover, the expressive meaning of *doch* as a particle changes depending on whether it is stressed or not, as seen in example (4). In (4a) *doch* is not stressed, while in (4b), it is.³

- (4) a. Das ist **doch** keine Kegelrobbe, das ist eine Ringelrobbe.
 that is DOCH no grey.seal that is a ringed.seal
 ‘What? That is not a grey seal, it is a ringed seal (and you should know that).’
 b. Das ist **DOCH** keine Kegelrobbe, das ist eine Ringelrobbe.
 that is DOCH no grey.seal that is a ringed.seal.
 ‘I was wrong! That is not a grey seal, it is a ringed seal.’

3. It should be noted that the status of accented *doch* as modal particle is still up for debate. Some authors treat it as a particle (Gutzmann, 2010; Hogeweg et al., 2011; Egg and Zimmermann, 2012; Rojas-Esponda, 2014a), while others consider it to be an adverb instead (Blühdorn, 2019; Abraham, 2020, p. 283).

The two particles studied in this article are *ja* and *doch*. As mentioned above, with the use of *ja*, the speaker assumes information to be known or highly salient to the hearer. It marks its prejacent proposition as not at-issue, indicating that it is not the main point itself, but relevant for understanding the main at-issue contribution (cf. Potts (2005) and e.g., Thurmair (1989); Kwon (2005)). With *doch*, a speaker signals that information is assumed to be known to the hearer, but not necessarily part of the currently activated shared knowledge. Without *doch*, the sentence in (4a) would simply state a fact about seal taxonomy. The use of *doch* expresses the speaker's belief that the hearer ought to already know this. The aspect of contrast is kept if *doch* is stressed, but the reference point for the contrast changes: In (4b), the speaker corrects their own previous statement/belief, not the listener's (Egg and Zimmermann, 2012; Rojas-Esponda, 2014b).

Based on their function of marking information as known, both particles should not co-occur with verbs presenting new information (Thurmair, 1989; Kratzer, 2004; Gutzmann, 2015; Viesel, 2017). Restrictions of modal particle use in embedded clauses are described in Coniglio (2012): Modal particles can only appear in adverbial sentences that have illocutionary force independent of the matrix clause. Therefore, it is argued that they cannot be used in embedded conditional or temporal clauses, see (5) and (6).

- (5) #Falls es **ja** regnet, muss ich die Blumen nicht gießen.
 If it JA rains must I the flowers not water
 Intended: 'I don't need to water the plants if it rains.'
- (6) #Während ich **ja** draußen die Blumen gieße, schaut die Katze von drinnen zu.
 While I JA outside the flowers water look the cat from inside at
 Intended: 'While I water the plants outside, the cat is watching me from inside the house.'

If a modal particle can be used in a sentence introduced by a conjunction typically categorized as temporal, like *nachdem* ('after'), Coniglio (2012) views this as evidence for the conjunction having grammaticalized from a strictly temporal reading to a causal or contrastive conjunction. In example (7a), *nachdem* can be replaced by *da* ('since'), illustrating its use as a causal conjunction. Adding the modal particle *ja* to the subordinate clause is infelicitous if the same sentence is modified to allow only a temporal reading.

- (7) a. Nachdem er **ja** immer gesagt hatte, ich könnte ihn jederzeit anrufen, habe ich das
 gestern auch gemacht.
 'Because he always said I could call him anytime, I did just that yesterday.'
- b. #Nachdem er **ja** immer gesagt hatte, ich könnte ihn jederzeit anrufen, hat er nun aller-
 dings sein Telefon ständig ausgeschaltet.
 Intended: 'After always saying that I could call him at any time, he has now switched
 his phone off all the time.'

(adapted from Hentschel, 1986, p. 203)

Before testing how these functions and restrictions translate to modal particle use in a discourse structure framework, the following sections offer an overview of the discourse model used in this study.

(8) a. [The cat got wet]_N [because it went outside during the rain.]_S (CAUSE)
 b. [In order to hunt its prey,]_S [the cat climbed the tree.]_N (PURPOSE)
 c. Yesterday, the cat did the following things: [It ate,]_N [it climbed a tree,]_N [it fell asleep on the bed.]_N (LIST)

Beyond the function of establishing text coherence, discourse relations are also thought to be mental representations of information conveyed in the text (Sanders et al., 1992), as well as processing instructions helping to construct the connection between segments of a text (Canestrelli et al., 2013). Earlier research focused on the differences in processing causal vs. additive relations, finding that causal relations between segments lead to better understanding and shorter reading times compared to additive relations (Sanders and Noordman, 2000; Sanders, 2005). Sanders (2005) proposes that this effect is due to language users assuming a causal relation between segments of text by default, so that establishing a non-causal relation between given segments leads to additional processing costs. Recently, more types of relations and their processing have been studied, finding that negative as well as subjective relations like CONCESSION or CONTRAST are harder to process compared to positive relations (Kuperberg et al., 2011; Canestrelli et al., 2013; Köhne-Fuetterer et al., 2021).

5

Furthermore, it has been found that the explicit marking of a relation with a connective or cue phrase – as in (8a) and (8b) – helps both in reducing processing time and improving understanding of a text because the connective functions as an instruction for identifying the discourse relation (Millis and Just, 1994; Degand and Sanders, 2002; Kamalski, 2007; Sanders and Spooren, 2007). Explicit marking of relations is especially helpful for negative and subjective relations that express a speaker’s reasoning or judgment, as it helps to speed up the additional step of processing the propositional content as well as the negation or the speaker’s stance towards it (Kleijn et al., 2019; Wei et al., 2019; Crible and Pickering, 2020).

4. Relation Signals

Signals that help with the identification and processing of relations can appear in many forms. Some of these are connectives like *because* or *in order to*; relations marked by a connective are referred to as ‘explicit’ relations, relations without a connective are ‘implicit’ (Prasad et al., 2008). Recent studies have focused on identifying other types of (non-connective) relation signals and their effects on text processing, as well as implicit discourse relations themselves. Das et al. (2015) present the RST Signalling Corpus, a study on the different kinds of relation signals in RST annotated texts. They find that most relations are signaled by non-connective signals or by a combination of connectives and other signals, and only 10.65% of all relations are exclusively signaled by connectives (Das and Taboada, 2019). Other signals the study identifies are lexical (non-connective cue words or phrases like *I concede*), morphological (tense), syntactic (word order, sentence mood) or semantic (synonymy, lexical chains). In addition, numerical (characters indicating lists) and graphical signals (punctuation, headings), genre and references to entities are identified as signals of relations (Das and Taboada, 2019). Crible (2020) expands on this investigation by studying which types of signals co-occur. She reports that 32.55% of all relations are marked exclusively by connectives, while all other cases feature combinations with other signals.

Many of the signals identified by Das and Taboada (2019) have been shown to affect the processing of discourse positively by speeding up the processing and/or leading to a better understanding of the presented text. This has been shown for e.g., antonyms and resultative verbs (Crible and Demberg, 2020), parallelisms (Crible and Demberg, 2020; Crible and Pickering, 2020), quantifying expressions (Scholman et al., 2020), enumerative structures (Péry-Woodley et al., 2017), and verbal tense (Grisot and Blochowiak, 2017). One general finding of these studies is that the signal disambiguates an ambiguous discourse context without the need for a connective to guide participants to an interpretation of the relation, as observed in detail by Das and Taboada (2019) and Crible (2020). Second, the effect of the signal is generally stronger if there is no connective present or if the connective is ambiguous, like *and* (Asr and Demberg, 2012; Crible, 2020; Das and Taboada, 2019).

Moreover, epistemic markers have also been shown to give processing instructions similar to connectives. Canestrelli et al. (2013) compared Dutch causal connectives without further marking and found that sentences with the subjective connective *want* (‘because’) were processed more slowly compared to those with the objective connective *omdat* (‘because’).⁵ This effect was canceled out by the addition of the epistemic marker *volgens* (‘according to’) to the first part of the

5. ‘Subjective’ and ‘objective’ refer to the source of coherence: A subjective connective might indicate that the speaker is presenting their own interpretation, while an objective connective might be used if observations without additional evaluation are reported (cf. e.g., Sanders, 2005; Canestrelli et al., 2013).

sentence because the additional processing of the subjectivity indicated by *want* already happened at the epistemic marker. In a similar vein, Wei et al. (2020) studied which (epistemic) stance markers and connectives co-occur in Chinese. They found that the neutral connective *suoyi* ('so') appears more frequently with epistemic markers than the corresponding connective *kejian* ('so') that marks an utterance as subjective. Both results indicate that as processing instruction, encoding subjectivity (= the speaker's/author's perspective, cf. Sanders et al. (2021)) at one point of an utterance is both satisfying for the speaker/author as well as sufficient for the hearer/reader. In German, modal particles are used to mark the epistemic states of the interlocutors (Doherty, 1985; Zimmermann, 2011; Dörre et al., 2018). This again motivates the question whether the German particles influence the perception of discourse relations, as has been shown for epistemic markers in other languages. Whether there is a similar interaction between modal particles and discourse connectives in German might be an interesting question for future work.

5. Previous Work on Modal Particles in Discourse

Modal particles have frequently been suspected to be markers in discourse, but few studies have explicitly investigated the contribution of German modal particles in a discourse structure framework. Other work has studied the processing of modal particles without including the discourse context. One of the early mentions of modal particles as indicators of text coherence is presented in Helbig and Buscha (2001):

- (9) a. Ich gehe nicht schwimmen, das Wasser ist **ja** noch viel zu kalt.
 I go not swimming the water is JA still much too cold
 'I don't go swimming because the water is still far too cold.'
 b. Ich gehe nicht schwimmen, **weil** das Wasser noch viel zu kalt ist.
 I go not swimming because the water still much too cold is
 'I don't go swimming because the water is still far too cold.'

(Helbig and Buscha, 2001, p. 429)

The authors argue that in (9), *ja* can serve a function similar to the conjunction *weil* ('because'), causally connecting the two arguments. The authors do not comment on the fact that the causality can be inferred even without any marker at all (cf. Sanders, 2005), nor do they further elaborate on the discourse connecting function of modal particles.

The first explicit connection between modal particles and discourse relations is drawn in a corpus study conducted by Döring and Repp (2019). The authors investigate whether there are modal particles that appear more often with certain discourse relations and whether speakers have preferences regarding modal particle choice in different relations. They find that the modal particles studied occur more or less often in certain discourse relations that are specific for each particle. For example, *ja* is frequently found in BACKGROUND or causal relations (e.g., EVIDENCE, JUSTIFY), but never in temporal relations like CIRCUMSTANCE or in CONDITION. In addition, the position of the particle in each relation differs, e.g., whether the particle is found in the nucleus or the satellite. While *ja* is found almost exclusively in the satellite of the relations it appears in, *doch* is found in the nucleus of certain relations such as CONCESSION as well – although in general, it is still more frequently used in the satellite. In example (10), discussed by Döring (2016, pp. 165-167), an EVIDENCE relation holds between [2] (nucleus) and [3] (satellite). Döring argues that using *ja* in the satellite of EVIDENCE, which is where the evidence for the claim (= the nucleus) is presented, the

evidence is marked as being already in the Common Ground. By this, a speaker might raise the listener's acceptance of the claim. Since rejecting backgrounded or inferred information is difficult, using modal particles for certain effects might be an argumentative strategy.

- (10) [1] Die Repräsentanten der Gewerkschaften wie auch Sie im Hause haben in
the representatives of.the unions like also you in.the house have in
Wahrheit **doch** erkannt – [2] das zeigt die Debatte heute –, [1] daß [...] die große
truth DOCH realized that shows the debate today that the great
Mehrheit unserer Mitbürgerinnen und Mitbürger längst erkannt hat, daß um der
majority of.our fellow.citizens long.ago realized has that for the
Sicherung der Zukunft willen Veränderungen [...] notwendig sind. [3] Wolfgang
security of.the future PR changes necessary are Wolfgang
Schäuble hat **ja** die neuesten Umfragedaten bekanntgegeben.
Schäuble has JA the latest survey.data announced
'In fact, the representatives of the unions and you here in this house have realized – this
shows today's debate – that the majority of our fellow citizens realized long ago that
changes are necessary to secure the future. Wolfgang Schäuble has announced the latest
survey data.'

(Döring, 2016, p. 165; KOHLCORPUS, Speech 4, #10690)

To confirm their corpus findings about the particles' preferences, Döring and Repp (2019) conduct a forced-choice study in which participants have to choose one out of three given modal particles for BACKGROUND or JUSTIFY contexts. They conclude that there is an interaction between modal particles and discourse structure, and that speakers can use the particles deliberately to convey a certain effect. In a larger corpus study investigating additional modal particles and how they interact with discourse structure, Döring (2016) concludes that the particles do not just co-occur randomly with discourse relations, but only in specific combinations.

Yet, experimental work on the processing of German modal particles is scarce and does typically not account for the discourse context. Dörre et al. (2018) compare modal particles with their homonyms to find out whether the modal particle (= the not-at-issue) reading of, e.g., *einfach*, leads to higher processing costs compared to its counterpart, the adjective *einfach* ('simple'). They find higher reading times for the modal particle readings of the target words compared to their counterparts. This finding is in accordance with studies on the processing of discourse relations, even though these studies did not test modal particles: Köhne-Fuetterer et al. (2021) report higher processing times for relations more complex than causality, Canestrelli et al. (2013) suggest that subjective relations have higher processing costs than objective relations, and Van Bergen and Bosker (2018) find increased processing times for markers of inter-subjective meaning. Further investigating the differences between modal particles and their homonyms, Reimer and Trotzke (2019) study whether reading times of *nur* and *bloß* (both 'only') differ depending on whether the particle is presented in its modal or focus particle reading. Surprisingly, even though both particles are generally more frequently used as focus particles than as modal particles, Reimer and Trotzke find higher reading times for *nur* in the modal particle reading and for *bloß* in the focus particle reading. They conclude that there is a preference for one interpretation given a certain particle, independent of the frequency of that reading.

And while it is frequently suspected that, given their discourse-structuring functions, modal particles might be signals or markers in discourse (e.g., in Dahl, 1988; Thurmair, 1989; König and Requard, 1991; Helbig and Buscha, 2001; Döring, 2016), this assumption is never put to the test explicitly. The present study aims at this research gap, presenting experimental evidence for modal particles as discourse signals.

6. Experiment 1: Acceptability of Modal Particles in Discourse Relations

The first experiment tested whether adding a modal particle to a sentence leads to lower mean acceptability ratings if the relation context presented is not expected to allow for the modal particle. This study is preregistered, access to the registration is provided in the Data Availability section.

6.1 Hypotheses

Previous results suggest that there is an interaction between modal particles and (the processing of) discourse. The goal of the present study is to offer more experimental evidence of this interaction, testing whether readers are sensitive to it, as well as delineating the particles' role as discourse signals. Even though Döring and Repp (2019) do verify their corpus results in a forced-choice study, this is so far the only experimental evidence for the interaction of modal particles and discourse relations. Therefore, using a different experimental design, we also tested some of the reported corpus results experimentally to further confirm that modal particles are restricted to certain discourse relations. In our acceptability study, we investigate the following hypotheses:

H1: Acceptability of sentences with modal particles is lower for certain discourse relations.

H2: The acceptability of a modal particle in a discourse relation also depends on the particle's position (= nucleus or satellite).

To ensure comparability with the findings of the corpus data presented in Döring and Repp (2019), the modal particles chosen for all experiments are *ja* and *doch*. In their study, the authors report frequent occurrences of *ja* and *doch* in the discourse relations EVIDENCE and JUSTIFY,⁶ and no occurrences in the relations CIRCUMSTANCE and CONDITION. Therefore, we refer to CIRCUMSTANCE and CONDITION as 'dispreferred' and to EVIDENCE and JUSTIFY as 'preferred' relations for these two modal particles, and we expect sentences with particles in dispreferred relations to have lower mean acceptability ratings than the same sentence without a particle. Even though Döring and Repp (2019) report on the frequency in more than four relations, these four relations will be used in the rating study because they show similar frequency patterns for both modal particles. We thus expect ratings for CIRCUMSTANCE and CONDITION to be lower than the ratings for EVIDENCE and JUSTIFY if the presented sentences contain a modal particle. In addition, we expect to see higher ratings if the particles are presented in a relation's satellite than for the nucleus, given that this is the position they are found in most often in the corpus. Based on these predictions, we can specify our hypotheses:⁷

6. The definition of JUSTIFY differs from the definition of the same relation in Stede (2016a), and is more similar to what is called REASON there. To construct the target sentences for this experiment, the relation definitions given in Döring and Repp (2019) have been used.

7. The specified hypotheses were not part of the original preregistration. We have nonetheless decided to include them here in order to further specify the vague original hypotheses.

H1': The acceptability of sentences with modal particles is lower in the relations CIRCUMSTANCE and CONDITION (= 'dispreferred') than in EVIDENCE and JUSTIFY (= 'preferred').

H2': The acceptability of sentences with modal particles is lower if the particle is presented in the relation's nucleus.

6.2 Participants

92 participants (mean age 35.5 years, SD 12.5 years), all native speakers of German were recruited via Prolific.⁸ Participants received compensation based on Germany's minimum wage (2 EUR for 10 minutes). One submission was excluded for failing two out of three attention checks, as predetermined in the study preregistration. Since modal particles are rarely found in written text and readers might tend to rate any written sentence containing a modal particle as 'bad', we instructed people as follows:

This experiment contains 40 sentences. The sentences have been produced by a software trying to change the style of formal texts to informal style. The software does so by inserting various words in the original, formal texts. Unfortunately, the software tends to be overzealous and also inserts words that don't fit the context. Your task is to rate the transformed sentences. A sentence is good if you find the style to be colloquial, but not odd. A sentence is bad if you stumble over certain words while reading.⁹

The instruction itself was kept in a colloquial style to prepare the participants for reading rather colloquial than formal written sentences.

6.3 Materials and Design

We used a 3 x 2 x 2 design that accounts for the modal particle (*ja*, *doch*, no particle), the discourse relation (preferred: EVIDENCE and JUSTIFY, dispreferred: CIRCUMSTANCE and CONDITION) as well as the position of the particle (nucleus or satellite). Definitions for the four discourse relations are given in Appendix A. For each discourse relation, we designed five experimental items that were presented in five different versions: With *ja* or *doch* in the nucleus, with *ja* or *doch* in the satellite, or without particle, see (11)–(14). All relations are explicitly marked by a connective: CIRCUMSTANCE is marked by *während* ('while'), CONDITION is marked by *falls* ('if'), EVIDENCE is marked by *da* ('because'), and JUSTIFY is marked by *weil* ('because'). The 20 experimental items were distributed over five pre-determined lists to ensure every participant reads sentences from all conditions, but each sentence only once.

(11) CIRCUMSTANCE

Während Ingo (Ø/*ja/doch*) in der Straßenbahn sitzt, liest er (Ø/*ja/doch*) die Zeitung.
 'While Ingo (Ø/*JA/DOCH*) is on the tram, he (Ø/*JA/DOCH*) reads the newspaper.'

8. <https://www.prolific.com/>

9. Original German instruction: 'Dieses Experiment besteht aus 40 Sätzen. Produziert wurden die Sätze von einer Software, die den Stil von existierenden Texten auflockern soll. Dies tut die Software unter anderem, indem sie verschiedene Worte in ursprünglich eher nüchtern geschriebene Texte einfügt. Leider ist die Software stellenweise etwas übereifrig und fügt manchmal auch Worte ein, die eigentlich an der jeweiligen Stelle gar nicht passen. Ihre Aufgabe ist die Bewertung der fertigen Sätze. Ein Satz ist gut, wenn Sie den Stil als (tendenziell) umgangssprachlich, aber trotzdem nicht merkwürdig empfinden. Ein Satz ist schlecht, wenn Sie beim Lesen über bestimmte Worte stolpern.'

- (12) **CONDITION**
 Falls Fiona (\emptyset /*ja/doch*) nicht bald aufbricht, wird sie (\emptyset /*ja/doch*) zu spät kommen.
 ‘If Fiona (\emptyset /JA/DOCH) doesn’t leave soon, she will (\emptyset /JA/DOCH) be late.’
- (13) **EVIDENCE**
 Da Helena (\emptyset /*ja/doch*) häufig auf Konzerten ist, kann sie (\emptyset /*ja/doch*) viele Lieder mitsingen.
 ‘Because Helena (\emptyset /JA/DOCH) frequently visits concerts, she can (\emptyset /JA/DOCH) sing along to many songs.’
- (14) **JUSTIFY**
 Weil Jakob (\emptyset /*ja/doch*) Medizin studieren will, muss er (\emptyset /*ja/doch*) einen guten Abschluss machen.
 ‘Because Jakob (\emptyset /JA/DOCH) wants to study medicine, he has (\emptyset /JA/DOCH) to get a good degree.’

All sentences were constructed according to the same pattern: The relation’s satellite including a typical connective first, and the nucleus second. If applicable, the modal particles were always placed after the name/pronoun, in the middle field of the sentence. It has to be noted that for some of the sentences containing *doch*, ambiguity with the particle’s adverbial reading cannot be avoided. This limitation will be further discussed in §6.6.

In addition to the 20 experiment items, participants read three training items, 17 fillers and three attention checks asking participants to choose a specific point of the rating scale. Some of the filler items contained (different) modal particles to draw attention away from *ja* and *doch* and to balance the number of sentences with/without particles as well as the fraction of sentences that are expected to have good or bad ratings.

After having read each sentence, participants were asked to rate the sentence on a 7-Point Likert Scale (1: ‘bad’, 7: ‘good’), prompted by the question “How well did the software transform the sentence?”.

6.4 Analysis

We fit two ordinal regression models¹⁰ with varying intercepts and slopes using the R package *ordinal* (Christensen, 2022). The two models are necessary to deal with model deficiency (arising from the constraint that ‘position’ can not be accounted for if no particle is present in a sentence) in a combined model. Given that Döring and Repp (2019) report the same occurrence pattern of *ja* and *doch* for the relations used in this study, we expect the same rating patterns for both particles. However, to verify that there are no statistically significant differences between the two, we fit our first model using Helmert coding to check for differences between *ja* and *doch* compared to the no particle condition in the ratings of the presented sentences. Possible differences between the two particles might arise due to their slightly different meaning contributions: Both signal known information, but only *doch* also indicates contrast. Since we did not find a statistically significant main effect for the difference between the two particles, we fit the second model to compare the ratings of both particles combined in the positions nucleus and satellite. In both models, we coded the relations

10. This deviates from the preregistered analysis plan because it was pointed out to us that this model fits the ordinal data from the Likert scale rating better. The analysis had to be split into two models to account for rank deficiency in the combined model.

according to our assumptions defined in the beginning of this section as ‘preferred’ (EVIDENCE, JUSTIFY) and ‘dispreferred’ (CIRCUMSTANCE, CONDITION) and report on the maximum random effect structure.

6.5 Results

Figure 1 provides an overview of the descriptive results of our study, the ratings for the four relations by modal particle and position. In all four relations, sentences without particle were rated highest. Furthermore, if there is a difference between particles or positions, there is a tendency for *ja* to be rated higher than *doch*, and for a particle in the satellite to be rated higher. One exception is the relation CONDITION: *Doch* in the satellite is rated as more acceptable than *ja*. Without modal particles, the ratings for CIRCUMSTANCE and JUSTIFY are high (median = 7), followed by CONDITION (median = 6) and EVIDENCE (median = 5).

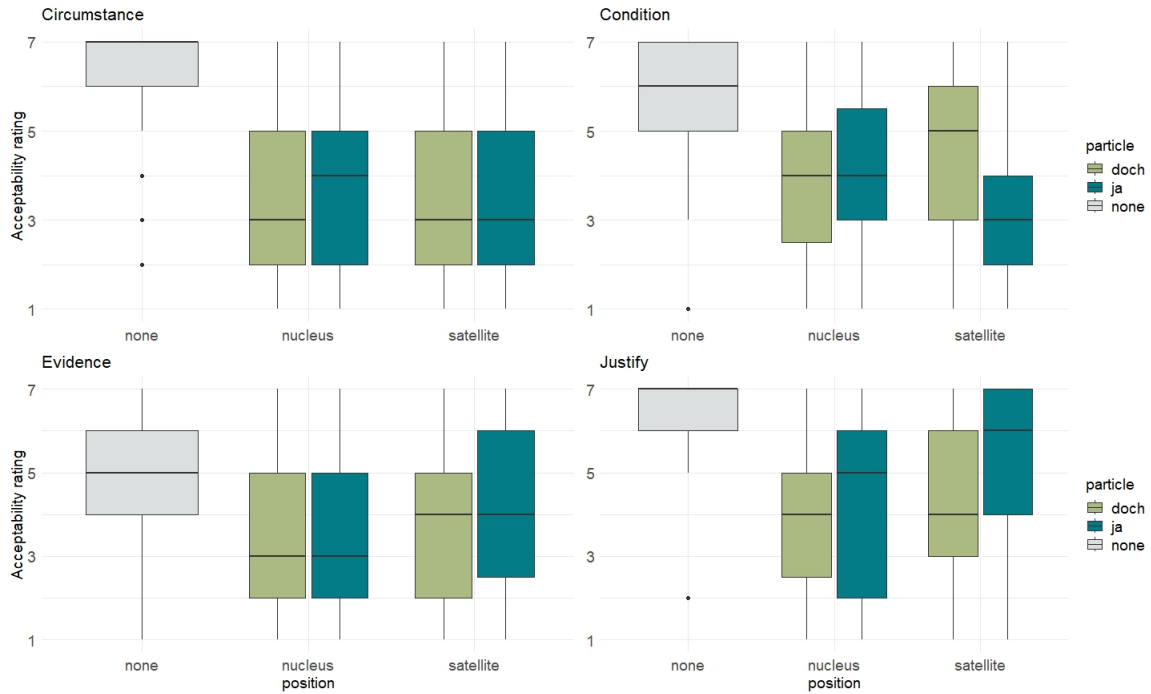


Figure 1: Acceptability ratings for the two ‘dispreferred’ discourse relations CIRCUMSTANCE, CONDITION and the two ‘preferred’ relations EVIDENCE and JUSTIFY. The rating scale ranges from 1 (= rejection) to 7 (= acceptance). Ratings are divided by the particle’s position (nucleus vs. satellite) and by modal particle (*doch* vs. *ja*).

The observations in Figure 1 are confirmed by our first model (shown in Table 1). The model shows that the comparison of ‘no particle’ vs. ‘particle’ is significant ($p < 0.001$), but the difference between the two particles *ja* and *doch* is not ($p = 0.15$). The main effect of preferred vs. dispreferred relation is not significant in general, but the significant interaction between relation and presence of particle indicates that the particle’s presence influences the ratings ($p < 0.05$). This is illustrated

in Figure 2: The mean rating of sentences in dispreferred relations drops more in the presence of a particle than in relations that are assumed to be acceptable with the given modal particle.

Fixed effects	Estimate	Std. Error	z-value	p-value
preferred vs. dispreferred relation	0.092	0.392	0.234	0.815
\emptyset vs. (<i>ja/doch</i>)	-3.030	0.291	-10.398	< 0.001
<i>ja</i> vs. <i>doch</i>	-0.235	0.166	-1.414	0.157
relation : \emptyset vs. (<i>ja/doch</i>)	1.026	0.507	2.023	< 0.05
relation : <i>ja</i> vs. <i>doch</i>	-0.913	0.325	-2.810	< 0.01

Table 1: Output of the cumulative link mixed model for ordinal data shown below. We use Helmert coding to compare first ‘no particle’ to (*ja/doch*) and *ja* and *doch* in a second step. We use sum coding for the relations, combined to ‘preferred’ (EVIDENCE, JUSTIFY) and ‘dispreferred’ (CIRCUMSTANCE, CONDITION).

```
clmm(rating ~ relation * particle +
      (1+particle*relation|participant) + (1+particle|item))
```

While the difference between the two particles in general is not significant, the interaction of relation and *ja* vs. *doch* is ($p < 0.01$). Since in three out of four relations, the two particles show similar rating trends – as visible in Figure 1 – this interaction is likely to be driven by the rating of *doch* in CONDITION. In the preferred relations, *ja* is rated as being more acceptable than *doch* in the relation’s satellite. No such difference is visible in the dispreferred relation CIRCUMSTANCE. Only in CONDITION, *doch* is rated as more acceptable than *ja*. This difference might arise from an ambiguity of the particle *doch* (which has an additional stressed reading, elaborated in section §6.6), rather than a difference between occurrences of the unstressed modal particles *ja* and *doch*.

The second model (Table 2) shows the comparison of a sentence with any of the two particles in the nucleus or satellite. Ratings for sentences without particles are not included.

Fixed effects	Estimate	Std. Error	z-value	p-value
preferred vs. dispreferred relation	0.4474	0.2994	1.494	0.13513
nucleus vs. satellite	-0.2588	0.2067	-1.252	0.21063
relation : position	-1.1468	0.3981	-2.880	< 0.01

Table 2: Output of the cumulative link mixed model for ordinal data shown below. We use sum coding to compare both position (nucleus, satellite) and relation, combined to ‘preferred’ (EVIDENCE, JUSTIFY) and ‘dispreferred’ (CIRCUMSTANCE, CONDITION).

```
clmm(rating ~ relation * position +
      (1+relation*position|participant) + (1+position|item))
```

The model shows no significant main effect for position ($p = 0.21$), but a significant interaction between relation type and position ($p < 0.01$). The differences in the ratings of sentences with modal particles in dispreferred/preferred relations by position are presented in Figure 3: Whether the modal particle appears in nucleus or satellite does not strongly influence the rating in dispreferred

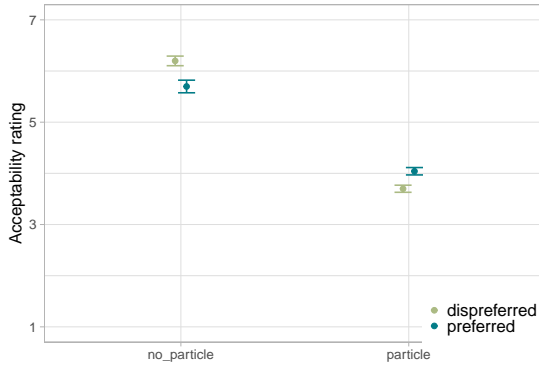


Figure 2: Mean rating for ‘preferred’ (EVIDENCE, JUSTIFY) and ‘dispreferred’ (CIRCUMSTANCE, CONDITION) relations, by particle presence. The rating scale ranges from 1 (= rejection) to 7 (=acceptance).

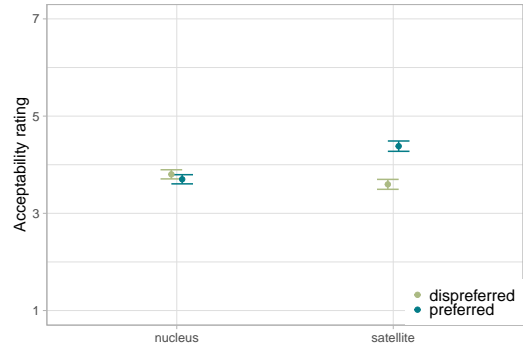


Figure 3: Mean rating for ‘preferred’ (EVIDENCE, JUSTIFY) and ‘dispreferred’ (CIRCUMSTANCE, CONDITION) relations, by position of the particle. Ratings for sentences without particle are excluded. The rating scale ranges from 1 (= rejection) to 7 (=acceptance).

relations, but does in preferred relations. Here, the presence of the modal particle in the satellite, as expected, leads to higher acceptability ratings.

6.6 Discussion

The rating study showed that the acceptability of sentences with modal particles varies depending on the discourse relation, thus confirming our H1: The acceptability of sentences with modal particles is worse for certain discourse relations. Still, in contrast to our expectation, modal particles in an EVIDENCE and JUSTIFY relation were not generally more accepted than particles in CIRCUMSTANCE or CONDITION, but showed this effect only if the particle was presented in the relation’s satellite. We can confirm our H2 as well: For the relations EVIDENCE and JUSTIFY, participants preferred the modal particle in the satellite of the relation over the particle being in the relation’s nucleus. The effect could not be shown for the relations CIRCUMSTANCE and CONDITION. This is in accord with expectations based on the particles’ semantics and the relation definitions: In EVIDENCE and JUSTIFY, the claim is presented in the nucleus, whereas the evidence/argument is presented in the satellite of the relation, matching the particles’ effect of marking information as known or salient.

Furthermore, we take our results to indicate an interaction of particles with the discourse relation and not simply co-occurrence with specific connectives. (15a) shows that in an implicit CONDITION relation, adding a modal particle to the satellite is still infelicitous.¹¹ On the other hand, adding *ja* or *doch* to the satellite of JUSTIFY as in (15b) is acceptable. While it will be up to future studies to verify these intuitions, we expect similar rating patterns for implicit relations.

11. A syntactic structure such as (15a), though without a modal particle, might be used to express a conditional sentence in German.

- (15) a. #Bricht Fiona *ja/doch* nicht bald auf, wird sie zu spät kommen.
 Intended: ‘If Fiona doesn’t leave soon, she will be late (as you knew before).’
 b. Jakob will *ja/doch* Medizin studieren, er muss einen guten Abschluss machen.
 ‘(As you know,) Jakob wants to study medicine, he has to get a good degree.’

The box plots showed that participants rated sentences in the EVIDENCE relation generally lower compared to all other relations. This is probably due to the order of the segments. To present all relations in the same manner, the satellite was always presented first. However, presenting the satellite (= the evidence) before the nucleus (= the claim) might be unusual for an EVIDENCE relation and lead to lower acceptability ratings (cf. Stede, 2016a). Even though employing a modal particle that marks the evidence as uncontroversial might be an effective rhetorical strategy (as indicated in section §3), the approach of presenting evidence first does seem to be less acceptable to participants in general.

The higher acceptability scores of *doch* in the CONDITION relation might be due to a limitation of this study: Context could not be provided for the sentences, as there is no context that allows for both modal particles *ja* and *doch* to be used in all stimuli. But without context, participants’ ratings of a sentence can differ depending on whether they read the particle in its stressed or unstressed version. This is exemplified in (16). While (16b) is a perfectly fine example, the sentence in (16a) is odd in German.

- (16) a. #Falls Fiona nicht bald aufbricht, wird sie *doch* zu spät kommen.
 Intended: ‘If Fiona doesn’t leave soon, she will be late (as you knew before).’
 b. Falls Fiona nicht bald aufbricht, wird sie *DOCH* zu spät kommen.
 ‘If Fiona doesn’t leave soon, she will be late (even though it seemed like she could make it in time).’

For future studies, this limitation of the prosodic stress influencing the particle’s rating outside of the experiment conditions can be avoided by reading the sentences aloud to participants. We acknowledge that further research is necessary to pinpoint the exact behavior of *doch* in different discourse contexts.

In this first experiment, we obtained results similar to previous findings on modal particles in discourse. Taking our results together with the experimental evidence presented in Döring and Repp (2019), an interaction of modal particles with discourse relations should be assumed. Our results go beyond their corpus findings insofar as they show that the dispreferred relations – that were never found with the tested modal particles in the corpus – are not completely unacceptable to the tested participants, just less acceptable than the preferred relations. After having identified that participants are sensitive to the interaction of particles and discourse, we proceed to studying this interaction further to find out whether modal particles are non-connective discourse signals.

7. Experiment 2: Modal Particles as Discourse Signals

The second experiment tested whether the presence of the modal particle *ja* in a sentence leads to participants disambiguating the sentence as causal, indicated by their choice of a causal connective. The preregistration can be accessed in the Data Availability section.

7.1 Hypotheses

Döring (2016) suggests that modal particles do not simply co-occur with discourse relations, but can be used to mark those relations. Taking the findings in Das and Taboada (2019) and studies on relation signals like Crible and Pickering (2020) as well as the particles' discourse-managing function into account, modal particles might be a non-connective discourse signal. To study this discourse-structuring characteristic, we tested the following hypothesis:

H3: As discourse signals, modal particles can help in disambiguating ambiguous relation contexts.

To test this hypothesis, we conducted a connective insertion study. Participants were presented with ambiguous contexts (causal-contrastive or causal-temporal) with a contrastive or temporal reading and an optional causal reading. The texts were presented with or without the modal particle *ja* and participants were asked to insert a connective out of a given set of alternatives. The relation types were chosen because they differ conceptually, but minimal pairs can be disambiguated without changing the sentence structure. Based on the results presented in Crible and Pickering (2020) – disambiguating a relation is easier if the second sentence contains a signal – we modified the stimuli to be presented with or without modal particle in the second clause.

We expect mainly contrastive respective temporal connectives to be chosen if the sentence is presented without *ja*, and a significant increase of the causal connective chosen if the sentence is presented with the particle. This expectation builds on the results presented in Döring (2016) as well as *ja*'s function of marking information as known or highly salient and not at-issue. This marking of an utterance makes *ja* a good choice for causal relations. Any statement containing *ja* might be more readily perceived as being uncontroversial or given – as long as the claim is not obviously false – and being a valid point in an argument. Thus the modal particle supports causality without being causal by itself: In the presence of *ja*, the causal connective *deshalb* ('that is why') should be chosen more frequently than without the particle.

7.2 Participants

For this experiment, 100 participants (mean age 35.3 years, SD 10.1 years) were recruited via Prolific and received compensation based on the minimum wage in Germany (1.60 EUR for 8 minutes). In this study, participants were instructed to choose one out of four connectives – the one they liked best in the given context – to connect the two sentence parts presented. Seven participants had to be excluded, two because German was not their first language, and five because they chose the syntactically mismatched distractor connective in critical items.

7.3 Materials and Procedure

We used a 2 x 2 design with the factors modal particle (*ja*, no particle) and ambiguous discourse context (causal-contrastive, causal-temporal). We designed eight experiment items for both discourse contexts that were presented in two versions: with or without the particle. The 16 items were split into two lists, ensuring every participant reads every sentence in exactly one version, and presented in random order. An example of a causal-contrastive context is given in (17), and a causal-temporal context is presented in (18). The low number of items is intended to avoid overexposure of participants to the modal particles. To balance the low number of observations per participant, we recruit a high number of individual participants.

- (17) *causal-contrastive*
 Raphael serviert normalerweise Zimtschnecken als Dessert, _____ backt Florian (Ø/*ja*) meistens Schokoladenkuchen.
 ‘Raphael usually serves cinnamon rolls as dessert, _____ Florian mostly bakes (Ø/*JA*) a chocolate cake.’
- (18) *causal-temporal*
 Die Züge warten auf die Behebung einer Signalstörung, _____ warten die Fahrgäste (Ø/*ja*) am Bahnhof.
 ‘The trains are waiting for a signalling error to be fixed, _____ the passengers are (Ø/*JA*) waiting at the train station.’

In addition, the experiment included two training items and eight items in unambiguously concessive contexts as fillers. None of the fillers contained any modal particles. When presented with each item, participants could choose between four different connectives: contrastive *dahingegen* (‘whereas’), causal *deshalb* (‘that is why’), concessive *obwohl* (‘although’), and temporal *währenddessen* (‘meanwhile’), and were asked to pick the one they prefer in this context. Except for *obwohl*, which serves as a distractor, the other three connectives could be used syntactically correctly in all critical items. A screenshot of the interface visible to participants is presented in Figure 4.

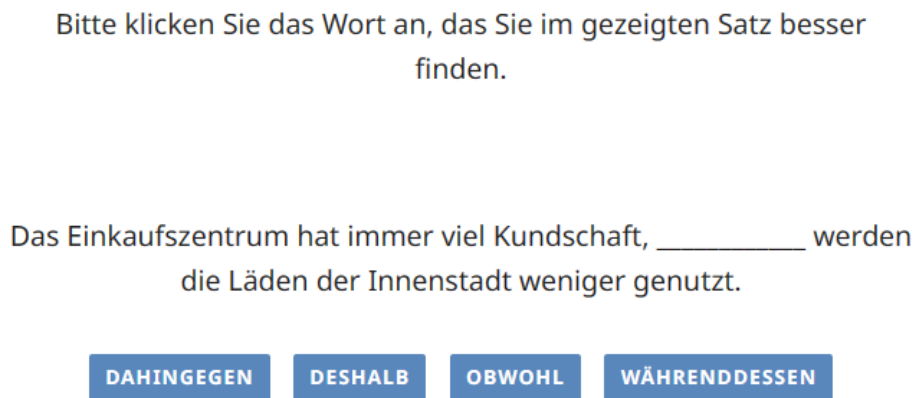


Figure 4: Interface visible for participants. English translation:
 “Please click on the word you like best in the presentend sentence.
 The shopping mall has many customers, _____ the stores in the city center are less frequented.”

The connectives were chosen based on the discourse relation they can signal and the sentence structure they fit in, not primarily based on their frequency. In the German connective lexicon (DiMLex; Scheffler and Stede, 2016), *dahingegen* is listed as the only contrastive connective that is not a signal for causal relations, too (Stede et al., 2019). Compared to the other connectives, it

appears much less frequently in German texts: The German *Referenz- und Zeitungskorpora* of the *Digitales Wörterbuch der Deutschen Sprache* (‘digital dictionary of the German language’),¹² see Geyken et al. (2017) list 243,969 matches for *deshalb* between 1990 and 2018 (305.62 per million token), but only 26 for *dahingegen* (0.03 per million token). To ensure participants will choose the contrastive connective even though it is used less frequently, we conducted a pilot study, showing that participants use *dahingegen* despite its rarity. In this pilot, we tested a sub-sample of the final stimuli with students. As the participants of the pilot did choose the rare connectives in the contexts we expected them to be chosen in, we concluded that the frequency of a connective does not keep participants from using it in the context of this study.

7.4 Analysis

We fit a generalized linear mixed model with sum coding for the independent variables ‘discourse relation’ and ‘particle’, using the *glmer* function in the R package *lme4* (Bates et al., 2015). We coded the connectives chosen as either ‘causal’ (= 1) or ‘not causal’ (= 0). Since the maximal random effect structure did not converge, we simplified the model to the maximal converging random effect structure, that is varying intercepts and varying slopes for particle presence, but not for relation. Details on the model used are presented together with the results.

7.5 Results

Overall, participants chose the causal connective more often if the sentence was presented with a modal particle. In the sentences without *ja*, the causal *deshalb* is chosen in 6.99% of all cases. In sentences presented with *ja*, this frequency increases to 17.34%. Figure 5 shows participants’ connective choices for both contexts. The results of the generalized linear mixed-effects model are presented in Table 3.

Fixed effects	Estimate	Std. Error	z-value	p-value
(Intercept)	-2.798	0.357	-7.841	< 0.001
∅ vs. <i>ja</i>	0.940	0.238	3.950	< 0.001
discourse relation	0.064	0.684	0.094	0.9254
particle : relation	0.743	0.429	1.731	0.0834

Table 3: Output of the generalized linear mixed-effects model shown below. Connective is coded as 1 = causal, 0 = not causal. Sum coding is used for both particle and relation.

```
glmer(connective ~ particle * relation + (1+particle||participant) +
(1+particle||item))
```

The model shows that the presence of *ja* has a significant effect on participants’ connective choice ($p < 0.001$). The main effect of discourse relation ($p = 0.9$) and the interaction between particle and relation ($p = 0.08$) are not statistically significant. In the causal-contrastive context, the frequency of causal *deshalb* increases from 7.79% to 15.59% in the presence of *ja*, in the causal-temporal context it increases from 6.18% to 19.08%. While the frequency increase is higher in the causal-temporal context, this difference is not found to be statistically significant by our model.

12. <https://www.dwds.de/d/korpora/public>

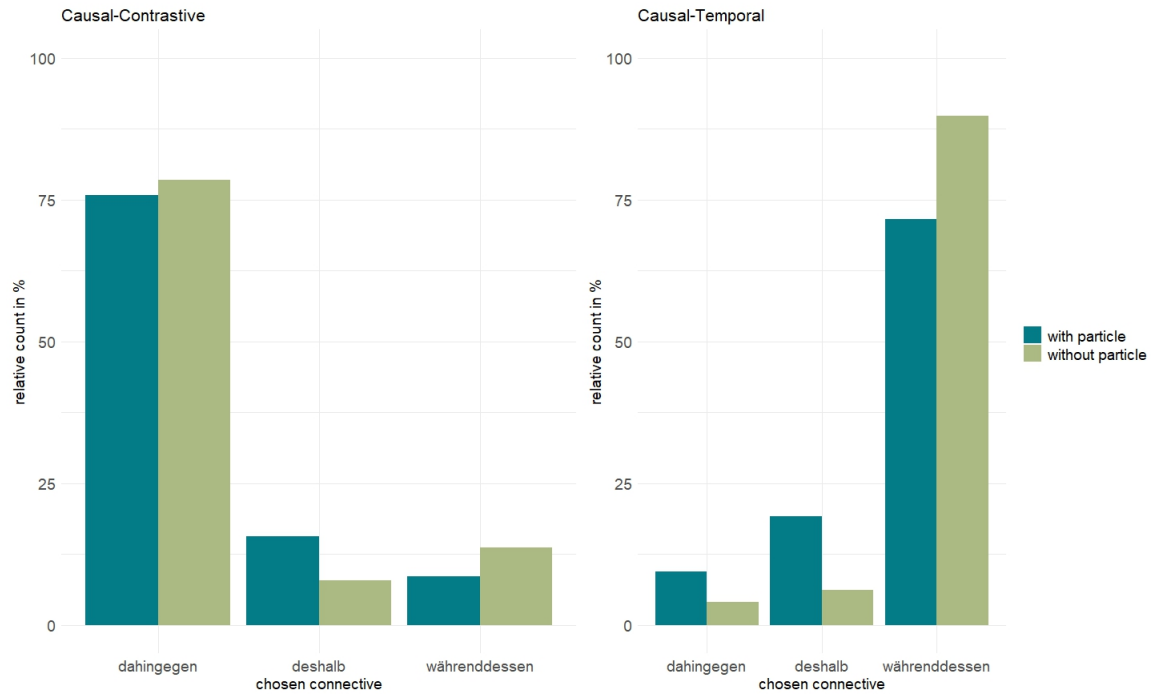


Figure 5: The bar plots show the absolute of the respective connective chosen (contrastive, causal, or temporal) based on whether the sentence was presented with or without the modal particle *ja* for both given contexts: causal-contrastive (left) and causal-temporal (right).

7.6 Discussion

Figure 5 illustrates that both discourse contexts are mainly perceived as intended, with a contrastive or temporal reading respectively as the default and optional causal reading. Both the graphs and the generalized linear mixed model show an increase of causal connectives chosen in the presence of *ja*. Given that neither the main effect of discourse relation nor the interaction of relation and particle are statistically significant, the effect is arguably the same for both discourse contexts tested: Adding the modal particle *ja* to an ambiguous discourse context leads to an increase of causal connectives chosen, independent of the type of relation displayed.

However, even if the frequency of the causal connective *deshalb* increases in the presence of *ja*, it is still not the connective used most in either context. If the modal particle were a discourse *connective*, the proportion of causal interpretations ought to be higher. Since discourse *signals* support certain interpretations rather than facilitating them on their own, they serve as main interpretation instructions only in the absence of connectives (Das, 2014). Often, discourse signals co-occur with connectives or other signals (Hoek et al., 2018; Das and Taboada, 2019). Hoek et al. (2018) distinguish three types of discourse signals: i) *division of labour*; signals that can fulfill the function of a connective in the absence of the same (e.g., lexical cues), ii) *agreement*; signals that fulfill the same function, but still co-occur together with a connective of similar type (e.g., modal verbs), and iii) *general collocation*; signals that always co-occur together with a connective (e.g., verbal tense).

How many signals occur (or co-occur) in a sentence depends on both the compatibility of signals present and the strength of the respective signals (cf. Asr and Demberg, 2012). In light of these observations, we consider *ja* to be type ii) or iii) marker that requires other signals to unambiguously indicate a certain discourse relation. Related to this point, it should be noted that while the experimental items did not contain additional connectives, the presence of additional signals compatible with various discourse relations – such as verbal tense – could not be avoided. But given that the presence of *ja* has a significant effect on participants’ connective choice, we consider *ja* to be a causal discourse signal.

This raises the question of the exact relation between *ja* and causality. As stated at the beginning of this section, we do not take the modal particle to be causal by itself. A causal connective will always lead to the relation between the two discourse segments it connects being interpreted as causal. Modal particles, like all phrase-internal signals, do not connect segments but occur in one of them. They might indicate how the relation holding between the segment carrying the signal and the other segment is to be interpreted, but the same signal can indicate different discourse relations, as discussed by Hoek et al. (2018): One of the non-connective discourse signals observed by Das and Taboada (2018), the numeral *five*, can indicate different discourse relations, depending on the context. If five elements in a list follow the numeral, an ELABORATION relation is signaled. But if the numeral is contrasted with another numeral, CONTRAST might be the correct relation to infer from the same signal. The same holds for modal particles as signals. If a causal interpretation of a relation is possible, then the presence of *ja* indicates that this causal interpretation might be valid. This effect is likely to arise from the particle’s function of backgrounding information and marking it as potentially previously known: A statement marked by *ja* as salient and difficult to refuse can facilitate drawing a conclusion in an argument – and thus a causal interpretation. At this point, it cannot be decided whether *ja* directly signals a causal relation or whether it merely marks information as salient, thus facilitating a causal interpretation in a second step. Independent of which explanation holds, the result of adding the particle *ja* to a sentence would be an increased probability of the discourse relation being interpreted as causal. Thus we take our results as the first experimental evidence of *ja* as a discourse signal, though further studies are needed to pinpoint the exact contribution of modal particles to the interpretation of discourse, and their interaction with connectives or additional signals such as verbal tense.

A last observation to be discussed is the frequent choice of both the contrastive *dahingegen* and the temporal *währenddessen*, even if sentences were presented with *ja*. In the case of contrastive *dahingegen*, this does not pose a problem, as both discourse signals and connectives can, in principle, signal any compatible relation.¹³ But the fact that ~70% of the participants chose the temporal *währenddessen* in the presence of *ja* might appear unexpected. However, the above-mentioned restrictions of modal particles in temporal clauses only concern subordinate clauses (Coniglio, 2012), and the sentences in our experiment were temporal main clauses. Combining the results of this experiment with the acceptability rating of *ja* in the temporal CIRCUMSTANCE relation (median = 3; which means this interpretation is disliked, but not completely unacceptable), this indicates that the modal particle is not generally infelicitous in the temporal context, but rather dispreferred.

13. Studies documenting the co-occurrence of signals have been discussed above. Lexicons of discourse connectives list more than one possible discourse relation for most of the included connectives (Stede et al., 2019), and studies of individual signals in various languages also conclude that marking multiple relations is possible (Schwenter, 2000; Miltsakaki et al., 2005; Pitler et al., 2008; Mortier and Degand, 2009; Meyer et al., 2011; Stede, 2014; Costăchescu, 2017).

8. Conclusion

The present study outlines the role of German modal particles as interpretation instructions for discourse relations. We approach this question by first experimentally confirming previous theoretical and corpus-based predictions about the interaction of modal particles and discourse structure in the form of discourse relations, and directly testing a modal particle’s effect on relation disambiguation in a second step. For the first step, we conducted an acceptability study, asking participants to rate sentences expressing four different discourse relations. In those relations, we expected *ja* to be dispreferred in CIRCUMSTANCE and CONDITION, but not in EVIDENCE and JUSTIFY. Results show that participants are indeed sensitive to the predicted interaction between particles and discourse relations, as the ratings for the two causal relations were higher than for the other two relations if the sentence was presented with a modal particle. Furthermore, we tested whether the position of the particle, nucleus (e.g., claim) or satellite (e.g., evidence or argument), of a relation influences the ratings. For the causal relations, we found higher ratings if the particles were presented in a relation’s satellite, but we could not show this effect for the dispreferred relations.

In the second step, we conducted a forced-choice experiment, where participants were tasked with inserting a connective into an ambiguous context to indicate the perceived discourse relation between the presented segments. We found that the presence of the modal particle *ja* led to a significant increase in the use of the causal connective *deshalb*, indicating that *ja* might be interpreted as a causal discourse signal.

With these results, our study presents first experimental evidence for modal particles serving as interpretation instructions in discourse. The experimental approach chosen here offers insights into the perception of modal particles, complementing previous corpus results on the distribution of modal particles over various discourse contexts. This study thus contributes to a series of studies on non-connective discourse signals as well as a discussion of the status of modal particles. But since we tested only one German modal particle, further research investigating modal particles’ effects on discourse is needed to pinpoint their exact contribution in discourse. Given that we find the particle’s disambiguation effect to be rooted in its semantic properties, we expect to find the same effect for all particles with similar properties. Testing not only more modal particles in German but also in other languages with particles that are used for similar functions – like Dutch – will offer more insights on this. In addition, studying the contribution of further modal particles compatible with different contexts will help answer the question of how modal particles guide the interpretation of discourse relations.

Data Availability

Both experiments were created in *magpie*¹⁴ and preregistered. The raw data, stimuli, JavaScript code, preregistration, and analysis for experiment 1 can be accessed via the Open Science Framework.¹⁵ All corresponding data for experiment 2 can be found in the corresponding repository.¹⁶

Declaration of Interest

The authors report there are no competing interests to declare.

14. <https://magpie-ea.github.io/magpie-site/>

15. https://osf.io/xhdfj/?view_only=63ad745bdaf44583aaf527876764b50d

16. https://osf.io/9gp6c/?view_only=088a957ed6054db7b614bb3f47c1f51a

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project ID 317633480 – SFB 1287.

We thank Merel Scholman for her insights and help with the statistical analysis, and our reviewers and editor for their detailed comments and suggestions. We appreciate their work and think our paper profited a lot from it.

Appendix

The presented definitions are based on the relation description defined in Stede (2016a) and Döring (2016). The examples given for each relation are taken from the Potsdam Commentary Corpus (Stede, 2016b), as translated by Stede et al. (2017).

CIRCUMSTANCE

The nucleus (N) is interpreted in light of the satellite (S), where S presents an event or state of affairs as context in time and space.

Example: [When Veag came under pressure because of the deregulation of the electricity market,]_S [they compensated for this by squeezing their suppliers.]_N (maz-5297, PCC)

CONDITION

The realization of N depends on the realization of S, where S presents a future event or hypothetical situation.

Example: [If the sanitary facilities are still not available in the coming season,]_S [Radewege is in danger of losing the competition for attracting boats to the campground.]_N (maz-6488, PCC)

EVIDENCE

N presents a claim. S presents information a reader is likely to accept. Accepting S raises a reader's willingness to accept N.

Example: *Debate about having two subjects 'Religion' and 'LER' at school* [And now even our state government seems determined to remove this apparent equality between the two subjects.]_N [Stolpe, Reiche and others do say Yes to the possible compromise offered by the Karlsruhe court, but they also decree: There cannot be any voluntary subject area LER/Religion.]_S (maz-6159, PCC)

JUSTIFY

Both N and S present subjective statements. S presents a reason or justification for the statement presented in N.

Example: [With each new day of air raids, the military operations of the U.S. lose more and more credibility.]_N [By means of comprehensive area-wide destruction you can't hit the Taliban, nor can you eliminate bin Laden.]_S (maz-5701, PCC)

References

- Werner Abraham. Discourse marker = discourse particle = thetical = modal particle? A futile comparison. In Josef Bayer and Volker Struckmeier, editors, *Discourse Particles*, pages 241–280. De Gruyter, 2017. ISBN 978-3-11-049715-1. doi: 10.1515/9783110497151-010.
- Werner Abraham. *Modality in Syntax, Semantics and Pragmatics*, volume 165 of *Cambridge Studies in Linguistics*. Cambridge University Press, 2020. doi: 10.1017/9781139108676.
- Fatemeh Torabi Asr and Vera Demberg. Measuring the Strength of Linguistic Cues for Discourse Relations. *Proceedings of the Workshop on Advances in Discourse Analysis and its Computational Aspects (ADACA), COLING, Mumbai*, pages 33–41, 2012.
- Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi: 10.18637/jss.v067.i01.
- Douglas Biber. Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5(2):97–116, 2006. ISSN 14751585. doi: 10.1016/j.jeap.2006.05.001.
- Hardarik Blühdorn. Diskursmarker: Pragmatische Funktion und syntaktischer Status. In Hardarik Blühdorn, Arnulf Deppermann, Henrike Helmer, and Thomas Spranz-Fogasy, editors, *Diskursmarker im Deutschen. Reflexionen und Analysen*, pages 311–336. Verlag für Gesprächsforschung, Göttingen, 2017.
- Hardarik Blühdorn. Modalpartikeln und Akzent im Deutschen. *Linguistische Berichte*, 259:275–317, 2019.
- Anneloes R. Canestrelli, Willem M. Mak, and Ted J. M. Sanders. Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements. *Language and Cognitive Processes*, 28(9):1394–1413, 2013. ISSN 0169-0965, 1464-0732. doi: 10.1080/01690965.2012.685885.
- R. H. B. Christensen. ordinal—regression models for ordinal data, 2022. R package version 2022.11-16. <https://CRAN.R-project.org/package=ordinal>.
- Marco Coniglio. Die Syntax der deutschen Modalpartikeln: Ihre Distribution und Lizenzierung in Haupt- und Nebensätzen. In *Die Syntax der deutschen Modalpartikeln*, volume 73 of *Studia grammatica*. Akademie Verlag, 2012. ISBN 978-3-05-005357-8. doi: 10.1524/9783050053578.
- Adriana Costăchescu. Discourse markers and discourse relations: The French DM quoui. In Chiara Fedriani and Andrea Sansó, editors, *Pragmatic Markers, Discourse Markers and Modal Particles. New perspectives*, volume 186 of *Studies in Language Companion Series*, pages 151–167. John Benjamins Publishing Company, Amsterdam, 2017. doi: 10.1075/slcs.186.06cos.
- Ludivine Crible. Weak and Strong Discourse Markers in Speech, Chat, and Writing: Do Signals Compensate for Ambiguity in Explicit Relations? *Discourse Processes*, 57(9):793–807, 2020. ISSN 0163-853X, 1532-6950. doi: 10.1080/0163853X.2020.1786778.
- Ludivine Crible and Vera Demberg. The role of non-connective discourse cues and their interaction with connectives. *Pragmatics & Cognition*, 27(2):313–338, 2020. ISSN 0929-0907, 1569-9943. doi: 10.1075/pc.20003.cri.

- Ludvine Crible and Pickering. Compensating for processing difficulty in discourse: Effect of parallelism in contrastive relations. *Discourse Processes*, 57(10):862–879, 2020.
- Johannes Dahl. *Die Abtönungspartikeln im Deutschen. Ausdrucksmittel für Sprechereinstellungen - mit einem kontrastiven Teil deutsch-serbokroatisch*, volume 7 of *Deutsch im Kontrast*. Julius Groos Verlag, Heidelberg, 1988.
- Debopam Das. *Signalling of Coherence Relations in Discourse*. Dissertation, Simon Fraser University, Canada, 2014.
- Debopam Das and Maite Taboada. RST Signalling Corpus: a corpus of signals of coherence relations. *Language Resources and Evaluation*, 52(1):149–184, 2018. ISSN 1574-020X, 1574-0218. doi: 10.1007/s10579-017-9383-x.
- Debopam Das and Maite Taboada. Multiple Signals of Coherence Relations. *Discours*, 24, 2019. ISSN 1963-1723. doi: 10.4000/discours.10032.
- Debopam Das, Maite Taboada, and Paul McFetridge. RST Signalling Corpus, 2015.
- Liesbeth Degand and Ted Sanders. The impact of relational markers on expository text comprehension in L1 and L2. *Reading and Writing*, 15(7-8):739–757, 2002.
- Liesbeth Degand, Bert Cornillie, and Paola Pietrandrea, editors. *Discourse markers and modal particles: categorization and description*. Number 234 in Pragmatics & beyond new series. John Benjamins Publishing Company, Amsterdam, Philadelphia, 2013. ISBN 978-90-272-5639-3.
- Gabriele Diewald. Abtönungspartikel. In *Handbuch der deutschen Wortarten*, pages 117–141. De Gruyter, Berlin, New York, 2009.
- Monika Doherty. *Epistemische Bedeutung*, volume 32 of *Studia grammatica*. Akademie Verlag, Berlin, Boston, 1985. doi: doi:10.1515/9783050067452.
- Sophia Döring. *Modal Particles, Discourse Structure and Common Ground Management*. Dissertation, Humboldt-Universität zu Berlin, 2016.
- Sophia Döring and Sophie Repp. The modal particles *ja* and *doch* and their interaction with discourse structure: Corpus and experimental evidence. In Sam Featherston, Robin Hörnig, Sophie von Wietersheim, and Susanne Winkler, editors, *Experiments in Focus*, pages 17–56. De Gruyter, 2019. ISBN 978-3-11-062309-3. doi: 10.1515/9783110623093-002.
- Laura Dörre, Anna Czypionka, Andreas Trotzke, and Josef Bayer. The processing of German modal particles and their counterparts. *Linguistische Berichte*, 255:313–346, 2018.
- Markus Egg and Malte Zimmermann. Accented Discourse Particles: The Case of ‘*doch*’. *Proceedings of Sinn und Bedeutung*, 16:225–238, 2012.
- Chiara Fedriani and Andrea Sansò, editors. *Pragmatic Markers, Discourse Markers and Modal Particles*, volume 186 of *Studies in Language Companion Series (SLCS)*. John Benjamins Publishing Company, Amsterdam, Philadelphia, 2017.
- Bruce Fraser. What are discourse markers? *Journal of Pragmatics*, 31(7):931–952, 1999.

- Alexander Geyken, Adrien Barbaresi, Jörg Didakowski, Bryan Jurish, Frank Wiegand, and Lothar Lemnitzer. Die korpusplattform des ‘digitalen wörterbuchs der deutschen sprache’ (dwds). *Zeitschrift für germanistische Linguistik*, 45(2):327–344, 2017. doi: doi:10.1515/zgl-2017-0017.
- Bethany Gray and Douglas Biber. Stance markers. In Karin Aijmer and Christoph Rühlemann, editors, *Corpus Pragmatics. A Handbook*, pages 219–248. Cambridge University Press, Cambridge, 2015.
- Cristina Grisot and Joanna Blochowiak. Temporal connectives and verbal tenses as processing instructions. *Pragmatics & Cognition*, 24(3):404–440, 2017.
- Daniel Gutzmann. Betonte Modalpartikeln und Verumfokus. In Theo Harden and Elke Hentschel, editors, *40 Jahre Partikelforschung*, volume Band 55 of *Stauffenburg Linguistik*, pages 119–138. Narr, Tübingen, 2010.
- Daniel Gutzmann. *Use-Conditional Meaning*. Number 6 in Studies in Multidimensional Semantics. Oxford University Press, Oxford, New York, 2015.
- Daniel Gutzmann. Modal particles \neq modal particles (= modal particles). In Josef Bayer and Volker Struckmeier, editors, *Discourse Particles*, pages 144–172. De Gruyter, 2016. ISBN 978-3-11-049715-1. doi: 10.1515/9783110497151-007.
- Dietrich Hartmann. Syntaktische Funktionen der Partikeln eben, eigentlich, einfach, nämlich, ruhig, vielleicht und wohl. Zur Grundlegung einer diachronischen Untersuchung von Satzpartikeln im Deutschen. In Harald Weydt, editor, *Die Partikeln der deutschen Sprache*, pages 121–138. De Gruyter, Berlin, New York, 1979.
- Dietrich Hartmann. Art. particles. In Jakob L. Mey, editor, *Concise Encyclopedia of Pragmatics*, pages 657–663. Elsevier, Oxford, 1998.
- Gerhard Helbig and Joachim Buscha. *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht*. Langenscheidt, Berlin, München, Wien, Zürich, New York, 2001.
- Elke Hentschel. Funktion und Geschichte deutscher Partikeln: Ja, doch, halt und eben. In *Funktion und Geschichte deutscher Partikeln*, volume 63 of *Reihe Germanistische Linguistik*. Max Niemeyer Verlag, 1986. doi: 10.1515/9783111371221.
- Jet Hoek, Sandrine Zufferey, Jacqueline Evers-Vermeul, and Ted J. M. Sanders. The linguistic marking of coherence relations: Interactions between connectives and segment-internal elements. *Pragmatics & Cognition*, 25(2):276–309, 2018. doi: 10.1075/pc.18016.hoe.
- Lotte Hogeweg, Stefanie Ramachers, and Verena Wottrich. *Doch*, *toch* and *wel* on the table. *Linguistics in the Netherlands*, 28:50–60, 2011. doi: 10.1075/avt.28.05hog.
- Judith Kamalski. *Coherence marking, comprehension and persuasion: on the processing and representation of discourse*. Dissertation, Universiteit Utrecht, Utrecht, 2007.
- Elena Karagjosova. *The Meaning and Function of German Modal Particles*. Dissertation, Universität des Saarlandes, Saarbrücken, 2004.

- Suzanne Kleijn, Henk L.W. Pander Maat, and Ted J.M. Sanders. Comprehension Effects of Connectives Across Texts, Readers, and Coherence Relations. *Discourse Processes*, 56(5-6):447–464, 2019. ISSN 0163-853X, 1532-6950. doi: 10.1080/0163853X.2019.1605257.
- Angelika Kratzer. Beyond ouch and oops: How descriptive and expressive meaning interact. In *Cornell conference on theories of context dependency*, volume 26. Cornell University Ithaca, NY, 1999.
- Angelika Kratzer. Interpreting focus: Presupposed or expressive meanings? A comment on Geurts and van der Sandt. *Theoretical Linguistics*, 30(1):123–136, 2004. doi: doi:10.1515/thli.2004.002.
- Gina R. Kuperberg, Martin Paczynski, and Tali Ditman. Establishing Causal Coherence across Sentences: An ERP Study. *Journal of Cognitive Neuroscience*, 23(5):1230–1246, 2011. doi: 10.1162/jocn.2010.21452.
- Min-Jae Kwon. *Modalpartikeln und Satzmodus. Untersuchungen zur Syntax, Semantik und Pragmatik der deutschen Modalpartikeln*. Dissertation, Ludwig-Maximilians-Universität, München, 2005.
- Judith Köhne-Fuetterer, Heiner Drenhaus, Francesca Delogu, and Vera Demberg. The online processing of causal and concessive discourse connectives. *Linguistics*, 59(2):417–448, 2021. ISSN 1613-396X, 0024-3949. doi: 10.1515/ling-2021-0011.
- Ekkehard König and Susanne Reuquard. A relevance-theoretic approach to the analysis of modal particles in German. *Multilingua. Journal of Cross-Cultural and Interlanguage Communication*, 10(1-2):63–78, 1991. doi: https://doi.org/10.1515/mult.1991.10.1-2.0.
- William C. Mann and Sandra A. Thompson. Rhetorical Structure Theory: Toward a functional theory of text organization. *Text - Interdisciplinary Journal for the Study of Discourse*, 8(3): 243–281, 1988. ISSN 0165-4888, 1613-4117. doi: 10.1515/text.1.1988.8.3.243.
- Daniel Marcu. *The Theory and Practice of Discourse Parsing and Summarization*. Cambridge, London, MIT Press, 2000.
- Thomas Meyer, Andrei Popescu-Belis, Sandrine Zufferey, and Bruno Cartoni. Multilingual annotation and disambiguation of discourse connectives for machine translation. In Joyce Y. Chai, Johanna D. Moore, Rebecca J. Passonneau, and David R. Traum, editors, *Proceedings of the SIG-DIAL 2011 Conference*, pages 194–203, Portland, Oregon, 2011. Association for Computational Linguistics. URL <https://aclanthology.org/W11-2022>.
- Keith K. Millis and Marcel A. Just. The Influence of Connectives on Sentence Comprehension. *Journal of Memory and Language*, 33(1):128–147, 1994.
- Eleni Miltsakaki, Nikhil Dinesh, Rashmi Prasad, Aravind Joshi, and Bonnie Webber. Experiments on Sense Annotations and Sense Disambiguation of Discourse Connectives. In *Proceedings of the 4th Workshop on Treebanks and Linguistic Theories*, Barcelona, Spain, 2005.
- Liesbeth Mortier and Liesbeth Degand. Adversative discourse markers in contrast: The need for a combined corpus approach. *International Journal of Corpus Linguistics*, 14(3):338–366, 2009. ISSN 1384-6655. doi: 10.1075/ijcl.14.3.03mor. Publisher: John Benjamins Publishing Company.

- Sonja Müller. *Modalpartikeln*. Kurze Einführungen in die germanistische Linguistik 17. Winter, Heidelberg, 2014. ISBN 978-3-8253-6365-9.
- Sonja Müller. Alte und neue Fragen der Modalpartikel-Forschung. *Linguistische Berichte*, 252: 383–441, 2017.
- Emily Pitler, Mridhula Raghupathy, Hena Mehta, Ani Nenkova, Alan Lee, and Aravind Joshi. Easily Identifiable Discourse Relations. In Donia Scott and Hans Uszkoreit, editors, *Coling 2008: Companion volume: Posters*, pages 87–90, Manchester, UK, 2008. Coling 2008 Organizing Committee. URL <https://aclanthology.org/C08-2022>.
- Christopher Potts. *The Logic of Conventional Implicatures*. Oxford University Press, Oxford, 2005.
- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. The Penn Discourse TreeBank 2.0. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, and Daniel Tapias, editors, *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC)*, Marrakech, Morocco, 2008. European Language Resources Association (ELRA).
- Marie-Paule Péry-Woodley, Lydia-Mai Ho-Dac, Josette Rebeyrolle, Ludovic Tanguy, and Cécile Fabre. A corpus-driven approach to discourse organisation: from cues to complex markers. *Dialogue & Discourse*, 8(1):66–105, 2017. doi: 10.5087/dad.2017.103.
- Laura Reimer and Andreas Trotzke. The Processing of Secondary Meaning: An Experimental Comparison of Focus and Modal Particles in Wh-questions. In Daniel Gutzmann, editor, *Secondary Content*, volume 37 of *Current Research in the Semantics / Pragmatics Interface*, pages 143–167. Brill, Leiden, Boston, 2019.
- Tania Rojas-Esponda. A discourse model for überhaupt. *Semantics and Pragmatics*, 7:1–45, 2014a. ISSN 1937-8912. doi: 10.3765/sp.7.1.
- Tania Rojas-Esponda. A QUD account of German doch. *Proceedings of Sinn und Bedeutung*, 18: 359–376, 2014b.
- Ted Sanders. Coherence, causality and cognitive complexity in discourse. In M. Aurnague and M. Bras, editors, *Proceedings of the First International Symposium on the Exploration and Modelling of Meaning*, pages 31–46, Université de Toulouse-Le Mirail, 2005.
- Ted Sanders and Leo Noordman. The Role of Coherence Relations and Their Linguistic Markers in Text Processing. *Discourse Processes*, 29(1):37–60, 2000. ISSN 0163-853X, 1532-6950. doi: 10.1207/S15326950dp2901_3.
- Ted Sanders and Wilbert Spooren. Discourse and text structure. In Dirk Geeraerts and Hubert Cuyckens, editors, *The Oxford handbook of cognitive linguistics*, pages 916–941. Oxford University Press, Oxford, New York, 2007.
- Ted Sanders, Wilbert Spooren, and Leo Noordman. Toward a taxonomy of coherence relations. *Discourse Processes*, 15(1):1–35, 1992.

- Ted Sanders, Vera Demberg, Jet Hoek, Merel Scholman, Fatemeh Torabi Asr, Sandrine Zufferey, and Jacqueline Evers-Vermeul. Unifying dimensions in coherence relations: How various annotation frameworks are related. *Corpus Linguistics and Linguistic Theory*, 17(1):1–71, 2021. doi: 10.1515/cllt-2016-0078.
- Tatjana Scheffler and Manfred Stede. Adding Semantic Relations to a Large-Coverage Connective Lexicon of German. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC)*, pages 1008–1013, Portorož, Slovenia, 2016.
- Merel Scholman, Vera Demberg, and Ted Sanders. Individual differences in expecting coherence relations: Exploring the variability in sensitivity to contextual signals in discourse. *Discourse Processes*, 57(10):844–861, 2020. ISSN 0163-853X, 1532-6950. doi: 10.1080/0163853X.2020.1813492.
- Scott Schwenter. Viewpoints and polysemy: Linking adversative and causal meanings of discourse markers. In Elizabeth Couper-Kuhlen and Bernd Kortmann, editors, *Cause - Condition - Concession - Contrast. Cognitive and Discourse Perspectives*, volume 33 of *Topics in English Linguistics*, pages 257–282. De Gruyter Mouton, Berlin ; New York, 2000.
- Manfred Stede. a prerequisite for discourse parsing: Resolving connective ambiguity. In Helmut Gruber and Gisela Redeker, editors, *The Pragmatics of Discourse Coherence: Theories and applications*, Pragmatics & Beyond New Series, pages 121–141. John Benjamins Publishing Company, 2014. doi: 10.1075/pbns.254.05ste.
- Manfred Stede, editor. *Handbuch Textannotation: Potsdamer Kommentarkorpus 2.0*. Number 8 in Potsdam cognitive science series. Universitätsverlag Potsdam, Potsdam, 2016a. ISBN 978-3-86956-343-5.
- Manfred Stede. Das Potsdamer Kommentarkorpus. In Hartmut Lenk, editor, *Persuasionsstile in Europa II. Kommentartexte in den Medienlandschaften europäischer Länder*, number 229-231 in Germanistische Linguistik, pages 177–202. Olms, Hildesheim, 2016b.
- Manfred Stede, Maite Taboada, and Debopam Das. Annotation Guidelines for Rhetorical Structure, 2017. Unpublished Manuscript.
- Manfred Stede, Tatjana Scheffler, and Amália Mendes. Connective-Lex: A Web-Based Multilingual Lexical Resource for Connectives. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics*, Discours (24), 2019. doi: 10.4000/discours.10098. Number: 24 Publisher: Presses universitaires de Caen.
- Maite Taboada. Discourse markers as signals (or not) of rhetorical relations. *Journal of Pragmatics*, 38(4):567–592, 2006. doi: 10.1016/j.pragma.2005.09.010.
- Maria Thurmair. *Modalpartikeln und ihre Kombinationen*. Number 223 in Linguistische Arbeiten. Niemeyer, Tübingen, 1989. ISBN 978-3-484-30223-5. doi: 10.1515/9783111354569.
- Maria Thurmair. Satztyp und Modalpartikeln. In Jörg Meibauer, Markus Steinbach, and Hans Altmann, editors, *Satztypen des Deutschen*, pages 627–651. De Gruyter, Berlin ; New York, 2013.

- Geertje Van Bergen and Hans Rutger Bosker. Linguistic expectation management in online discourse processing: An investigation of Dutch *inderdaad* 'indeed' and *eigenlijk* 'actually'. *Journal of Memory and Language*, 103:191–209, 2018. doi: 10.1016/j.jml.2018.08.004.
- Yvonne Viesel. Discourse particles “embedded”: German *ja* in adjectival phrases. In Josef Bayer, editor, *Discourse Particles. Formal Approaches to their Syntax and Semantics*, Linguistische Arbeiten, pages 173–202. De Gruyter, 2017. doi: 10.1515/9783110497151-008.
- Yipu Wei, Willem M. Mak, Jacqueline Evers-Vermeul, and Ted J. M. Sanders. Causal connectives as indicators of source information: Evidence from the visual world paradigm. *Acta Psychologica*, 198:Article 102866, 2019. doi: 10.1016/j.actpsy.2019.102866.
- Yipu Wei, Jacqueline Evers-Vermeul, and Ted J.M. Sanders. The Use of Perspective Markers and Connectives in Expressing Subjectivity: Evidence from Collocational Analyses. *Dialogue & Discourse*, 11(1):62–88, 2020. ISSN 2152-9620. doi: 10.5087/dad.2020.103.
- Malte Zimmermann. Discourse Particles. In P. Portner, C. Maienborn, and K. von Heusinger, editors, *Semantics*, volume 2 of *Handbücher zur Sprach- und Kommunikationswissenschaft HSK*, pages 2011–2038. Mouton de Gruyter, Berlin, 2011.