Glossa
a journal of general linguistics

Question-sensitive discourse particles at the interfaces of syntax, semantics and pragmatics – an experimental approach

RESEARCH

ANNA CZYPIONKA
MARIBEL ROMERO
JOSEF BAYER
*Author affiliations can be found in the back matter of this article

]u[ubiquity press

ABSTRACT

Question-sensitive discourse particles (QDiPs) like German *denn* introduce non-at-issue meaning that intuitively reshapes the Force of the interrogative clause. QDiPs have interesting licensing conditions: While Q-operators do not license QDiPs across clause boundaries, embedded QDiPs can be licensed if the *wh*-element was moved from the clause containing the QDiP. We present the results of two rating and two self-paced reading studies, with the following main results. First, outright licensing violations cause strong effects. Second, for embedded DiPs, a mild increase is found in processing cost for successful long-distance licensing. Third, effects for violations of syntactic locality are surprisingly weak in offline and online measures. We discuss two potential ways to account for the last findings. On the one hand, we consider an explanation in terms of processing errors. On the other hand, we offer a characterization of pragmatic aspects of QDiP licensing via focus association that may contribute to non-syntactic/non-semantic QDiP licensing.

CORRESPONDING AUTHOR:
Anna Czypionka
University of Konstanz, Fach 185, 78457 Konstanz, DE
anna.czypionka@uni-konstanz.de

KEYWORDS:
discourse particles; questions; *wh*-movement; long-distance licensing; sentence comprehension; German

TO CITE THIS ARTICLE:
Czypionka, Anna, Maribel Romero and Josef Bayer. 2021. Question-sensitive discourse particles at the interfaces of syntax, semantics and pragmatics – an experimental approach. *Glossa: a journal of general linguistics* 6(1): 24. 1–34. DOI: *https://doi.org/10.5334/gjgl.1203*

1 INTRODUCTION

Discourse particles (in the following, DiPs) in German are words that intuitively reshape the illocutionary force of an utterance. These DiPs are words like *denn*, *schon* ('already') or *bloß* ('only'). DiPs are pervasive in daily life communication. All DiPs have a non-DiP reading, and their reading crucially depends on context. Another interesting property is that DiPs are sensitive to clause type, with certain types of DiPs being dependent on occurring in declaratives or interrogatives (question-sensitive discourse particles, or QDiPs). QDiPs are subject to a number of different licensing constraints, thus forming a long-distance dependency at the interfaces of syntax, semantics and pragmatics. Although there is a long tradition in studying QDiPs in syntax and semantics, very little is known about their role in processing.

In the following, we will give an outline of the syntactic and semantic licensing constraints for QDiPs. We will then proceed with a short overview of related phenomena in language comprehension research, followed by a more detailed restatement of our research questions. We will present the outcome of four experimental studies on offline and online processing of QDiP licensing. The general discussion will link the experimental results to the theoretical accounts discussed in this Introduction.

1.1 CONTRIBUTION OF DIPS TO INTERROGATIVE MEANING

To get an idea of the role of DiPs in questions, let us consider the *wh*-question in (1) and the meaning that results from combining it with a DiP, e.g. with *denn*:

(1) Wo wird Max wohnen?
where will Max stay
'Where will Max stay?'

By its form, (1) is a simple information seeking question by which the speaker expresses the desire to identify a place x of which it will or may be true that Max stays at x. Provided that we know who Max refers to, (1) could be used as an out-of-the blue question. Things change when we add the question-sensitive DiP (QDiP) *denn* as in (2):

(2) Wo wird Max denn wohnen?
where will Max QDiP stay
'Where will Max DENN stay?'

The addition of *denn* in (2) forces a link between the current question and the discourse situation (König 1977; Thurmair 1991, i.a.). For example, König (1977) proposes that, for a *denn*-question to be felicitous, the *denn*-question has to be motivated by the previous interactional context. If no preceding context is provided, or if the preceding context is completely irrelevant to the question and does not lead to a reason to ask it, the *denn*-question is pragmatically infelicitous.[1]

As we shall see shortly, *denn* is Q(uestion)-sensitive.[2] Assertive sentences with *denn* are generally ill-formed. As we will see, in addition to *denn*'s context dependency, *denn* and other QDiPs[3] show interesting sentence-internal licensing constraints, which will be the central concern of our studies.[4]

1 We will not be concerned in this article with the exact lexical meaning of *denn*. Nevertheless, since *denn*'s lexically imposed conditions also have an impact on the acceptability of the sentence in context, we will briefly come back to them in section 7. See also footnote 7.

2 The rough empirical generalization is that, as a DiP, *denn* can occur in *wh*-questions and yes/no-questions and in no other environment. Interesting exceptions exist, in two directions. On the one hand, DiP *denn* can occur within *if*-clauses (Csipak & Zobel 2016). See Rawlins (2008), Onea & Steinbach (2012) and Romero (2015), a.o., for syntactic and semantic/pragmatic similarities between interrogative and conditional clauses and see Theiler (2020) for a unified lexical entry of *denn* in these two environments. On the other hand, DiP *denn* is not compatible with all kinds of *wh*-questions and yes/no-questions. For example, echo questions with *wh*-in-situ like (i) do not tolerate DiP *denn*:

(i) *Max wird wo denn wohnen?
Max will where QDiP stay
'Max will DENN stay where?'

3 There are other QDiPs such as *bloß, nur, schon, wohl* which we cannot discuss here for reasons of space and methodological constraints (see Bayer 2018; Dörre et al. 2018).

4 There is a certain similarity with the licensing of a negative polarity item (NPI) by a c-commanding carrier of negation or similar operator. In the absence of such an operator, the NPI gives rise to ungrammaticality.

DOI: 10.5334/gjgl.1203

In the following two subsections, we briefly present two existing accounts of the sentence-internal licensing constraints of QDiPs: a syntactic account (Bayer et al. 2016) and a semantic account (Romero 2017).

1.2 SYNTACTIC LICENSING OF QDIPS

1.2.1 Syntactic integration

In German(ic) root-questions, *wh*-questions as well as polar questions, the seat of interrogative Force must be in the C-projection, i.e. either in the C-position itself or in an interrogative operator in the CPs specifier (SpecCP). For the purposes of this article, we can ignore the question whether a distinction should be made between clause type and a speech act projection representing the speaker as proposed in Coniglio and Zegrean (2012) and Bayer et al. (2016). Here we speak of (illocutionary) Force or more precisely QForce. The DiP does not occur as high as Force but rather in the middle field of the clause. Weak pronouns obligatorily precede the DiP, and XPs with the status of discourse topics may precede it. There is much evidence that the DiPs are not adjoined like adverbs but are syntactic heads which as such project a particle phrase (PrtP; Bayer 1996; 1999; 2012; 2018; Munaro & Poletto 2004; Struckmeier 2014; Petrova 2017). This suggests the simplified clause structure in (3).

(3) $[_{\text{ForceP}}$ Force $[_{\text{TopP}}$ (...) $[_{\text{PrtP}}$ Prt° $[_{v\text{P}}$ ...]]]]

Along this structure, *wh*-clauses with *denn* can be analyzed as follows:

(4) $[_{\text{ForceP}}$ Wem$_1$ $[_{\text{Force'}}$ hat $[_{\text{PrtP}}$ denn $[_{v\text{P}}$ Max t_1 die Professorin vorgestellt?]]]]
who has QDiP Max the professor introduced
'Who did Max introduce the professor to?'

(5) $[_{\text{ForceP}}$ Wem$_2$ $[_{\text{Force'}}$ hat $[_{\text{TopP}}$ er$_1$ Max$_3$ $[_{\text{PrtP}}$ denn $[_{v\text{P}}$ t_1 t_2 t_3 vorgestellt?]]]]]
who has he Max QDiP introduced
'Who did he introduce Max to?'

The DiP requires a focal element in its scope: the entire propositional *v*P may be in focus ("broad focus") or some proper part of it ("narrow focus"). For example, in (5), focus on the verbal form *vorgestellt* is sufficient. This distinction will have no relevance for the present work and can thus be ignored here. The different word orders can be derived by scrambling elements out of *v*P across the DiP into the topic field.

But given that the DiP is distant, how can it communicate with Force? And how can the DiP ultimately contribute to Force in such a way that the different readings result? Following Bayer & Obenauer (2011), Bayer (2012) and Bayer et al (2016), we assume that the DiP is in an agreement relation with Force by virtue of a feature that mirrors the clause type in which it occurs. Importantly, this relation does not require any LF-style movement of DiP to Force. If Force has an interpretable interrogative feature *i*Q, and the DiP has a corresponding uniterpretable feature *u*Q, there can be feature sharing between a probe, namely force, and a goal, namely the Q-sensitive DiP as proposed in Pesetsky & Torrego (2007) and adopted by Bayer and colleagues. Agreement can be seen as feature sharing which is expressed here by an arbitrary value by which the uninterpretable Q-feature disappears and the semantic features inherent in the DiP become part of the illocutionary meaning of interrogative force. Assume the following simplified representations in which we choose the arbitrary value 5 and indicate feature-deletion by strike-out.

(6) $[_{\text{ForceP}}$ Force $_{i\text{Q}[\,]}$ [... $[_{\text{PrtP}}$ Prt $_{u\text{Q}[\,]}$ $[_{v\text{P}}$ ...]]]] $=$ AGREE $\Longrightarrow$

(7) $[_{\text{ForceP}}$ Force $_{i\text{Q}[5]}$ [... $[_{\text{PrtP}}$ Prt $_{\not{u}\text{Q}[5]}$ $[_{v\text{P}}$ ...]]]]

Agreement can only operate within a local domain. Locality is understood in such a way that probe and goal must not be separated by "phases", in GB also called "cyclic nodes" or "barriers". According to standard assumptions of minimalist syntax, *v*P and CP (here called ForceP) are phases. Prt is outside *v*P and is as such accessible to agreement with Force. This implements local agreement. Notice that Prt, in our case a DiP, does not undergo movement. It stays put. This is desirable because, as we have argued above, the focus in the scope of Prt can vary. If Prt would raise to Force, the resulting information-structural differences would be neutralized.

Importantly, probe-goal agreement does not interfere with word order in the local domain, and it is fully in line with the traditional insight that DiPs do not undergo movement, see Thurmair (1989). According to Bayer (2012) and previous work, DiPs are functional heads that build up clause structure and are immobile throughout.

1.2.2 When DiPs are too far away from Force

Given what has been said so far, it is unexpected that QDiPs may occur in complement clauses that are non-interrogative. Nevertheless, relevant examples can be found.

(8) *http://mein-kummerkasten.de/142829/fremdgehen.html* (July 18th 2013)
Wie denkst du, dass es denn weitergehen soll mit euch?
how think you that it QDiP on-go should with you.PL
'How do you think that the two of you should carry on?'

The matrix-clause is a question but the embedded clause is not as can easily be verified. The verb *denken* cannot select an interrogative complement clause. Thus, it looks as if QForce and QDiP are separated by a non-interrogative CP-node; thus, QForce seems to be "too far away" from the DiP to be able to agree with it. Notice that there are examples which verify this.[5]

(9) *Wer hat bedauert [dass die Gäste denn zu spät gekommen sind]?
who has regretted that the guests QDiP too late come are
'Who regretted that the guests arrived too late?'

(10) *Wer hat sich gefreut [dass die Kinder denn einen Hund bekommen haben]?
who has REFL rejoiced that the children QDiP a dog gotten have
'Who was happy that the children got a dog?'

Syntactic theory offers a straightforward explanation for the contrast between (8) and (9)–(10). In (9) and (10), the local domain of *denn* offers no access to interrogative Force, but it does in (8). The reason is that – once again due to syntactic locality – the *wh*-element *wie* (how) must have passed through the specifier of CP (SpecCP) or what we call here SpecForceP.

(11) [$_{\text{ForceP}}$ wie $_{iQ[7]}$ [denkst du [$_{\text{ForceP}}$ wie $_{uQ[7]}$ [dass es denn $_{uQ[7]}$ [$_{vP}$ weitergehen soll ...]]]]]

(8)/(11) is a question by virtue of *wh*-movement to the root clause. Thus, the Q-feature must be interpretable at this position. In the intermediate position, it cannot be interpretable as this would be in conflict with the verb *denken* ('to think'). Nevertheless, the clause in its complement position remains interrogative by virtue of the uninterpretable Q-feature *u*Q. This feature is locally accessible to the matching *u*Q feature of *denn*. Thus, we see that cyclic *wh*-movement sponsors the occurrence of a phonetically remote QDiP. This is possible because the agreement chain between QForce and QDiP is decomposed into smaller chains. This is clearly not the case in examples like those in (9)–(10). There, the QDiP is indeed too far away from interrogative Force.

1.3 SEMANTIC LICENSING OF QDIPS

Romero (2017) presents a semantic analysis of the distributional restrictions of QDiPs. While the syntactic approach exploits local feature agreement (along a chain) between the QDiP and the Force head, the semantic approach capitalizes on a Hamblin-style treatment of *wh*-phrases and on the intuitive semantic type of QDiPs.[6] We will see first how the semantic approach applies to simple interrogative clauses and then how it handles QDiP in embedded clauses.

5 Maybe even clearer cases are adjunct islands, i.e., clauses that are typically strict islands for syntactic movement, as in (i)–(ii):

(i) *Wer hat den Unfall fotografiert [bevor die Polizei denn am Ort war]?
who has the accident photographed before the police QDiP at place was
'Who took picture of the accident before the police was present?'

(ii) *Wem hat der Papst zugewunken [ohne ihn denn anzusehen]?
who has the pope waved-at without him QDiP look-at
'Who did the pope wave at without looking at him?'

6 These two lines of explanation of the distribution of QDiPs are, as they stand, independent from each other. We leave the possibility of combining them together for future research.

DOI: 10.5334/gjgl.1203

1.3.1 Simple interrogative clauses

Romero's (2017) analysis uses the following two ingredients.

First, following Hamblin (1973) and a long tradition thereafter, *wh*-phrases are interpreted in base position and introduce sets of alternatives. In run-of-the-mill declaratives like (12), a simple NP like *Anne* denotes a concrete individual, as in (13a), and combines with the remaining elements in *v*P to produce a single proposition, as in (13b); the *v*P-proposition in turn keeps composing with other potential elements until it reaches ForceP, as in (13c). In contrast, in interrogative clauses like (14), a *wh*-phrase like *who* denotes a set of individuals, as in (15a), and combines with the rest of the elements in the *v*P to produce a set of propositions, as in (15b); this set of propositions, again, keeps composing with other potential constituents until ForceP is reached, as in (15c):

(12) a. Max saw Anne.
b. LF: $[_{\text{ForceP}} \ldots [_{vP}$ Max saw Anne]]

(13) a. ⟦Anne⟧ = a(nne)
b. $⟦[_{vP}$ Max saw Anne]⟧ = λw. max see anne in w
c. $⟦[_{\text{ForceP}} \ldots [_{vP}$ Max saw Anne]]⟧ = λw. max see anne in w

(14) a. Who did Max see?
b. LF: $[_{\text{ForceP}}$ Q … $[_{vP}$ Max saw who]]

(15) a. ⟦who⟧ = {x: x is human}
$=_{\text{e.g.}}$ {a(nne), b(eatrice), c(arl)}
b. $⟦[_{vP}$ Max saw who]⟧ = {λw. max see x in w : x is human}
c. $⟦[_{\text{ForceP}} \ldots [_{vP}$ Max saw who]]⟧
= {λw. max see x in w : x is human}
$=_{\text{e.g.}}$ {λw. max see a in w, λw. max see b in w, λw. max see c in w}

Second, QDiPs take as argument the (ordinary) semantic value of their syntactic sister and require it to be of question type, i.e., they require it to be a set of propositions (type <<s,t>,t>). The job of the QDiP is to pass up this question meaning as the at-issue content and to predicate some relation between the question meaning and the current CG as non-at-issue-content, as illustrated in (16) for *denn*:[7]

(16) $⟦\text{DENN}_{<<<s,t>,t>,<<s,t>,t>>}⟧(Q_{<<s,t>,t>})$
a. At-issue meaning: Q
b. Non-at-issue meaning: λw. Q is motivated by the previous interactional context in w

These two ingredients derive the distribution of QDiP across clause types – declarative vs. interrogative – as follows. When the QDiP is inserted in a simple interrogative like (17), the QDiP will encounter a sister-*v*P of the appropriate semantic question type <<s,t>,t> and the semantic derivation will succeed, as in (18).[8] But, when the QDiPs is inserted in a declarative like (19), the

7 Since we are not concerned with the exact lexical meaning of *denn* (see fn. 1), we will simply state its non-at-issue meaning in terms of König's (1977) intuitive motivation through the interactional context, as in (16b). For more recent formulations of the link to the previous context, see e.g. Gutzmann (2015) and Theiler (2020). (Theiler's formulation finds, additionally, a common lexical core between the QDiP *denn* and the causal conjunction *denn*). For additional lexical pre-conditions, see Rapp's (2018) (28) below.

8 In the case of yes/no-questions, overt or covert *whether* serves as a *wh*-phrase ranging over possible values of the polarity head (Han & Romero 2004; Guerzoni & Sharvit 2014): the positive polarity value (i.e., the identity function $\lambda p_{<s,t>}.p$) and the negative value (i.e., $\lambda p_{<s,t>}.\neg p$). There is yet another construction, known as '*Why*-like-*what*' and illustrated in (i), in which the QDiP works fine but the *wh*-phrase has been argued to be base-generated directly in SpecCP (Bayer & Obenauer 2011). For the present semantic analysis to derive this case, either the *wh*-phrase would have to be generated as part of the propositional content of the sentence (i.e., under the Q-morpheme) or *denn* would have to target the entire CP as its semantic argument. We leave the choice open for future research.

(i) Was schaust du denn (*was) so dumm!
what look you QDiP (*what) so stupid
'Why do you look so stupid!'

DOI: 10.5334/gjgl.1203

simple propositional type <s,t> provided by its syntactic sister does not match the type required by the QDiP. This means that the semantic derivation crashes, leading to ungrammaticality, as in (20):

(17) a. Who did Max DENN see?
b. LF: [$_{\text{ForceP}}$ Q … DENN [$_{\text{PrtP}}$ [$_{vP}$ Max saw who]]]

(18) ⟦[$_{\text{PrtP}}$ DENN [$_{vP}$ Max saw who]]⟧
= ⟦DENN⟧ (⟦[$_{vP}$ Max saw who]⟧)
= ⟦DENN$_{\text{<<s,t>,t>, <<s,t>,t>>}}$⟧ ({λw. max see x in w : x is human}$_{\text{<s,t>,t>}}$)

(19) a. *Max hat Anne DENN gesehen.
b. LF: [$_{\text{ForceP}}$ … DENN [$_{\text{PrtP}}$ [$_{vP}$ Max saw Anne]]]

(20) ⟦[PrtP DENN [$_{vP}$ Max saw Anne]]⟧
= ⟦DENN⟧ (⟦[$_{vP}$ Max saw Anne]⟧)
= ⟦DENN$_{\text{<<s,t>,t>, <<s,t>,t>>}}$⟧ ([λw. max see anne in w]$_{\text{<s,t>}}$)
Type Mismatch!

1.3.2 Complex interrogative clauses

Recall that, when the QDiP is located in an embedded clause, there is a contrast between acceptable sentences like (8), which feature long extraction of the *wh*-phrase from the complement clause, and unacceptable sentences like (9)–(10), with short *wh*-extraction from the matrix clause. The ingredients introduced above derive the contrast between long and short extraction as follows.

Following Hamblin (1973), the grammatical (21a) has the LF representation in (21b), with the *wh*-phrase *how* in base position. The *wh*-phrase denotes a set of alternatives, as illustrated in (22a), which combines with the remaining elements in *v*P2 to produce the set of propositions (22b). As we saw, the QDiP must then combine with the meaning of its sister *v*P2, with the prerequisite that this meaning be a set of propositions (type <<s,t>,t>). Since *v*P2 provides an object of the required type <<s,t>,t>, the semantic derivation proceeds normally and the sentence is acceptable:

(21) a. How do you think [that DENN the two of you should carry on]?
b. LF: [$_{\text{ForceP}}$ ~~how~~ Q … [$_{vP1}$ you think [$_{CP2}$ ~~how~~ [$_{\text{PrtP}}$ DENN [$_{vP2}$ the two of you should carry on how]]]]]

(22) a. ⟦how⟧ = {m: m is a manner}
=$_{\text{e.g.}}$ {m(anner)$_1$, m$_2$, m$_3$}
b. ⟦[$_{vP2}$ the two of you should carry on how]⟧
= {λw. the two of you should carry on in manner m in w : m is a manner}
=$_{\text{e.g.}}$ {λw. the two of you should carry on in manner m$_1$ in w,
λw. the two of you should carry on in manner m$_2$ in w,
λw. the two of you should carry on in manner m$_3$ in w}
c. ⟦[$_{\text{PrtP}}$DENN [$_{vP2}$ the two of you should carry on how]]⟧
= ⟦DENN⟧ (⟦[$_{vP2}$ the two of you should carry on how]⟧)
= ⟦DENN⟧ ({λw. the two of you should carry on in manner m in w : m is a manner})

The ungrammatical (23a) has the LF representation in (23b). Crucially, now the base position of the *wh*-phrase is within the matrix *v*P1, not within the embedded *v*P2. This means that, when computing the meaning of the embedded *v*P2, no trigger of alternatives is present and no set of propositions is produced; the semantic value of *v*P2 is the single proposition in (24a) (type <s,t>). The QDiP must then combine with this meaning. But, as the QDiP requires a <<s,t>,t> object as argument but only an <s,t> object is encountered, a type mismatch arises. Hence, the semantic derivation cannot proceed and the sentence is unacceptable:

(23) a. Who regretted [that the guests DENN arrived too late]?
b. LF: [$_{\text{ForceP}}$ ~~who~~ Q … [$_{vP1}$ who regretted [$_{CP2}$ that [$_{\text{PrtP}}$ DENN [$_{vP2}$ the guests arrived too late]]]]]

(24) a. ⟦[$_{vP2}$ the guests arrived too late]⟧
= λw. the guests arrived too late in w

b. ⟦[$_{PrtP}$ DENN [$_{vP2}$ the guests arrived too late]]⟧
= ⟦DENN⟧ (⟦[$_{vP2}$ the guests arrived too late]⟧)
= ⟦DENN⟧ (λw. the guests arrived too late in w)
Type Mismatch!

To sum up sub-sections 1.2 and 1.3, while availing themselves to different tools, the syntactic approach and the semantic approach make the same predictions with regard to the distribution of QDiPs. They predict that: (i) QDiPs are ungrammatical in declarative clauses, (ii) they are licit in interrogative sentences if the QDiP occurs on the path of the *wh*-phrase and (iii), crucially, they are *i*llicit in interrogative sentences if the QDiP does *not* occur on the path of the *wh*-phrase.

1.4 RELATED PHENOMENA IN PSYCHOLINGUISTICS

The majority of the related psycholinguistic literature[9] deals with the processing of either questions or discourse particles, but rarely with both at the same time.

Prominent topics in the literature on online processing of German discourse particles are crosslinguistic comparisons of strategies for marking discourse (see, e.g., Dimroth et al. 2010; Turco et al. 2014), and the processing of different readings of DiPs and their counterparts; especially the increase in processing cost associated with the non-at-issue reading of discourse particles relative to their at-issue counterparts (Bayer 1991; see also Dörre et al. 2018; Dörre 2018 for an overview).[10] We are aware of only one study investigating the licensing constraints of QDiPs empirically with the help of a quantitative method (Bayer et al. 2016, outlined in more detail below).

In the literature on the processing of questions, a long line of research is concerned with the relation between the *wh*-filler and its associated gap. Earlier findings suggest that the parser employs an active filler strategy, i.e., that speakers use structural information to actively predict gaps (Stowe 1986; Frazier & Flores d'Arcais 1989). In multiple *wh*-dependencies, the active search for multiple gap sites is visible in online processing in coordinate structures, but not in adjuncts, suggesting that parsing decisions are informed by detailed grammatical constraints (Wagers & Phillips 2009). The workload associated with processing *wh*-questions and filler-gap dependencies includes the cost of keeping the filler in working memory, and of integrating a filler with its gap; the two are associated with different EEG correlates which are sensitive to different manipulations of the stimuli (Fiebach et al. 2001; Felser et al 2003). Based on evidence from the processing of Japanese questions, Aoshima et al. (2004) conclude that dependency formation is not delayed in verb-final languages, and that the constraints to be satisfied during the processing of questions are satisfied incrementally. They also conclude that the preference for assuming short-filler gap dependencies in English is not universal. Investigating the processing of *wh*-in-situ elements in Mandarin, Xiang and colleagues found that the covert dependency between the clause-initial scope position and the *wh*-in-situ element is reflected in processing (Xiang et al. 2014; 2015).

9 For reasons of space, we omit an in-depth discussion of the licensing of negative polarity items (NPIs), which could be argued to have comparable properties (see above). See Baker (1970); Linebarger (1987); Kadmon & Landman (1993); Krifka (1995); Israel (2004); Chierchia (2006); Giannakidou (2006); Horn (2010); Giannakidou (2011) for theoretical descriptions of NPI licensing conditions and constraints, Saddy et al. (2004); Drenhaus et al. (2005); Vasishth et al. (2008); Xiang et al. (2009; 2013); Yurchenko et al. (2013); Liu et al. (2019) for licensing violations and intrusive licensing effects in sentence comprehension, and Vasishth et al. (2008); Xiang et al. (2009); Parker & Phillips (2016) for different accounts of intrusive licensing. We will briefly return to this topic in the general discussion. For similar reasons we are omitting a discussion of the much richer literature on focus particles, which, to some extent, are related to discourse particles (see Gotzner & Spalek 2017; Spalek & Zeldes 2017; Gotzner & Spalek 2019 for recent overviews).

10 To illustrate the difference between the discourse particle and counterpart readings, consider *nur* (literally 'only'):

(i) Iss **nur** von dem Kuchen!
eat NUR from the cake
Discourse particle reading: 'Go ahead and have some of this cake! (Don't be shy.)'
Counterpart reading: 'Eat **only** from the cake! (Don't touch the other food you may see.)'

To our knowledge, there is only one study investigating the role of questions with long-distance licensing of QDiPs, namely Bayer et al. (2016). In this article, the authors set out to monitor the reflections of cyclic *wh*-movement in offline acceptability ratings. QDiP licensing is used as a diagnostic tool to reveal if *wh*-elements move to an intermediate landing position in SpecCP of embedded clauses before arriving at their final sentence-initial landing site. In a first rating study, the authors investigate the licensing of *denn* with *wh*-elements extracted either from the root clause (short extraction) or the embedded clause (long extraction). They find that if the *wh*-element is extracted from the root clause, ratings drop for *denn* in embedded (z = –.33) relative to root (z = –.42) positions. This indicates that, in short extractions environments, *denn* in embedded clauses is not properly licensed. However, if the *wh*-element is extracted from the embedded clause, ratings are similar for *denn* in either position (z = –.15 for root clauses, and –.19 for embedded clauses). This is in line with the syntactic and semantic accounts of QDiP licensing outlined above. In a second rating study, long extraction conditions are replaced by partial *wh*-movement. Again, *denn* is less acceptable in embedded than in root clauses (with short extraction, z = .42 in root clauses, and –.44 in embedded clauses); however, this drop in acceptability is greatly reduced if the *wh*-element is partially extracted from the embedded clause (with partial *wh*-movement, z = .44 in root clauses, and .00 in embedded clauses). The authors interpret their findings as reflecting cyclic *wh*-movement through the embedded SpecCP, in line with Bayer (2012) and Bayer & Obenauer (2011). The qualitative difference between long extraction and partial *wh*-movement for *denn* ratings is explained as a reflection of the syntactic differences between both types of *wh*-extraction.

In Bayer et al. (2016), the focus was on assessing the existence of cyclic *wh*-movement, rather than on the processing of questions. However, the results of this study raise new questions on the influence of different licensing violations, and on how this complex licensing process affects sentence comprehension in real time. These questions are related to (i) the severity of different QDiP licensing violations, and (ii) to the processing cost for successful QDiP licensing:

(i) **Licensing violations:** Experiment 1 of Bayer et al. (2016) shows a drop in acceptability ratings for embedded *denn* with short extraction. However, ratings for this ungrammatical condition do not drop much below the level of the grammatical conditions with long extraction, hinting that participants did not perceive the violation as very severe. This surprising finding is not discussed in great detail in the original study. One possible explanation is that some of the stimulus properties (see below for details) might have a subtle influence on the outcomes. In addition, there are no conditions with *denn* in declaratives, which would allow for a comparison between a completely absent licenser and an inaccessible one. Therefore, it is necessary to replicate the original findings with adapted stimuli that contain conditions with completely absent licensers (i.e., *denn* in declaratives), and with maximally parallel conditions to allow transfer to real-time processing measures.

(ii) **Successful licensing of QDiPs:** Based on the theoretical accounts outlined above, we assume that successful licensing of QDiPs involves checking a number of different licensing constraints. If these theoretical processes are reflected in sentence comprehension, we expect increases in processing cost for licensed QDiPs relative to baselines that do not require licensing.

1.5 RESEARCH QUESTIONS

Based on the background outlined above, we set out to answer the following research questions:

- With respect to QDiP violating environments:
  - What is the relative severity of different types of QDiP licensing violations? Are ratings for QDiPs with syntactically/semantically inaccessible licensers (i.e., embedded QDiP with short extraction) similar to those for QDiPs without licensers (i.e., QDiPs in declaratives)?
  - Are licensing violations associated with increases in processing cost during online processing?

- With respect to QDiP licensing environments:
  - How acceptable are QDiPs in embedded clauses in long *wh*-extraction environments compared to other baselines?
  - Is successful QDiP licensing associated with increases in processing cost, compared to a non-QDiP baseline?
- How do the findings relate to the syntactic and semantic theories of QDiP licensing?

In the following, we will present the language material for the experiments presented in this study.

2 LANGUAGE MATERIALS

The language materials for our experiments consist of two stimulus sets. Each stimulus set was used in an offline rating study and in a self-paced reading time study.

The first stimulus set contains interrogatives with short *wh*-extraction and declaratives. This allows us to assess the severity of different violations of QDiP licensing: In declaratives, there simply is no licenser, and the violation should be severe. In interrogatives with short extraction, there is a licenser, but it is syntactically inaccessible for the QDiP in the subordinate clause.

The second stimulus set contains interrogatives with short and long *wh*-extraction. This allows us reassess the findings reported in Bayer et al. (2016) that QDiPs can be licensed by means of *wh*-traces.

In both sets of stimuli, the embedding verbs are *meinen* 'mean', *glauben* 'believe', *denken* 'think' and *sagen* 'say'.

2.1 STIMULUS SET 1: INTERROGATIVES VS. DECLARATIVES

Stimulus sentences consisted of a root clause followed by an embedded clause. We compared critical conditions containing the QDiP *denn* to baseline conditions containing the non-QDiP *jetzt*.[11]

We manipulated the following factors:

i. **DiP type**: Stimulus sentences contained either the QDiP *denn* or the non-QDiP *jetzt*.
ii. **DiP position**: The DiPs were positioned either (a) in the root clause, at the position directly preceding the final participle, or (b) in the embedded clause, at the position preceding the infinitive, i.e., the second-to-last position.
iii. **Clause type**: Stimulus sentences were either interrogative or declarative. For interrogative clauses, the *wh*-element was always extracted from the root clause. Interrogatives began with *Welche von diesen Leuten* ('Which of these people'), declaratives began with *Manche von diesen Leuten* ('Some of these people').

Root clauses were in present perfect and contained nonseparable verbs. Embedded clauses were in present tense and contained a transitive separable particle verb (occurring in their nonseparate infinitive form) and a modal verb. Subjects in the embedded clause were always animate, objects were always inanimate. Stimulus sentences were kept strictly parallel to minimize noise (e.g., due to implicit prosodic effects of separable/nonseparable verbs etc.).

Following this pattern, 40 stimulus item sets were constructed. An example of a stimulus quartet with DiPs in the root clause is given in ***Example 1***, and with DiPs in the embedded clause in ***Example 2***.

11 The counterpart reading of *denn* cannot be used as a baseline because it is a connective and cannot occur in the same syntactic position as the discourse particle *denn*. *Jetzt* can serve both as a temporal adverb ('now') and as a modal particle. The particle reading is similar to the English discourse marker 'now' (see Detges 2007 for a description of *jetzt* in questions). This particular word was chosen because its felicity is not affected by clause type, it is of comparable length and frequency to *denn* (type frequency is 78282 for *jetzt*, and 81340 for *denn* according to the dlexDB corpus, Heister et al. 2011, accessed July 20th 2020), and it does not trigger temporal implicatures or life time effects to the same extent as other temporal expressions (e.g., *damals*, 'back then'). We will also refer to *jetzt* as a non-QDiP, to indicate that it is equally felicitous in interrogatives and declaratives. By this, we do not imply that *jetzt* actually lacks its modal particle reading in all of our stimuli.

(A) Interrogative, *denn*:
Welche von diesen Leuten haben **denn** gemeint, dass der Gärtner das
which of these people have QDiP meant that the gardener the
Gemüse anbauen darf?
vegetables plant may
'Which of these people thought DENN that the gardener may plant the vegetables?'

(B) Interrogative, *jetzt*:
Welche von diesen Leuten haben **jetzt** gemeint, dass der Gärtner das
which of these people have non-QDiP meant that the gardener the
Gemüse anbauen darf?
vegetables plant may
'Which of these people thought JETZT that the gardener may plant the vegetables?'

(C) Declarative, *denn*:
*Manche von diesen Leuten haben **denn** gemeint, dass der Gärtner das
some of these people have QDiP meant that the gardener the
Gemüse anbauen darf.
vegetables plant may
*'Some of these people thought DENN that the gardener may plant the vegetables.'

(D) Declarative, *jetzt*:
Manche von diesen Leuten haben **jetzt** gemeint, dass der Gärtner das
some of these people have non-QDiP meant that the gardener the
Gemüse anbauen darf.
vegetables plant may
'Some of these people thought JETZT that the gardener may plant the vegetables.'

Example 1 Stimulus set 1, glossed example of a typical sentence quartet, DiP in the root clause

(A) Interrogative, *denn*:
*Welche von diesen Leuten haben gemeint, dass der Gärtner das Gemüse
which of these people have meant that the gardener the vegetables
denn anbauen darf?
QDiP plant may
*'Which of these people thought that the gardener may DENN plant the vegetables?'

(B) Interrogative, *jetzt*:
Welche von diesen Leuten haben gemeint, dass der Gärtner das Gemüse
which of these people have meant that the gardener the vegetables
jetzt anbauen darf?
non-QDiP plant may
'Which of these people thought that the gardener may JETZT plant the vegetables?'

(C) Declarative, *denn*:
*Manche von diesen Leuten haben gemeint, dass der Gärtner das Gemüse
some of these people have meant that the gardener the vegetables
denn anbauen darf.
QDiP plant may
*'Some of these people thought that the gardener may DENN plant the vegetables.'

(D) Declarative, *jetzt*:
Manche von diesen Leuten haben gemeint, dass der Gärtner das Gemüse
Some of these people have meant that the gardener the vegetables
jetzt anbauen darf.
non-QDiP plant may
'Some of these people thought that the gardener may JETZT plant the vegetables.'

Example 2 Stimulus set 1, glossed example of a typical sentence quartet, DiP in the embedded clause

2.2 STIMULUS SET 2: SHORT VS. LONG EXTRACTION

Stimulus sentences were essentially parallel to those in stimulus set 1, consisting of a root clause followed by an embedded clause. We compared between critical conditions containing the QDiP *denn* and baseline conditions containing the non-QDiP *jetzt*. In contrast to the first stimulus set, all sentences were interrogatives.

We manipulated the following factors:

i. **DiP type**: Stimulus sentences contained either the QDiP *denn* or the non-QDiP *jetzt*.

ii. **DiP position**: The DiPs were positioned either (a) in the root clause, at the position directly preceding the final participle, or (b) in the embedded clause, at the position preceding the infinitive, i.e., the second-to-last position.

iii. **(*Wh*-)extraction**: The *wh*-element was either extracted from the root clause (short extraction, asking for the subject of the root clause) or from the embedded clause (long extraction, asking for the object of the embedded clause).[12]

Following this pattern, 56 stimulus items were constructed. An example of a stimulus quartet with DiPs in the root clause is given in ***Example 3***, and with DiPs in the embedded clause in ***Example 4***.

(A) Short extraction, *denn*:
Wer hat **denn** gemeint, dass der Türsteher den Musiker abweisen soll?
who has QDiP meant that the bouncer the musician away.turn should
'Who thought DENN that the bouncer should turn away the musician?'

(B) Short extraction, *jetzt*:
Wer hat **jetzt** gemeint, dass der Türsteher den Musiker abweisen soll?
who has non-QDiP meant that the bouncer the musician away.turn should
'Who thought JETZT that the bouncer should turn away the musician?'

(C) Long extraction, *denn*:
Wen hast du **denn** gemeint, dass der Türsteher abweisen soll?
who.ACC have you QDiP meant that the bouncer away.turn should
'Who did you DENN think that the bouncer should turn away?'

(D) Long extraction, *jetzt*:
Wen hast du **jetzt** gemeint, dass der Türsteher abweisen soll?
who.ACC have you non-QDiP meant that the bouncer away.turn should
'Who did you JETZT think that the bouncer should turn away?'

Example 3 Stimulus set 2, glossed example of a typical sentence quartet, DiP in the root clause

The goal of the second experiment was twofold: (1) to replicate the findings for short vs. long *wh*-extraction reported in Bayer et al. (2016) with a new stimulus set, and (2) to establish a maximally parallel stimulus set showing a comparable effect, that can also serve for recording online measures.

To enhance comparability with the earlier study, we outline the main differences to the stimuli used by Bayer et al. (2016) here:

i. Tense: In the current stimulus set, root clauses are in present perfect, and embedded clauses are in present tense. All embedded clauses end with a modal verb. In the older stimulus set, root clauses are in simple past, and embedded clauses are in simple past for short extraction conditions and in present perfect for long extraction conditions.

12 We chose to ask for the object instead of the subject of the embedded clause to avoid a garden path. The root clause verbs do not have animate direct objects, so the *wh*-element *wen* (who.ANIM.ACC, 'who') unambiguously refers to the object of the embedded clause.

(A) Short extraction, *denn*:
*Wer hat gemeint, dass der Türsteher den Musiker **denn** abweisen soll?
who has meant that the bouncer the musician QDiP away.turn should
*'Who thought that the bouncer should DENN turn away the musician?'

(B) Short extraction, *jetzt*:
Wer hat gemeint, dass der Türsteher den Musiker **jetzt** abweisen soll?
who has meant that the bouncer the musician non-QDiP away.turn should
'Who thought that the bouncer should turn away JETZT the musician?'

(C) Long extraction, *denn*:
Wen hast du gemeint, dass der Türsteher **denn** abweisen soll?
who.ACC have you meant that the bouncer QDiP away.turn should
'Who did you think that the bouncer should DENN turn away?'

(D) Long extraction, *jetzt*:
Wen hast du gemeint, dass der Türsteher **jetzt** abweisen soll?
who.ACC have you meant that the bouncer non-QDiP away.turn should
'Who did you think that the bouncer should JETZT turn away?'

Example 4 Stimulus set 2, example of a typical sentence quartet, DiP in the embedded clause

ii. Voice: In the current stimulus set, all clauses are active. In the old stimulus set, embedded clauses with short extraction are passive, while all other clauses are active.

iii. Parallelism: In the current stimulus set, all clauses are strictly parallel across conditions. For each item, the same embedding verbs and arguments are used. The interrogatives refer to arguments of the root clause and embedded clause verbs (subjects of root clauses, and objects of embedded clauses). In the old stimulus set, the sentences belonging to one item follow a semantically similar storyline, but do not necessarily use identical lexical material. In addition, embedding verbs in the long and short extraction conditions have different numbers of arguments.

If the new stimulus set will show comparable acceptability for *denn* in embedded clauses with long *wh*-extraction, it would replicate earlier findings reported by Bayer et al. (2016), and support the view that these findings were not due to the non-relevant differences between stimulus conditions outlined above. In addition, the new stimuli can be used in experiments using time-sensitive measures like self-paced reading.

3 EXPERIMENT 1: ACCEPTABILITY RATINGS, INTERROGATIVES VS. DECLARATIVES

In the first experiment, we compared the acceptability of *denn* with accessible, inaccessible and absent licensing *wh*-elements. The goal of this experiment was to assess the relative severity of different types of QDiP licensing violations, either by absent licensers (i.e., QDiPs in declaratives) or by inaccessible licensers (i.e., QDiPs in interrogatives, but with a CP boundary between the *wh*-element and the QDiP). Following the theoretical description outlined in 1.2 and 1.3, both violations should make the sentences ungrammatical.

3.1 MATERIALS AND METHODS

Participants 58 participants were recruited via the SONA systems database of the University of Konstanz. Participants were between 18 and 35 years of age. All participants spoke German as their only native language. All participants had normal or corrected-to-normal vision, and reported no neurological or reading-related disorders. The data from one participant were removed before the final data analysis because he/she had reported having switched the order of the rating scale during the experiment. The remaining participants were between 18 and 30 years old, their mean age was 23 years (s.d. = 3 years). 46 participants were female.

Language material The language material was the first stimulus set, detailed in section 2.1. In total, each participant saw 120 experimental sentences. Of these, 80 were critical sentences

(10 per condition), and 40 were filler sentences (20 grammatical and 20 ungrammatical). Each participant saw two out of the eight sentences from each item set; these sentences were different with respect to clause type (either interrogative or declarative) and to the DiP (either *denn* or *jetzt*). Before the presentation of the stimuli, participants saw five practice items. After 46 sentences, participants were offered a break.

Procedure The experiment was run in the Psycholinguistics Lab of the University of Konstanz as a Magnitude Estimation study, following the procedure outlined in Bader (2012). All sentences were rated relative to a reference sentence. The reference sentence was *Die Mitarbeiter haben dass der Chef Probleme hat wohl nicht sofort bemerkt* ('Apparently, the coworkers did not notice right away that the boss was having problems.'). The acceptability for this reference sentence was set to 50. Participants were instructed to rate sentences with higher acceptability with higher values, and sentences with lower acceptability with lower values. The lower limit for (bad) ratings was 1; there was no upper limit to the possible ratings. Before the start of the actual experiments, participants rated 5 practice sentences. Every participant saw 80 critical sentences (10 per condition), interspersed with 40 filler sentences. The experiment was performed using a 17" cathode ray tube monitor (Sony Trinitron Multiscan G400), connected to a Fujitsu personal computer. Stimuli were presented and ratings were recorded in Linger (Rohde 2003).

Data preparation and analysis Data were prepared for statistical analysis in R (R Development Core Team 2019), using core functions and the packages *reshape* (Wickham 2007), *plyr* (Wickham 2011), and *car* (Fox & Weisberg 2011). The worst value that could be assigned to a sentence was 1. 42 ratings (0.6% of the data) were removed because participants had rated sentences as '0'.

Following Bader (2012), the remaing rating values were normalized by dividing each rating by 50 (the reference value) and subsequent log-transformation. Outliers were defined as values that deviated more than two standard deviations from a participant's mean per condition, and were removed before the final data analysis. 3.2 % of the raw data were removed as outliers. Ratings were *z*-scaled for each participant.

z-scaled ratings were analyzed using a linear mixed effects model in R. Statistical analyses were performed on log-transformed *z*-scaled ratings with linear mixed models, using the package *lme4* (Bates et al. 2015, *lme4* function), and *LMERConvenienceFunctions* (Tremblay & Ransijn 2015, *summary* function). We defined the following factors: CLAUSE TYPE (declarative or interrogative); DiP TYPE (QDiP *denn* or non-QDiP *jetzt*); and POSITION (root or embedded). We began our analyses with a maximal random effects structure and then reduced the random effects structure, beginning on random effects structures for items, until the model would converge.

3.2 RESULTS

Mean ratings per condition are given in ***Table 1***. A visualization of the mean ratings is given in ***Figure 1***. Descriptively speaking, declaratives with *denn* (the question-sensitive discourse particle) received very low ratings for both DiP positions, while declaratives with the neutral discourse particle *jetzt* received ratings indicating that they were perceived as grammatical. For interrogative clauses, sentences with *jetzt* received relatively high ratings, as did those with *denn* in the root clause position. The ratings dropped for *denn* in the embedded clause, although less dramatically than for the declaratives containing *denn*.

CONDITION	NORM. RATINGS
interrogative-*denn*-root	.52 (.31)
interrogative-*denn*-embedded	–.24 (.58)
interrogative-*jetzt*-root	.44 (.28)
interrogative-*jetzt*-embedded	.48 (.32)
declarative-*denn*-root	–.90 (.56)
declarative-*denn*-embedded	–.86 (.50)
declarative-*jetzt*-root	.43 (.34)
declarative-*jetzt*-embedded	.59 (.27)

Table 1 Normalized *z*-scaled mean ratings over participants for all conditions, Experiment 1. Standard deviations are given in parentheses.

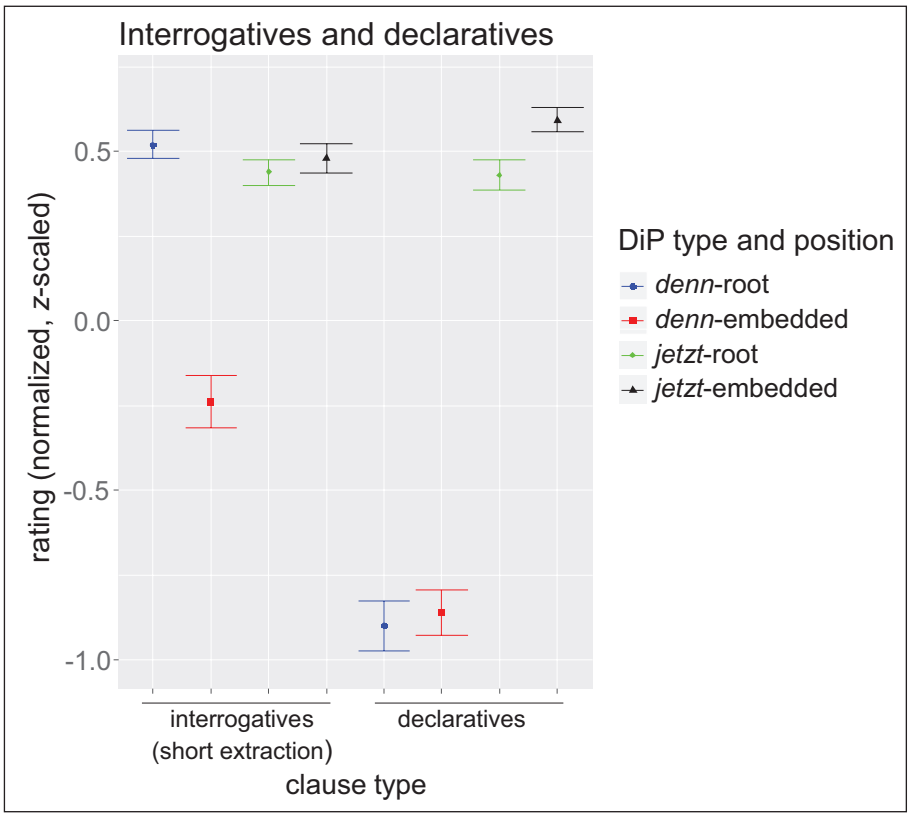

Figure 1 Normalized *z*-scaled mean acceptance rates per condition, Experiment 1. Error bars indicate standard error of the mean.

Density plots of ratings are depicted in ***Figure 2***. The density plots show that ratings for each condition are normally or close to normally distributed, i.e., that the mean ratings for each condition are not the result of widely varying ratings.

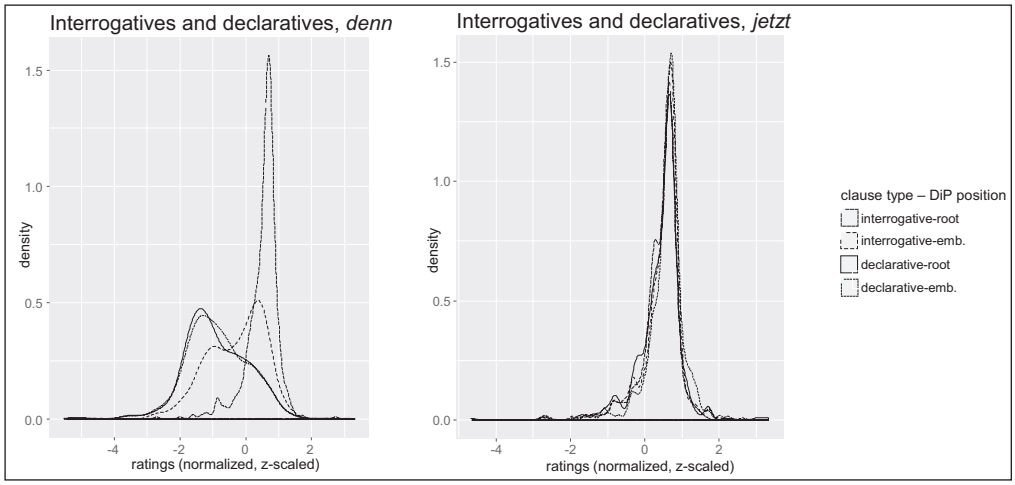

Figure 2 Density plots of ratings, Experiment 1. Conditions with *denn* are depicted in the left panel, conditions with *jetzt* are depicted in the right panel.

We analyzed the differences between individual conditions with a series of linear mixed-effects models. Full tables of the fixed effects of all models outlined below are given in the Appendix in Table 5. For Model 1.1, we analyzed the whole dataset, specifying the main effects and interactions of all three factors as fixed effects, and participant and item as random intercept. In addition, the interactions of DiP TYPE and POSITION, and of DiP TYPE and CLAUSE TYPE were specified as random slopes for participant.

The analysis revealed a statistically significant main effect of DiP TYPE ($t = 15.32$, $p < .001$), and of CLAUSE TYPE ($t = 15.80$, $p < .001$), and interactions of DiP TYPE with CLAUSE TYPE ($t = -15.82$, $p < .001$), POSITION and CLAUSE TYPE ($t = -15.23$, $p < .001$) and of DiP TYPE, POSITION and CLAUSE TYPE ($t = 9.20$, $p < .001$).

In the next step, we analyzed declaratives and interrogatives separately.

Declaratives While it was descriptively clear that declaratives containing the question-sensitive discourse particle *denn* were rated as ungrammatical, and those containing the

discourse particle/adverb *jetzt* were rated as grammatical, we wanted to assess the subtle differences in acceptability between these conditions (especially for the *denn* cases). For Model 1.2, we specified the main effects and interactions of POSITION and DiP TYPE as fixed effects, and participant and item as random intercept. In addition, we specified the main effect and interaction of POSITION and DiP TYPE as random slope for participant. The analysis revealed a statistically significant main effect of DiP TYPE (t = 13.44, p < .001). No main effects and no interactions of POSITION were found.

This suggests that the position of the DiP did not have a big impact on the ratings, and that both kinds of *denn*-containing declaratives were rated equally badly.

Interrogatives We used the same Model 1.2 for analyzing the interrogative clauses. The analysis revealed statistically significant main effects of DiP TYPE (t = –2.57, p < .05 and POSITION (t = –8.66, p < .001), and a statistically significant interaction of DiP TYPE and POSITION (t = 8.45, p < .001).

To pursue the interaction of DiP TYPE and POSITION, we analyzed the main effect of DiP TYPE separately for each level of POSITION, and the main effect of POSITION separately for each level of DiP TYPE.

For Model 1.3, we specified the main effect of DiP TYPE as fixed effect, and participant and item as random intercepts. In addition, we specified DiP TYPE as random slope for participant.

For interrogatives with the DiP in the root clause, we found a statistically significant main effect of DiP TYPE (t = –2.85, p < .01). Ratings for interrogatives containing *denn* in the root clause were slightly better than for those containing *jetzt* in the root clause.

For interrogatives with the DiP in the embedded clause, we found a statistically significant main effect of DiP TYPE (t = 8.06, p < .001). Ratings for interrogatives containing *denn* in the embedded clause were much lower than those for interrogatives containing *jetzt* in the embedded clause.

We analyzed the interrogatives with *denn* and *jetzt* separately, using Model 1.4. For Model 1.4, we specified the main effect of POSITION as fixed effect, and participant and item as random intercepts. In addition, we specified POSITION as random slope for participant. For interrogatives with the DiP *denn*, we found a statistically significant main effect of POSITION (t = –8.64, p < .001). Interrogatives with *denn* in the embedded clause were rated as much worse than those with *denn* in the root clause.

For interrogatives with the DiP *jetzt*, we found no statistically significant difference between sentences with *jetzt* in the embedded and the root clause.

In general, mean ratings were lower for interrogatives with *denn* in embedded clauses than for the other interrogatives. For the non-QDiP *jetzt*, embedded positions were favoured only very slightly over root clause positions.

3.3 DISCUSSION

The results of the first rating study show that QDiPs need to be licensed by a *wh*-element. While interrogatives with QDiPs in the root clause receive high ratings, ratings drop when there is no licenser (i.e., declaratives), or when the licenser is not in the same clause as the *wh-element* (i.e., interrogatives with *denn* in the embedded clause). This is in line with native speakers' intuition and the basic assumptions underlying Bayer et al. (2016). A comparison between conditions with *denn* and *jetzt* shows that the ratings are not influenced by a general preference for interrogatives over declaratives in our stimuli. Interestingly, the ratings are worse for *denn* in declaratives than for embedded *denn* in interrogatives. This is not predicted by the theoretical accounts of QDiP licensing outlined in subsections 1.2 and 1.3. We will discuss possible explanations for this unexpected difference in acceptability ratings in the general discussion.

4 EXPERIMENT 2: ACCEPTABILITY RATINGS, SHORT VS. LONG EXTRACTION

In the second acceptability rating study, we compared the acceptability of QDiPs in questions with long and short *wh*-extraction. The goal of this study was to monitor the licensing of QDiPs by *wh*-traces, replicating and extending the findings reported in Bayer et al. (2016). In particular, we were interested to see whether the drop in acceptability for QDiPs in embedded

compared to root clauses would not only occur when *wh*-element is extracted from the root clause (short extraction) (see Experiment 1) but also when the *wh*-element is extracted from the embedded clause (long extraction).

4.1 MATERIALS AND METHODS

Participants 56 participants were recruited via the SONA systems database of the University of Konstanz. All participants spoke German as their only native language. All participants had normal or corrected-to-normal vision, and reported no neurological or reading-related disorders. The data from one participant were removed before the final data analysis because he/she repeatedly assigned improbably high ratings to random sentences. The remaining participants were between 19 and 34 years old, their mean age was 23 years (s.d. = 3). 39 participants were female.

Language material The language material was the second stimulus set, described in section 2.2. In total, each participant saw 196 sentences. Of these, 112 were critical sentences (14 per condition), and 84 were filler sentences (56 grammatical and 28 ungrammatical). Each participant saw two of the eight sentences from each item set; these sentences were different with respect to extraction (either short or long extraction) and to the DiP (either *denn* or *jetzt*). Before the presentation of the stimuli, participants saw five practice items. After each 46 sentences, participants were offered a break.

Procedure The procedure was the same as for Experiment 1, detailed in section 3.

Data preparation and analysis Data were prepared and analyzed in the way described for Experiment 1 in 3. 77 data points (0.7%) were removed because participants had rated sentences as '0'. Data from one participant were removed from the data set because he/she repeatedly assigned arbitrary and absurdly high values to individual ratings. 4.5 % of the raw data were removed as outliers. Statistical analysis was performed on normalized and *z*-scaled ratings.

For data analysis, we defined the following factors: DiP TYPE (*denn* or *jetzt*) and POSITION (root or embedded); and EXTRACTION (short or long).

4.2 RESULTS

Mean ratings per condition are given in ***Table 2***. A visualization of the mean ratings is given in ***Figure 3***. Descriptively speaking, conditions with long extractions received worse ratings than conditions with short extraction, but were still close to the grammatical reference sentence. For short extraction conditions, *denn* in the root clause received better ratings than *denn* in the embedded clause, while *jetzt* conditions in embedded and root clauses received equally high ratings, close to the ones for interrogatives with *denn* in the root clauses. This finding mirrors the findings for interrogatives from Experiment 1. For long extraction conditions, there was no big difference in ratings between root and embedded clause *denn*.

CONDITION	NORM. RATINGS
short-*denn*-root	.55 (.35)
short-*denn*-embedded	.09 (.55)
short-*jetzt*-root	.43 (.36)
short-*jetzt*-embedded	.44 (.37)
long-*denn*-root	–.15 (.47)
long-*denn*-embedded	–.24 (.54)
long-*jetzt*-root	–.17 (.45)
long-*jetzt*-embedded	–.08 (.47)

Table 2 Normalized *z*-scaled mean ratings over participants for all conditions, Experiment 2. Standard deviations are given in parentheses.

Density plots of ratings are depicted in ***Figure 4***. The density plots show that ratings for each condition are normally or close to normally distributed, i.e., that the mean ratings for each condition are not the result of widely varying ratings.

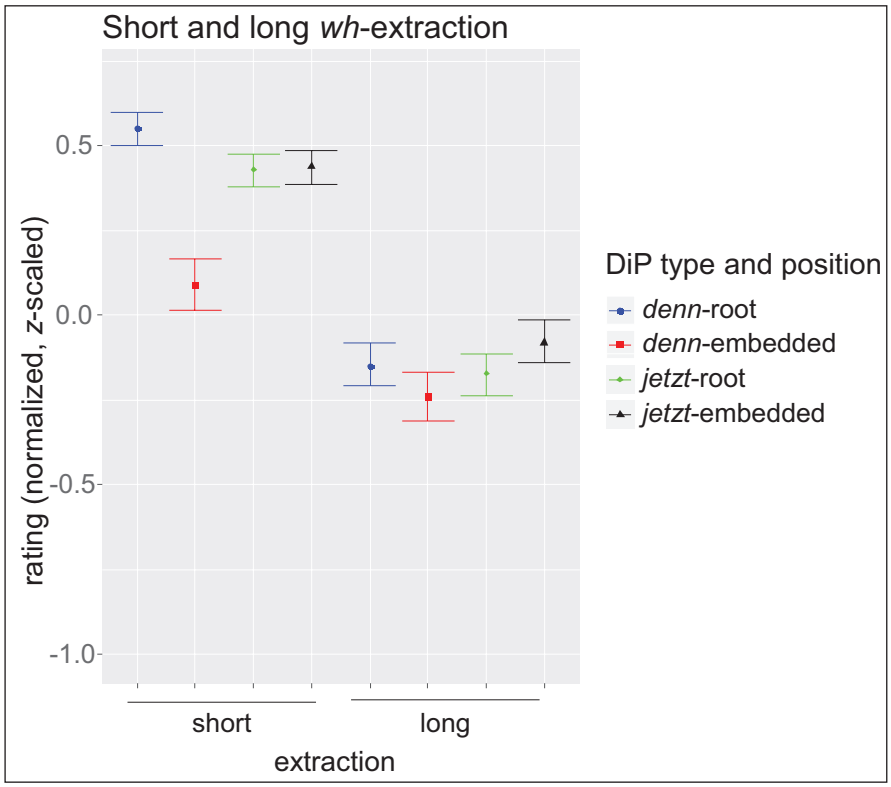

Figure 3 Normalized *z*-scaled mean acceptance rates per condition, Experiment 2. Error bars indicate standard error of the mean.

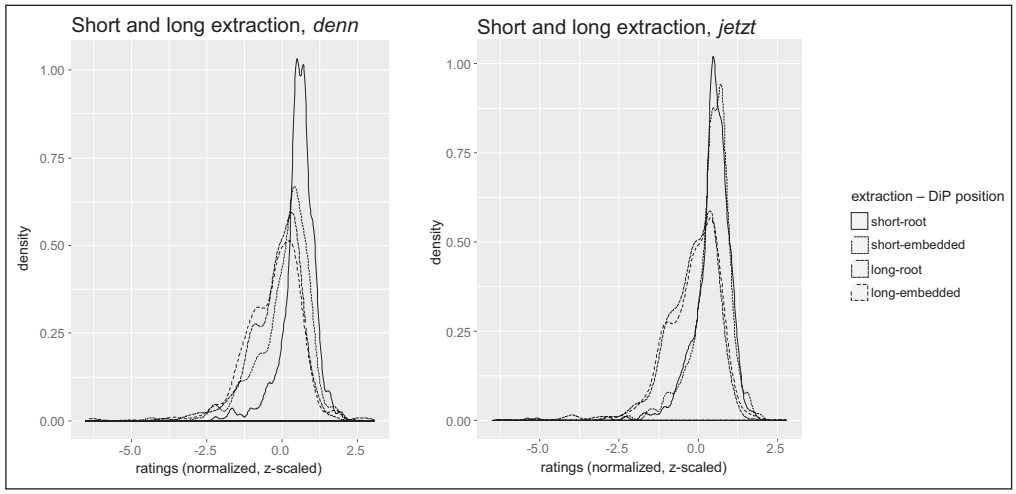

Figure 4 Density plots of ratings for Experiment 2. Conditions with *denn* are depicted in the left panel, conditions with *jetzt* are depicted in the right panel.

We analyzed the differences between individual conditions with a series of linear mixed-effects models. Full tables of the outcomes for the fixed effects of the models described below are given in the Appendix in Table 6. For Model 2.1, we analyzed the whole data set, specifying the main effects and interactions of all three factors as fixed effects, and participant and item as random intercept. In addition, the interactions of DiP TYPE and POSITION, and of DiP TYPE and EXTRACTION were specified as random slopes for participant. The analysis revealed statistically significant main effects of DiP TYPE (t = –3.08, p < .01), POSITION (t = –6.08, p < .001) and EXTRACTION (t = –8.30, p < .001), and interactions of DiP TYPE with POSITION (t = 5.60, p < .001), POSITION with EXTRACTION (t = 7.48, p < .001), and of DiP TYPE, POSITION and EXTRACTION (t = –4.16, p < .001). In addition, there was a marginally significant interaction of DiP TYPE with EXTRACTION (t = 1.88, p < .07).

In the next step, we analyzed short and long extraction conditions separately.

Short extraction Short extraction conditions were identical to the interrogative conditions in Experiment 1, and patterned in a descriptively similar fashion: While the position of the DiP *denn* in the embedded clause led to lower ratings, the position of the DiP *jetzt* did not affect ratings. For Model 2.2, we specified the main effects and interactions of POSITION and DiP TYPE as fixed effects, and participant and item as random intercept. In addition, we specified the main effect and interaction of POSITION and DiP TYPE as random slope for participant. The analysis revealed

statistically significant main effects of POSITION (t = –5.24, p < .001) and DiP TYPE (t = –3.82, p < .001), and a statistically significant interaction of DiP TYPE and POSITION (t = 5.01, p < .001).

To pursue the interaction of DiP TYPE and POSITION, we analyzed the main effect of DiP TYPE separately for each level of POSITION, and the main effect of POSITION separately for each level of DiP TYPE.

For Model 2.3, we specified the main effect of DiP TYPE as fixed effect, and participant and item as random intercepts. In addition, we specified DiP TYPE as random slope for participant.

For short extraction conditions with the DiP in the root clause, we found a statistically significant main effect of DiP TYPE (t = –4.53, p < .001). Ratings for interrogatives containing *denn* in the root clause were slightly better than for those containing *jetzt* in the root clause.

For short extraction conditions with the DiP in the embedded clause, we found a statistically significant main effect of DiP TYPE (t = 4.33, p < .001). Ratings for interrogatives containing *denn* in the embedded clause were lower than those for interrogatives containing *jetzt* in the embedded clause.

We analyzed the short extraction conditions with *denn* and *jetzt* separately, using Model 2.4. For Model 2.4, we specified the main effect of POSITION as fixed effect, and participant and item as random intercepts. In addition, we specified POSITION as random slope for participant. For interrogatives with *denn*, we found a statistically significant main effect of POSITION (t = –5.30, p < .001). Interrogatives with *denn* in the embedded clause were rated worse than those with *denn* in the root clause.

For interrogatives with *jetzt*, there was no statistically significant effect of POSITION (p > .8).

In general, mean ratings were lower for short extraction conditions with *denn* in embedded clauses than for the other short extraction conditions. For *jetzt*, the position of the DiP did not affect ratings.

Long extraction Descriptively, all long extraction conditions were rated worse than short extraction conditions. There did not seem to be a strong influence of either DiP TYPE or POSITION distinguishing the long extraction conditions, perhaps with slightly better ratings for *jetzt* in embedded clauses than for the other three conditions. Ratings for long extraction conditions were analyzed with Model 2.2, outlined above for the short extraction conditions. The analysis revealed a statistically significant interaction of DiP TYPE and POSITION (t = 2.37, p < .05).

To pursue the interaction of DiP TYPE and POSITION, we analyzed the main effect of DiP TYPE separately for each level of POSITION, and the main effect of POSITION separately for each level of DiP TYPE. We analyzed the long extraction conditions with positions root clause and embedded clause separately, using Model 2.3, outlined above. For long extraction conditions with the DiP in the root clause, we found no statistically significant effect of DiP TYPE (p > .7). For long extraction conditions with the DiP in the embedded clause, we found a statistically significant main effect of DiP TYPE (t = 2.28, p < .05). Ratings for long extraction conditions containing *denn* in the embedded clause were slightly lower than those for long extraction conditions containing *jetzt* in the embedded clause.

We analyzed the long extraction conditions with *denn* and *jetzt* separately, using Model 2.4, outlined above for short extraction conditions. For long extractions with *denn*, we found no statistically significant effect of POSITION (p > .2). For long extractions with *jetzt*, we found a very marginally significant effect of POSITION (t = 1.75, p < .09).

In general, mean ratings were similar for all long extraction conditions, with slightly better ratings for *jetzt* in embedded clauses.

4.3 DISCUSSION

The findings from the second acceptability rating study partly replicate the findings from the first study (namely, those for interrogatives). With short *wh*-extraction, ratings are lower for *denn* in the embedded clause than in the root clause, suggesting that QDiP licensing does not work smoothly across CP-boundaries. In contrast, there is no difference in acceptability for *denn* in root and embedded clauses if the *wh*-element was extracted from the embedded clause.

This latter finding is in line with the idea of cyclic *wh*-movement, and also with the findings reported in Bayer et al. (2016). Long extraction conditions are rated worse than short extraction conditions, irrespective of DiP type; however, even the dispreferred long extraction conditions are close to the rating for the grammatical baseline sentence (see Bayer et al. 2016 p.608 for a similar finding, and Phillips et al. 2005 for a discussion of the processing costs of long-distance dependencies).[13] Interestingly, the ungrammatical condition (short *wh*-extraction with *denn* in the embedded clause; i.e., a QDiP with an inaccessible licenser) is not rated worse than the long extraction conditions, indicating that it is dispreferred, rather than perceived as unacceptable. This matches the results of Experiment 1, where QDiPs with inaccessible licensers received better ratings than QDiPs without licensers, but is unexpected under the theoretical accounts explained in subsections 1.2 and 1.3. We will return to this finding in the general discussion.

5 EXPERIMENT 3: SELF-PACED READING TIMES, INTERROGATIVES VS. DECLARATIVES

The third experiment was a self-paced reading time study using Stimulus set 1 (see section 2.1). This experiment had two goals: The first was to assess if the decrease in acceptability found for QDiPs in declaratives and in embedded clauses translates to an increase in reading times, signaling increased processing load. Another goal was to assess if there was an increase in reading times for QDiPs (*denn*) compared to non-QDiPs (*jetzt*) in grammatical conditions (i.e., in root clauses of interrogatives); this would suggest that the successful checking of the licensing constraints outlined in subsections 1.2 and 1.3 is reflected in online processing.

5.1 MATERIALS AND METHODS

Participants 52 participants were invited via the SONA systems participant database the University of Konstanz. All participants spoke German as their only native language. All participants had normal or corrected-to-normal vision, and reported no neurological or reading-related disorders. Before the final data analysis, one participant was removed because he/she had unusually long reaction times (see below). The remaining participants were between 18 and 35 years old, their mean age was 22 years (s.d. = 3 years). 35 participants were female.

Language material The language material was Stimulus set 1, outlined in detail in section 2.1. In total, each participant saw 120 sentences. 80 of these were critical sentences (10 per condition), and 40 were filler sentences. Each participant saw two of the eight sentences from each item set; these sentences were different with respect to clause type (either interrogative or declarative) and to the DiP (either *denn* or *jetzt*). After 30 of the sentences, participants were asked to answer a question. Before the presentation of the critical sentences, each participant completed a practice phase with six practice items.

Procedure The experiment was run in the Psycholinguistics Lab of the University of Konstanz. Sentences were presented in in a word-by-word, non-cumulative self-paced reading paradigm. Stimulus presentation and recording were performed using the same hardware and software as described for Experiment 1 in section 3. Words were presented in a black font on a white screen. Presentation was centered word-by-word presentation.

Data preparation and analysis Data were prepared for statistical analysis in R (R Development Core Team 2019), using core functions and the packages *reshape* (Wickham 2007), *plyr* (Wickham 2011), *car* (Fox & Weisberg 2011) and *Rmisc* (Hope 2013). Graphs were prepared using the *ggplot2* package (Wickham 2009). Reading times longer than 6000 ms and shorter than 200 ms were removed, leading to the removal of 4.3% of the data. Outliers were defined as values that deviated more than two standard deviations from a participant's mean for the respective position, and were removed before the final data analysis. 4.3% of the data were removed as outliers.

13 In comparison with English, long extraction from *dass*-clauses in German has stirred controversial discussion from early on. Paul (1920), referring to them as "Satzverschlingungen" ('sentence intertwinings'), saw them as semi-grammatical. Andersson & Kvam (1984) present a first detailed investigation. See also Fanselow et al. 2005 for variation in the acceptability of the construction. Transclausal movement is stigmatized by normative grammar but is clearly available to different degrees in all varieties of German. It occurs predominantly in spoken language (dialects) and can be observed in earlier stages of the language in which norms did not play much of a role.

Reading times were analyzed for the first position following the DiP. [14] Statistical analyses were performed on log-transformed reading times with linear mixed models, using the packages *lme4* (Bates et al. 2015, *lmer* function) and *LMERConvenienceFunctions* (Tremblay & Ransijn 2015, *summary* function). We defined the following factors: CLAUSE TYPE (declarative or interrogative); DiP TYPE (QDiP *denn* or non-QDiP *jetzt*); and DiP position (root or embedded). Full tables of the fixed effects for the models outlined below are given in the Appendix in Table 7.

5.2 RESULTS

Answer accuracy The mean error rate for critical conditions was 22.7% (s.d. = 16.5%), the mean error rate for the fillers was 11.2% (s.d. = 2.4%). No participants were removed because of their answer accuracy.

Reading times Mean reading times per condition for the DiP and the two following words are given in ***Table 3***. A visualization of the mean log-transformed reading times is given in ***Figure 5*** for root clause conditions, and in ***Figure 6*** for embedded clause conditions.

DIP IN ROOT CLAUSE			
CONDITION	**DIP**	***GESAGT,***	***DASS***
interrogative-*denn*	360 (82)	366 (92)	360 (83)
interrogative-*jetzt*	353 (73)	362 (109)	367 (92)
declarative-*denn*	360 (93)	387 (119)	368 (89)
declarative-*jetzt*	363 (94)	367 (108)	360 (83)
DIP IN EMBEDDED CLAUSE			
CONDITION	**DIP**	***ABWEISEN***	***SOLL?***
interrogative-*denn*	358 (81)	387 (109)	422 (134)
interrogative-*jetzt*	357 (73)	375 (99)	416 (131)
declarative-*denn*	354 (70)	399 (106)	418 (127)
declarative-*jetzt*	351 (74)	374 (102)	422 (127)

Table 3 Mean reading times over participants per condition in ms, Experiment 3. Standard deviations are given in parentheses

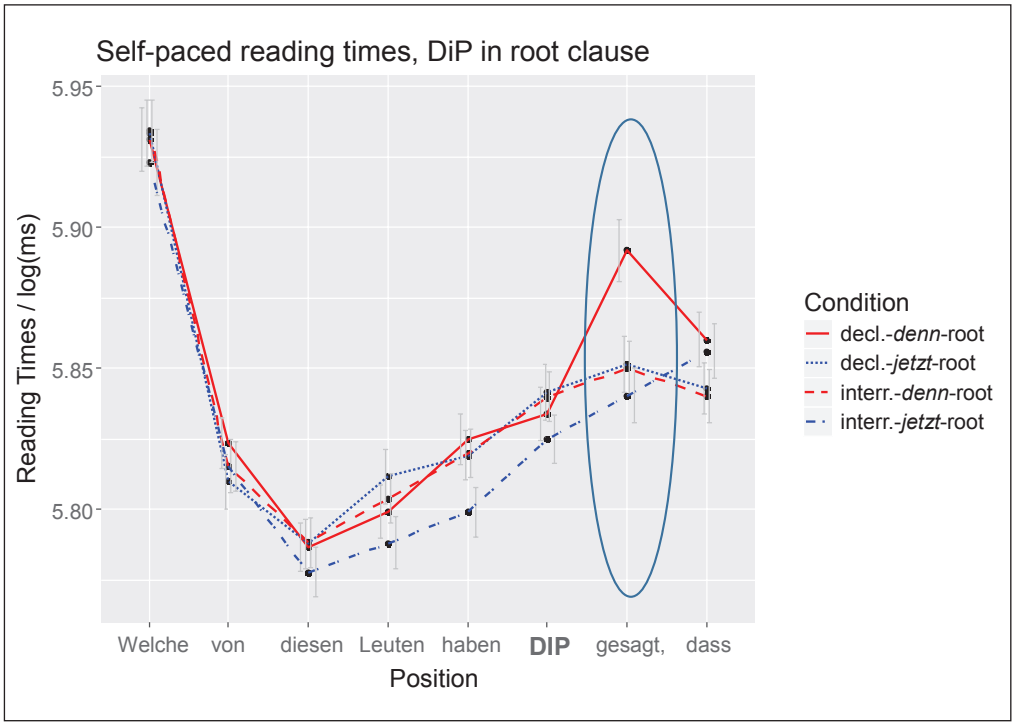

Figure 5 Mean log-transformed reading times over participants per position and condition, root clause conditions, Experiment 3. Error bars indicate standard error of the mean.

14 We chose to analyze this position because we assume that increases in processing cost are unlikely to surface on either *denn* and *jetzt*, given that both are short, highly frequent words, and highly repetitive in the context of the experiment.

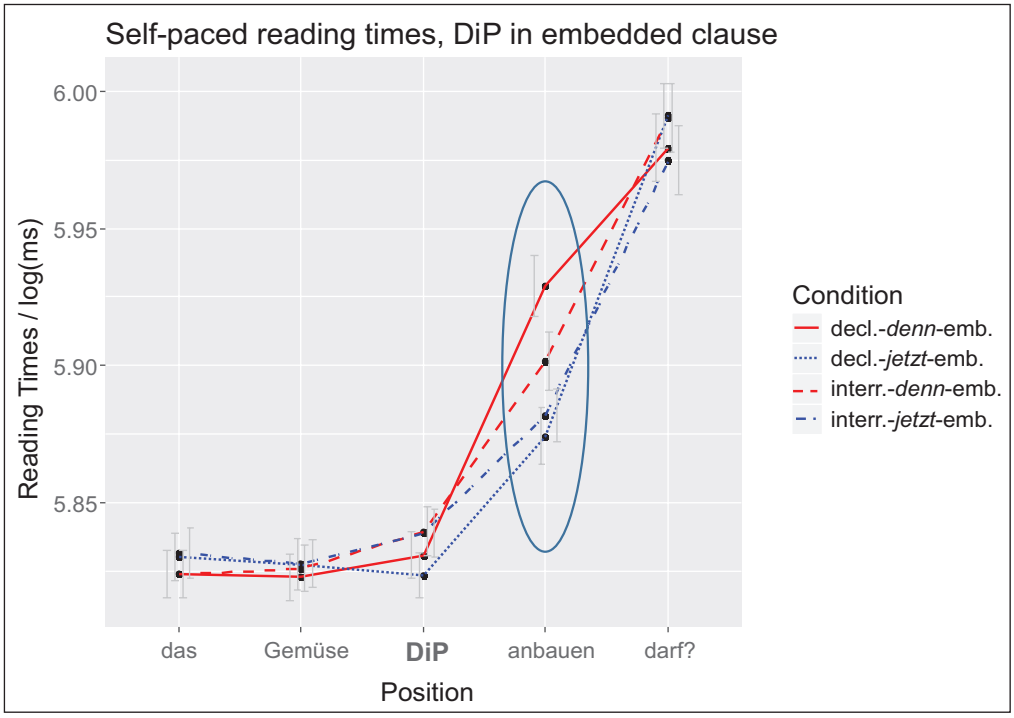

Figure 6 Mean log-transformed reading times over participants per position and condition, embedded clause conditions, Experiment 3. Error bars indicate standard error of the mean.

Reading times were analyzed for the position directly following the DiP. In root clauses it was the clause-final participle, in embedded clauses, it was the infinitive before the clause-final finite verb. We analyzed the differences between individual conditions with a series of linear mixed-effects models.[15] For Model 3.0, we specified the main effects and interactions of all three factors as fixed effects, and participant and item as random intercept. In addition, the main effects of DiP TYPE, POSITION and CLAUSE TYPE were specified as random slopes for participant. The joint analysis of all post-particle words revealed statistically significant main effects of DiP TYPE (t = –3.47, p < .001), POSITION (t = –2.20, p < .05) and CLAUSE TYPE (t = –2.13, p < .05), and a marginally significant interaction of DiP TYPE and CLAUSE TYPE (t = 1.93, p < .06). To pursue the interaction between DiP TYPE and CLAUSE TYPE, we analyzed the main effect of DiP TYPE separately for each clause type, and of CLAUSE TYPE separately for each DiP. For model 3.1, we specified the main effect of DiP TYPE as fixed effect, and participant and item as random intercepts. In addition, DiP TYPE was specified as random slope for participant. There was a statistically significant main effect of DiP TYPE for declaratives (t = –3.68, p < .001), but not for interrogatives (p > .3).

For model 3.2, we specified the main effect of CLAUSE TYPE as fixed effect, and participant and item as random intercepts. In addition, CLAUSE TYPE was specified as random slope for participant. There was a statistically significant main effect of CLAUSE TYPE for *denn* (t = –3.42, p < .001), but not for *jetzt* (t = –.08, p > .9).

5.3 DISCUSSION

The results of the first reading time study show longer processing time for *denn* in declaratives compared to all other conditions, i.e., there is a strong penalty for unlicensed QDiPs. The results for both conditions with *jetzt* show that this is not due to declaratives generally being associated with longer reading times than interrogatives. There is no visible increase in reading times for licensed QDiPs (i.e., interrogatives with *denn* in the root clause) compared to non-QDiPs, suggesting that any potential processing load associated with successful QDiP licensing is not strong enough to surface in the current experiment with comparatively easy structures.

Surprisingly, there is no three-way interaction of DiP TYPE, CLAUSE TYPE and DiP POSITION. This interaction could have been expected based on the theoretical assumptions outlined in

15 The random effects structure was chosen by beginning with a full random effects structure and then reducing gradually until the models would converge. Reduction began with random slopes for items, and continued with random slopes for participants. We then pursued the occurring interactions, striving to keep the same factors as fixed effects and random slopes. The models used in the analysis reported here are the most maximal models bearing straightforward reduction in the follow-up models pursueing interactions.

subsections 1.2 and 1.3. It would also have matched the findings of Experiments 1 and 2, showing that with short *wh*-extraction, *denn* is less acceptable in embedded clauses than in root clauses. However, this drop in acceptability was not reflected in longer reading times (while there is a descriptively visible difference between *denn* and *jetzt* in interrogatives, this difference did not reach statistical significance). We will return to this finding in the general discussion.

6 EXPERIMENT 4: SELF-PACED READING TIMES, SHORT VS. LONG EXTRACTION

The fourth experiment was a self-paced reading time study using stimulus set 2, described in section 2.2. The goal of this experiment was twofold. The first goal was to assess if the difference between licensing of QDiPs by *wh*-elements or by their traces is reflected in reading times. The second goal was to assess if there is a general increase in processing load associated with *successful* licensing of QDiPs relative to a non-QDiP baseline, reflecting the processing associated with the different checking procedures outlined in subsections 1.2 and 1.3.

6.1 MATERIALS AND METHODS

Participants 56 participants were invited via the SONA systems participant database of the University of Konstanz. All participants spoke German as their only native language. All participants had normal or corrected-to-normal vision, and reported no neurological or reading-related disorders. No participants were removed before the final data analysis. Participants were between 19 and 32 years old, their mean age was 24 years (s.d. = 3). 41 participants were female.

Language material The stimulus material was Stimulus set 2, outlined in detail in section 2.2. In total, each participant saw 196 sentences. 112 of these were critical sentences (14 per condition), 84 were filler sentences. Each participant saw two of the eight sentences from each item set; these sentences were different with respect to extraction (either short or long extraction) and to the DiP (either *denn* or *jetzt*).

After 50 of the sentences, participants were asked to answer a question. Before the presentation of the critical sentences, each participant completed a practice phase with six practice items.

Procedure The procedure was the same as described for Experiment 3, described in section 6.

Data preparation and analysis Data were prepared for analysis in the same way as described for Experiment 3. 1.8% of the data were removed as extreme values, 4.4% of the data were removed as outliers. The same statistical procedures and software as described for Experiment 3 were used for data analysis. For analysis, we defined the following factors: EXTRACTION (short or long); DiP TYPE (QDiP *denn* or non-QDiP *jetzt*); and POSITION (root or embedded).

6.2 RESULTS

Answer accuracies The mean error rate for the critical conditions was 38.5% (s.d. = 18.7%). The mean error rates for fillers were 12.8% (s.d. = 4.2%) for grammatical fillers, and 67.0% (S.D = 12.8%) for ungrammatical fillers. [16]

Reading times We analyzed self-paced reading times at the position directly following the DiP. Mean reading times per condition for the DiP and the two following words are given in ***Table 4***. A visualization of the mean log-transformed reading times is given in ***Figure 7*** for root clause conditions, and in ***Figure 8*** for embedded clause conditions.

Descriptively speaking, reading times for the position following the DiP were similar for all conditions in root clauses. In embedded clauses, reading times were longer for *denn* than for *jetzt* conditions, and slightly longer for conditions with short extraction than for those with long extraction.

16 The unusually high error rate for ungrammatical fillers seem to reflect an (unrelated) attraction effect in these conditions, which made the ungrammaticality difficult to detect in online processing. It does not signal low attention in our participants, and is no cause of concern for the processing of the critical sentences.

DIP IN ROOT CLAUSE			
CONDITION	**DIP**	***GESAGT***	***DASS***
short-*denn*	385 (82)	413 (105)	403 (95)
short-*jetzt*	389 (87)	416 (117)	405 (101)
long-*denn*	388 (91)	412 (124)	418 (144)
long-*jetzt*	395 (91)	419 (128)	438 (161)
DIP IN EMBEDDED CLAUSE			
CONDITION	**DIP**	***ABWEISEN***	***SOLL?***
short-*denn*	421 (100)	459 (149)	485 (193)
short-*jetzt*	418 (97)	436 (149)	474 (178)
long-*denn*	407 (97)	432 (116)	492 (205)
long-*jetzt*	406 (105)	419 (108)	478 (183)

Table 4 Mean reading times over participants per condition in ms, Experiment 4. Standard deviations are given in parentheses.

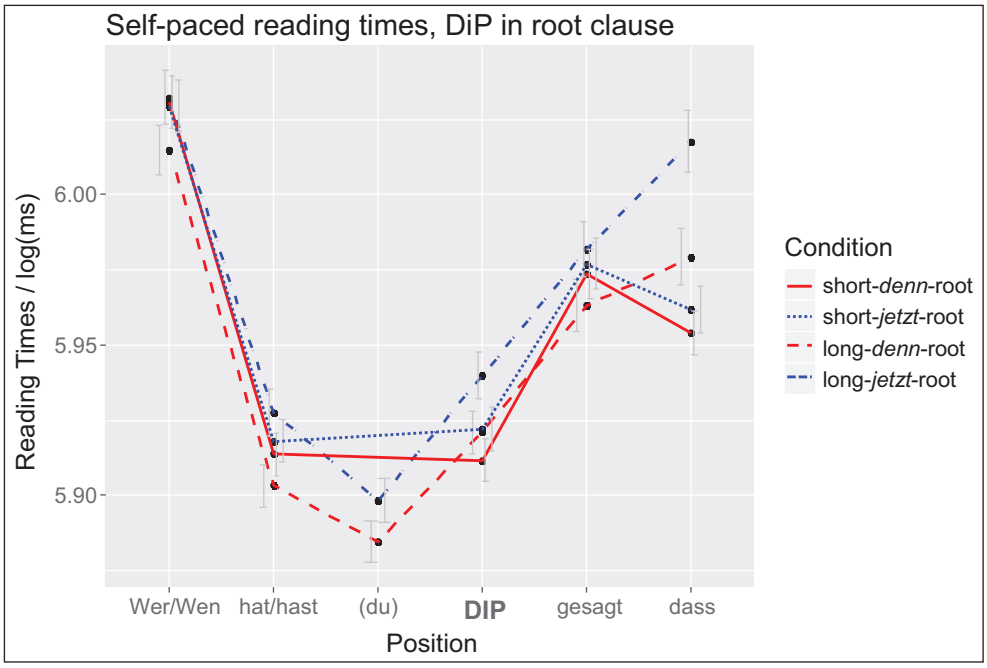

Figure 7 Mean log-transformed reading times per position and condition for root clause conditions, Experiment 4. Error bars indicate standard error of the mean.

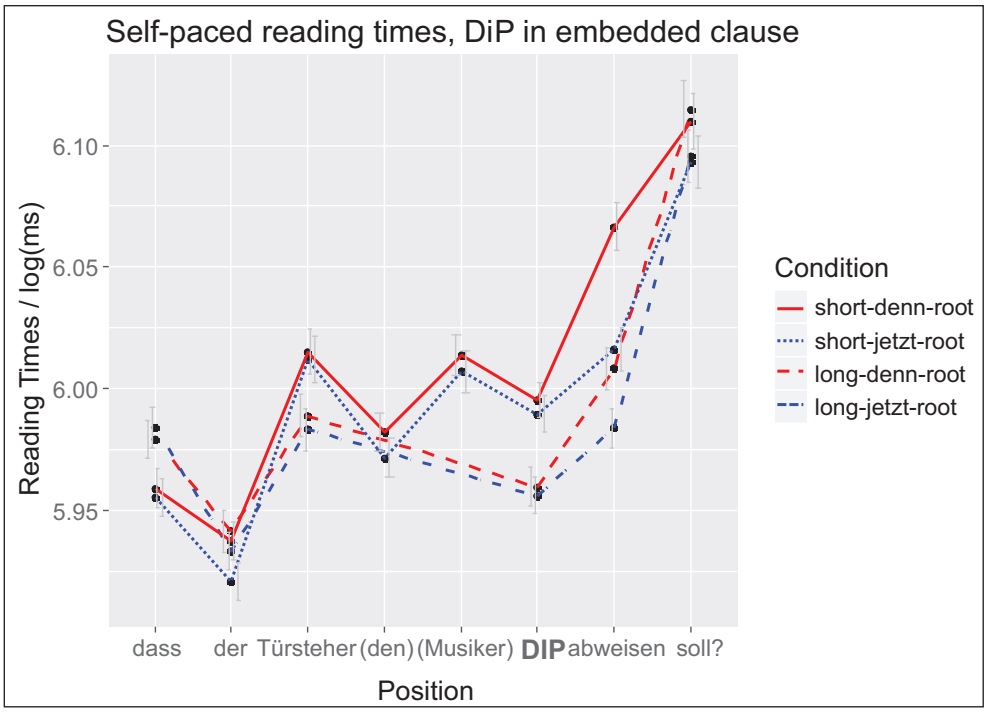

Figure 8 Mean log-transformed reading times per position and condition for embedded clause conditions, Experiment 4. Error bars indicate standard error of the mean.

We analyzed the differences between individual conditions and positions with a series of linear mixed-effects models. Full tables of the fixed effects for the models outlined below are given in the Appendix in Table 8. Random effects structures for all models were chosen in the same way as for Experiment 3.

Reading times for the post-particle verb from root clauses and embedded clauses were analyzed using Model 4.0. For this model, we specified the main effects and interactions of all three factors as fixed effects, and participant and item as random intercept. In addition, the main effects of DiP TYPE and EXTRACTION were specified as random slopes for participant. The analysis revealed statistically significant main effects of POSITION (t = –3.81, p < .001) and EXTRACTION (t = 4.68, p < .001) and a marginal effect of DiP TYPE (t = –1.94, p < .06). In addition, there was an interaction of POSITION and DiP TYPE (t = 2.55, p < .05), and of POSITION and EXTRACTION (t = –2.84, p < .01).

To pursue the interaction of POSITION and DiP TYPE, we analyzed the main effect of POSITION separately for each level of DiP TYPE, and the main effect of DiP TYPE separately for each level of POSITION. For model 4.1, we specified the main effect of DiP TYPE as fixed effect, and participant and item as random intercepts. In addition, EXTRACTION TYPE was specified as random slope for each participant. There was a statistically significant main effect of DiP TYPE for embedded clauses (t = –4.39, p < .001), but not for root clauses (t = 1.34, p > .1).

For model 4.2, we specified the main effect of POSITION as fixed effect, and participant and item as random intercepts. In addition, POSITION was specified as random slope for each participant. There was a statistically significant effect of POSITION for *denn* (t = –4.95, p < .001), but not for *jetzt* (t = –1.58, p > .12).

To pursue the interaction of POSITION and EXTRACTION, we analyzed the main effect of POSITION separately for each level of EXTRACTION, and the main effect of EXTRACTION separately for each level of POSITION. For model 4.3, we specified the main effect of EXTRACTION as fixed effect, and participant and item as random intercepts. In addition, EXTRACTION was specified as random slope for each participant. There was a statistically significant main effect of EXTRACTION in embedded clauses (t = 4.56, p < .001), but not in root clauses (t = –.19, p > .8).

For model 4.4, we specified the main effect of POSITION as fixed effect, and participant and item as random intercepts. In addition, POSITION was specified as random slope for each participant. There was a statistically significant main effect of POSITION on short extraction conditions (t = –4.33, p < .001), but not for long extraction conditions (t = –1.55, p > .1).

6.3 DISCUSSION

The second self-paced reading time study revealed no differences in reading times for the words following *denn* and *jetzt* in the root clauses, both for short and long extraction conditions. This is in line with the assumption that all of these conditions are grammatical, and is also in line with the findings of Experiment 3. In the embedded clause, reading times for conditions with *denn* are slightly longer than those for conditions with *jetzt*. Our general interpretation of this finding is that the increase in reading times for conditions with *denn* reflects the workload associated with checking the multiple licensing constraints outlined in subsection 1.2 and 1.3. This difference did not occur in root clause positions in the current experiment, and neither did it in Experiment 3. We assume that this increase in reading times is relatively subtle, and that it did not occur in conditions for which participants performed close to ceiling (i.e., short extraction conditions).

There is no statistically significant difference between embedded *denn* with short and long extraction. This latter finding is surprising given our theoretical assumptions outlined in subsection 1.2 and 1.3: Following these assumptions, embedded *denn* should be ungrammatical with short *wh*-extraction, and grammatical with long *wh*-extraction. This difference in grammaticality is not reflected in reading times. We will discuss possible reasons for this in the general discussion.

7 GENERAL DISCUSSION AND CONCLUSION

In general, the results of the four experiments match the theoretical assumptions outlined in subsections 1.2 and 1.3: The QDiP *denn* must be licensed – syntactically or semantically – by a

locally accessible Q-element. If the Q-element is absent (in declaratives) or out of reach of the QDiP (in interrogatives with short extraction and embedded QDiPs), sentences are rated as less acceptable than those with properly licensed QDiPs, or than parallel baseline sentences with non-QDiPs. In addition, there are some findings that may reflect an increased processing cost for successful QDiP licensing, and finally a surprising finding from both rating and self-paced studies.

QDiPs in declaratives Sentences with QDiPs in declaratives are rated as unacceptable in offline ratings, irrespective of QDiP position. In self-paced reading times, reading times are longer for the first position after the QDiP in declaratives, again independently of QDiP position. This indicates a severe violation that is detected quickly, in line with native speakers' intuitions and with the theoretical accounts outlined above.

Successful QDiP licensing According to the theoretical accounts in the introduction, QDiPs are successfully licensed in interrogatives in the root clause, and additionally in embedded clauses if the *wh*-element is extracted from the same embedded clause. Our findings from rating studies match these predictions, showing that QDiPs in root clauses and with long *wh*-extraction also in embedded clauses receive acceptability ratings that are comparable to the corresponding non-QDiP baselines. The results of Experiment 2 replicate earlier findings (Bayer et al. 2016), and are in line with the predictions from the theory outlined in the introduction. The results for long extraction conditions in particular (Experiment 2) support the relevance of intermediate traces in sentence processing in the syntactic approach (see Phillips & Parker 2014; Pickering & Barry 1991; Bayer et al. 2016 for discussions) and of reconstruction of *wh*-phrases to their base position in the semantic approach (Hamblin 1973; Rullmann & Beck 1998).

The results of Experiment 4 show increased online processing cost for embedded *denn* relative to *jetzt* with long *wh*-extraction. This increase in processing cost cannot be associated with licensing violations. Furthermore, it is unlikely that more basic differences between *denn* and *jetzt* are responsible – both are of similar length and highly frequent in German and the results of Experiment 2 do not suggest a marked difference in plausibility. We therefore assume that this increase in reading times must reflect some process related to successful QDiP licensing. We assume that this increase in processing cost did not become visible for QDiPs in root clause positions because the dependency between the *wh*-element and the QDiP is easier to parse, and participants were performing at ceiling. QDiP licensing involves (a) the checking of licensing constraints as outlined in the introduction, (b) the semantic computation not just of at-issue content but also of non-at-issue content[17], and finally (c) the integration of the meaning contribution of the QDiP with the previous discourse context. While (c) is unlikely to be a contributing factor in our experiment (which used single sentences without a context), (a) and (b) could be reflected in the longer reading times found for *denn* relative to *jetzt*. The possible processes included under (a) are too varied and numerous to allow for an in-depth discussion of their relation to different models of dependency processing here (but see our discussion on surprising findings below).

Explanation (b) (increased processing cost due to the semantic computation of non-at-issue content) would fit in with recent findings on the processing of German discourse particles (Dörre 2018; Dörre et al. 2018). In one of a series of self-paced reading experiments, Dörre et al. presented exactly identical sentences containing a German discourse particle. The sentences were felicitous both with the discourse particle and the counterpart reading; readings were distinguished via a preceding context. Reading times were longer on the immediate spillover region for discourse particle readings than for counterpart readings. This was attributed to the processing of non-at-issue content introduced by discourse particles, as opposed to the at-issue content introduced by the counterparts. Some caveats remain for simply adopting this explanation for our own data, the most important being that reading times for *denn* were not increased relative to its counterpart, but rather relative to a different, albeit controlled word *jetzt*. Still, it makes sense to assume that the processing of non-at-issue content contributes to the processing load for all well-formed conditions containing *denn*.

Taken together, while our findings provide an interesting first data point, more research is needed to replicate our finding and to systematically investigate the different possible explanations.

17 See also Dillon et al. (2014) on evidence that non-at-issue content in the sense of Potts 2005 is processed independently of at-issue content.

Surprising findings Consider the condition with short *wh*-extraction and the DiP situated in the embedded clause, predicted to be ungrammatical both by the syntactic and the semantic appraoch:

(25) Short extraction, embedded *denn*:
*Wer hat gemeint, dass der Türsteher den Musiker **denn** abweisen soll?
who has meant that the bouncer the musician QDiP away.turn should
'Who said that the bouncer should DENN turn away the musician?'

(26) Condition: short *wh*-extraction, embedded DiP:
*[$_{CP}$ WhP$_i$ Force^0 [$_{IP}$ ... t$_i$ ... [$_{CP}$ ... DiP ...]]].

There was a drop in acceptability ratings for this condition, both relative to QDiPs in root clauses and to the corresponding non-QDiP baseline condition. This finding is in line with the claims outlined in the introduction (subsections 1.2 and 1.3) that QDiP licensing does not work smoothly across clause boundaries. However, a closer look at the ratings revealed that the drop in acceptability ratings for these conditions was less dramatic than expected for a condition considered ungrammatical: Ratings did not drop to the level of ungrammatical sentences (QDiPs in declaratives) in Experiment 1, and did not even drop to the level of the dispreferred, but grammatical long *wh*-extraction conditions in Experiment 2.[18] It is surprising that a condition violating the syntactic and semantic licensing constraints outlined above should reliably be considered acceptable by our participants. Self-paced reading times associated with this condition were similar to those for grammatical conditions (non-QDiP baselines in Experiment 3, and embedded QDiPs with long *wh*-extraction in Experiment 4), indicating that the violations of the syntactic and semantic licensing constraints did not measurably affect online processing load. This is all the more surprising since the violation caused by QDiPs in declaratives had an immediate and marked effect.[19]

There are two possible lines of explanation for these unexpected findings. The first is that the unexpectedly good ratings for QDiPs with syntactically inaccessible licensers reflects an error in linguistic processing, i.e., a linguistic illusion. A comparable phenomenon would be intrusive licensing of negative polarity items (NPIs) like *ever*. This refers to the finding that NPIs with syntactically inaccessible licensers (as in *A man [who had **no** beard] was **ever** happy*) receive better than expected ratings, and are associated with no increases in processing cost or smaller ones than NPIs with completely absent licensers (as in *A man [who had a beard] was **ever** happy*) (Saddy et al. 2004; Drenhaus et al. 2005; Vasishth et al. 2008; Xiang et al. 2009; 2013; Yurchenko et al. 2013), even though both sentences are clearly ungrammatical. One point in favor of this explanation for our findings is that self-paced reading times reveal no increase in processing cost for embedded QDiPs with short extraction relative to grammatical conditions, while the rating studies revealed at least a drop relative to root clause QDiPs. This would suggest that the ungrammaticality of embedded QDiPs with out-of-reach licensers is more likely to go unnoticed in time-sensitive measures than in non-speeded offline rating studies (see Parker & Phillips 2016, Experiments 1 and 2, for a similar difference between speeded and non-speeded ratings for intrusive NPI licensing). However, our own rating studies do not seem to be completely immune to this, with surprisingly good ratings for ungrammatical conditions with embedded QDiPs.

There are differing accounts of intrusive NPI licensing, some focusing on errors during cue-based retrieval (Vasishth et al. 2008), others on an overapplication of pragmatic licensing processes (Xiang et al. 2009; Parker & Phillips 2016). Our currently available data from QDiP licensing would fit with both types of illusory licensing accounts. A cue-based retrieval approach

18 Interestingly, the results reported for the corresponding conditions in Bayer et al. (2016) show similar, surprisingly good ratings for embedded QDiPs with short extraction. This finding was not discussed in detail in the earlier study.

19 While interrogative-embedded-*denn* patterns like grammatical conditions in Experiment 3 and 4, there is nevertheless an interesting difference: For interrogative-embedded conditions, the contrast between *denn* and *jetzt* is significant in Experiment 4, but not in Experiment 3. Our tentative explanation is that this difference reflects an effect of stimulus context. In Experiment 4, interrogative-embedded-*denn* occurred together with acceptable conditions without licensing violations, which might have made its reduced acceptability stand out and led to more in-depth scrutiny. In Experiment 3, interrogative-embedded-*denn* occurred together with striking *denn*-licensing violations in declaratives, which might have made the reduced acceptability of the interrogative-embedded-*denn* conditions less prominent in comparison.

would explain the surprisingly good ratings for *denn* with an out-of-reach licenser as a partial match phenomenon (the root clause *wh*-element matches the +interrogative cue, but not the structural cues for QDiP licensers; see Vasishth et al. 2008 for intrusive NPI licensing). An account based on pragmatic licensing gone wrong could follow the proposal by Parker and Phillips for intrusive NPI licensing (Parker & Phillips 2016: 336). This would suggest that at the point the QDiP was encountered, the semantic/pragmatic representation of the QDiP licensing context was not yet complete, leading to an interpretation following simple heuristics rather than in-depth analysis. Another pragmatics-based approach could follow the one outlined in detail in Xiang et al. (2009: 53) for intrusive NPI licensing, assuming that spurious pragmatic inferences lead speakers to interpret the embedded clause as interrogative, thus licensing embedded QDiPs.[20]

Future studies will have to reveal to which extent a linguistic illusion contributes to the unexpectedly good ratings for QDiPs with locally inaccessible *wh*-licensers, and whether this illusion will be better explained with cue-based retrieval or pragmatic accounts of illusory licensing.

The second explanation for our unexpected findings is centered on the theoretical background for QDiP licensing. Our findings match the syntactic and semantic licensing properties and constraints outlined in the introduction, in that with short extraction, embedded QDiPs are rated as less acceptable than root clause QDiPs. However, these theoretical accounts alone would lead us to expect a drop in acceptability to the levels of QDiPs in declaratives. This was not the case. It is therefore possible that the theoretical accounts outlined in the introduction correctly model the syntactic and semantic contributions to QDiP licensing, but that there are additional licensing possibilities and constraints that are not captured by these accounts. This would suggest that the surprisingly good ratings for embedded QDiPs with short *wh*-extraction do not reflect errors in processing, but rather the result of an alternative licensing strategy. We give an outline of an account of pragmatic QDiP licensing below.

Pragmatic QDiP licensing Our surprising finding is, as discussed, that sentences with the structure (26) – with *wh*-short extraction and an embedded QDiP – are somewhat degraded but not completely ungrammatical. The syntactic and semantic approaches presented in sections 1.2 and 1.3 predict straightforward ungrammaticality for such sentences. The question arises whether any of those approaches can be "relaxed" so as to explain the intermediate status of sentences with this configuration.

In fact, independently of the intermediate status of these sentences, the syntactic and semantic approaches sketched above have two limitations.

The first limitation is that these approaches are too weak, in the sense that, by themselves, they fail to rule out infelicitous occurrences of *denn*. This is because these accounts were meant only as part of the story – i.e., as restrictions on the syntactic or semantic *composition* – to which the exact meaning contribution of *denn* still has to be added – as a *lexical* restriction. Two kinds of lexical restrictions have been tracked in the literature.

On the one hand, as mentioned in the introduction, *denn* has been argued to flag a dependency between the *denn*-sentence and the previous context: it explicitly signals that the *denn*-utterance is motivated by the previous interactional context (König 1977) or has special relevance given that context (Bayer 2012). This lexical condition rules out *denn* in out-of-the-blue questions like (27) (see König 1977; Thurmair 1991; Wegener 2002; Grosz 2005; Bayer 2012):

(27) [S wakes H up in the middle of the night.]
S: Wie spät ist es (#denn)?
How late is it QDiP
S: 'How late is it?'

On the other hand, *denn* has been argued to lexically require that some attitude holder – the speaker or the subject of an attitude verb – be in a "want-to-know" relation with the *denn*-clause, as formulated in (28) (Rapp 2018). This accounts for the contrast between examples (29a)–(29b) – where Christine can easily be assumed to be in a "want-to-know" relation with the embedded question – and example (29c) – where she cannot. Note that, if this lexical

20 Note that before an approach along these lines can be adopted or rejected, the pragmatic licensing conditions for QDiPs will need to be spelled out in greater detail than before; see the following paragraph on pragmatic QDiP licensing.

DOI: 10.5334/gjgl.1203

condition were not factored in, the syntactic and semantic approaches above would wrongly predict all the examples in (29) to be equally acceptable, since the relevant syntactic *u*Q feature is checked and an appropriate $Q_{<<s,t>,t>}$ is found in the embedded clause:

(28) Lexical pre-condition introduced by *denn* according to Rapp (2018):
A *denn*-clause *[denn p]* is felicitous only if there is an attitude holder that wants to know whether or to what extent *p* applies.

(29) a. Christine fragt, warum Klaus denn so blass ist.
Christine asks why Klaus QDiP so pale is
'Christine asks why Klaus is DENN so pale.'

b. Christine weiß nicht, warum Klaus denn so blass ist.
Christine knows not why Klaus QDiP so pale is
'Christine doesn't know why Klaus is DENN so pale.'

c. *Christine weiß, warum Klaus denn so blass ist.
Christine knows why Klaus QDiP so pale is
'Christine knows why Klaus is DENN so pale.'

The second limitation of the syntactic and semantic analyses above is that they are too strong in that they wrongly rule out naturally occurring examples like (30), noted in Bayer et al. (2016), and (31), from Fortmann (2017). These examples show that, in certain contexts, embedded QDiPs can be perfectly acceptable even if there is no long *wh*-extraction:

(30) Glaubst du, dass dieser Mann denn ernsthaft eine Beziehung führen möchte?
Believe you that this man QDiP seriously a relationship wants conduct
'Do you think that this man DENN seriously wants to be in a relationship?'

(31) (...) ob Adorno (...) glaube, dass man denn wirklich alle Menschen belasten
(...) if Adorno (...) believe that one QDiP really all people bother
müsse mit Grundsatzproblematik und Reflexionsaufwand, (...)
should with fundamental.problems and reflection.trouble (...)
'(...) whether Adorno thinks that one DENN really should bother everybody with fundamental problems and reflection (...)'

In (30), *denn* intuitively takes as complement the *wh*-question 'To what extent/How seriously does this man want to have a relationship?' or the polar question 'Does this man seriously want to have a relationship?', even though there is no *wh*-element base-generated in the embedded clause. Similarly, in (31), *denn* intuitively takes as complement the question 'Must one really bother everybody with fundamental problems and reflection?' even though there is no *wh*-element in the embedded clause. Interestingly, what we do find in both embedded clauses is some element in focus: *ernsthaft* 'seriously' in (30) and *wirklich alle* 'really everybody' in (31).

To account for these naturally occurring data, Romero (2017) "relaxes" her semantic analysis into a pragmatic analysis, roughly as follows. As before, *denn* needs to combine with a question meaning (type <<s,t>,t>). This question meaning may be provided directly by the semantics of its syntactic sister, as before, or *indirectly* by the *pragmatics* of its sister, namely by the Question-Under-Discussion (QUD) indicated by the focus structure of its sister. To see the intuition behind the latter possibility, note that focused elements invoke alternatives (formally, via their focus value $[\![.]\!]^f$), as in (32), that these focus alternatives percolate up the sentence, as in (33), and that the focus alternatives of a declarative clause must match the QUD that it serves as answer to (Rooth 1992; Krifka 2008), as in (34). Thus, in a natural and rich enough context, focus will enable the hearer to retrieve the corresponding QUD from the context and to use it as the $Q_{<<s,t>,t>}$-argument of *denn*.

(32) $[\![\text{seriously}_F]\!]^f$ = {'seriously', 'tentatively', 'jokingly',...}

(33) $[\![\text{this man seriously}_F \text{ wants a relationship}]\!]^f$ = {'that this man seriously wants a relationship', 'that this man tentatively wants a relationship', 'that this man jokingly wants a relationship',...}

DOI: 10.5334/gjgl.1203

(34) QUD: 'To what extent does this man want a relationship?'
Answer: This man seriously$_F$ wants a relationship.

Hence, while some sentences provide the desired <<s,t>,t>-argument of *denn* directly via a *wh*-phrase, sentences like (30)–(31) provide the <<s,t>,t>-argument indirectly via focus. Either way, *denn* combines with this $Q_{<<s,t>,t>}$-argument and explicity predicates of it that is is motivated by the previous interactional context (à la König 1977).

We are now ready to return to our surprising finding on the intermediate acceptability of structure (35). We will see that, now that the pragmatic account is in place to allow for naturally occurring data like (30)–(31), it can be combined with the lexical conditions – most importantly, with Rapp's (28) – to explain the intermediate status of (35) and the overall pattern found in our experiments.

(35) Condition: short *wh*-extraction, embedded DiP:
$[_{CP}$ WhP$_i$ Force^0 $[_{IP}$ ... t$_i$ ... $[_{CP}$ ... DiP ...]]]. (= (26))

We start with the perfectly acceptable sentences (29a)–(29b) and (36)–(37). In all these sentences, the *Q*-argument is directly provided via the *wh*-phrase in the semantic derivation; hence, there is no need for a natural rich-context setting to retrieve it via the pragmatics of focus. Furthermore, it is easy to assume Rapp's "want-to-know" relation between an attitude holder – the speaker in (36)–(37) and Christine in (29a)–(29b) – and the *Q*-argument. Since both the compositional and the lexical restrictions are fulfilled, the sentences are fully acceptable.

(36) Wen hat Max denn gesehen?
whom hat Max QDiP seen
'Who did Max see?'

(37) Wie denkst du, dass es denn weitergehen soll mit euch?
how think you that it QDiP on-go should with you
'How do you think that the two of you should carry on?' (=8)

At the opposite end, we have clearly unacceptable sentences like (38) and (29c). Here Rapp's lexical restriction is not satisfied: a "want-to-know" relation can be established neither with the speaker nor with other attitude holder. Thus, regardless of whether the *Q*-argument is provided semantically via a *wh*-phrase – as in (29c) – or it should be pragmatically retrieved from a natural rich-context setting via focus – as in (38) with some focus –, the lexical restriction is not satisfied and, thus, the sentences are unacceptable.

(38) *Max hat Anne denn gesehen.
Max has Anne QDiP seen.
Max saw Anne. (=19a)

Finally, we come to the surprising, partially degraded short extraction sentences with embedded *denn* in our experimental studies. Note that these sentence items are completely parallel to the naturally occurring data (30)–(31). Both instantiate the structure (35): the DiP is located in the embedded clause, and there is no *wh*-chain stemming from the embedded clause but only a *wh*-chain stemming from the matrix clause. Furthermore, in both cases it is easy to assume Rapp's "want-to-know" relation between the speaker and the *Q*-argument. However, while the naturally occurring data are judged as perfectly acceptable in their context of utterance, the experimental items are judged as somewhat degraded. Where does the difference lie? The difference, we suggest, lies in the lack vs. presence of sufficient elements to identify the intended QUD-meaning. While the presence of *ernsthaft* 'seriously'/*wirklich* 'really' and, importantly, the natural rich context converge on a given QUD in (30)–(31), our experimental setting provides no clearly focused items and, crucially, no previous context that could help the reader identify the intended QUD-meaning that *denn* is supposed to operate on. Since the pragmatic search for the intended QUD-meaning fails, the sentences are judged as somewhat degraded for pragmatic reasons. This contrasts with the hard violation of Rapp's lexical condition, which we saw unavoidably leads to unacceptability.

Conclusion and outlook The results of our studies replicate and extend earlier findings (Bayer et al. 2016), showing that the question-dependent German discourse particle (QDiP) *denn* needs to be licensed (syntactically by a [+Q] head or semantically via types): In offline ratings, immediate licensing violations conditioned by the absence of interrogativity lead to a drop in acceptability. In addition, self-paced reading times show longer reading times for such outright licensing violations, and a mild increase in processing cost for successful QDiP licensing relative to non-QDiPs. To our knowledge, we are the first to report online measures for the licensing of QDiPs, and the first to find correlates of successful QDiP licensing. These findings match the predictions from syntactic and semantic accounts of QDiP licensing as outlined in the introduction.

However, some of our results do *not* match the expectations of current theoretical approaches: *denn* in interrogatives violating syntactic locality and delivering a semantic type mismatch lead to surprisingly small drops in acceptability ratings, and no marked increases in reading times. These findings open at least two possible future lines of research. One is to pursue the contribution and nature of a potential linguistic illusion (along the lines of intrusive NPI licensing). Another one is to pursue alternatives to syntactic/semantic QDiP licensing via pragmatics, possibly along the lines sketched in the general discussion (see above).

In the broader picture, we have seen that a phenomenon whose effect is ultimately pragmatic in nature – namely the QDiP's re-shaping or fine-tuning of the illocutionary act in uttering the interrogative clause – is nevertheless highly constrained by grammar-internal mechanisms: by feature agreement in syntactic structure, by meaning composition via *wh*-alternatives or, as suggested in view of unexpected results, possibly by meaning composition via focus alternatives. This, in turn, means that the online processing of QDiPs involves not just pragmatic factors but also the monitoring of syntactic dependencies and/or of meaning construction. Taken together, our studies offer new insights into licensing processes at the interfaces of syntax, semantics and pragmatics.

ABBREVIATIONS

DiP = discourse particle, QDiP = question-sensitive discourse particle, NPI = negative polarity item, ACC = accusative, ANIM = animate

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Appendix.** Fixed effects tables for statistical analyses. DOI: *https://doi.org/10.5334/gjgl.1203.s1*

ACKNOWLEDGEMENTS

The authors would like to thank Yvonne Viesel and Laura Dörre for helpful discussions on modal particles, and Andrea Beltrama for helpful discussions on NPI licensing. They gratefully acknowledge the assistance of Christina Gozebina, Alexandra Engel, Charlotte Englert, Marc Meisezahl and Uwe Braun with stimulus construction and the running of experiments. We also thank the audience of the workshop 'Doing experiments with theoretical linguistics' (April 4th 2019, University of Amsterdam) for valuable feedback.

FUNDING INFORMATION

We thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for funding within project BA 1178/10-2 *The Role of German Particles in Questions* of the FOR2111 "Questions at the Interfaces".

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Anna Czypionka *orcid.org/0000-0003-2247-0110*
University of Konstanz, Konstanz, DE
Maribel Romero *orcid.org/0000-0001-9776-1479*
University of Konstanz, Konstanz, DE
Josef Bayer *orcid.org/0000-0002-6445-5333*
University of Konstanz, Konstanz, DE

REFERENCES

Andersson, Sven-Gunnar & Sigmund Kvam. 1984. *Satzverschränkung im heutigen Deutsch.* Tübingen: Narr.

Aoshima, Sachiko, Colin Phillips & Amy Weinberg. 2004. Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51(1). 23–54. DOI: *https://doi.org/10.1016/j.jml.2004.03.001*

Bader, Markus. 2012. The German *bekommen* passive: A case study on frequency and grammaticality. *Linguistische Berichte* 231.

Baker, C. Lee. 1970. Double negatives. *Linguistic Inquiry* 1(2). 169–186.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixede effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: *https://doi.org/10.18637/jss.v067.i01*

Bayer, Josef. 1991. German particles in a modular grammar: Neurolinguistic evidence. In Werner Abraham (ed.), *Discourse particles: Descriptive and theoretical investigations on the logical, syntactic and pragmatic properties of discourse particles in German,* 253–302. Amsterdam: John Benjamins. DOI: *https://doi.org/10.1075/pbns.12.09bay*

Bayer, Josef. 1996. *Directionality and logical form: On the scope of focusing particles and wh-in-situ.* Dordrecht: Kluwer.

Bayer, Josef. 1999. Bound focus in German or how can association with focus be achieved without going semantically astray? In Georges Rebushi & Laurice Tuller (eds.), *The grammar of focus,* 55–82. Amsterdam: John Benjamins. DOI: *https://doi.org/10.1075/la.24.03bay*

Bayer, Josef. 2012. From modal particle to interrogative marker: A study of *German denn.* In Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro & Cecilia Poletto (eds.), *Functional heads (The cartography of syntactic structures)*, vol. 7, 13–28. New York: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:oso/9780199746736.003.0001*

Bayer, Josef. 2018. Criterial freezing in the syntax of particles. In Jutta Hartmann, Marion Jäger, Andreas Kehl, Andreas Konietzko & Susanne Winkler (eds.), *Freezing: Theoretical approaches and empirical domains,* 225–263. Berlin & New York: Mouton de Gruyter. DOI: *https://doi.org/10.1515/9781501504266-007*

Bayer, Josef, Jana Häussler & Markus Bader. 2016. A new diagnostic for cyclic *wh*-movement: Discourse particles in German questions. *Linguistic Inquiry* 47(4). 591–629. DOI: *https://doi.org/10.1162/LING_a_00224*

Bayer, Josef & Hans-Georg Obenauer. 2011. Discourse particles, clause structure, and question types. *The Linguistic Review* 28(4). 449–491. DOI: *https://doi.org/10.1515/tlir.2011.013*

Chierchia, Gennaro. 2006. Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry* 37(4). 535–590. DOI: *https://doi.org/10.1162/ling.2006.37.4.535*

Csipak, Eva & Sarah Zobel. 2016. Discourse particle denn in the antecedent of conditionals. In Christopher Piñón (ed.), *Empirical Issues in Syntax and Semantics 11*, 31–60. Paris: CSSP.

Detges, Ulrich. 2007. Altfranzösisch *or* und deutsch *jetzt* in Fragen [Old French or and German *jetzt* in questions]. In Elisabeth Stark, Roland Schmidt-Riese & Eva Stoll (eds.), *Romanische Syntax im Vergleich,* 417–436. Tübingen: Narr.

Dillon, Brian, Charles Clifton Jr. & Lyn Frazier. 2014. Pushed aside: Parentheticals, memory and processing. *Language, Cognition and Neuroscience* 29(4). 483–498. DOI: *https://doi.org/10.1080/01690965.2013.866684*

Dimroth, Christine, Cecilia Andorno, Sandra Benazzo & Josje Verhagen. 2010. Given claims about new topics. How Romance and Germanic speakers link changed and maintained information in narrative discourse. *Journal of Pragmatics* 42(12). 3328–3344. DOI: *https://doi.org/10.1016/j.pragma.2010.05.009*

Dörre, Laura. 2018. *Primary and secondary meaning in the processing of German particles.* University of Konstanz doctoral dissertation.

Drenhaus, Heiner, Douglas Saddy & Stefan Frisch. 2005. Processing negative polarity items: When negation comes through the backdoor. *Linguistic evidence: Empirical, theoretical, and computational perspectives* 145–165. DOI: *https://doi.org/10.1515/9783110197549.145*

Dörre, Laura, Anna Czypionka, Andreas Trotzke & Josef Bayer. 2018. The processing of German modal particles and their counterparts. *Linguistische Berichte* 255. 58–91.

Fanselow, Gisbert, Reinhold Kliegl & Matthias Schlesewsky. 2005. Syntactic variation in German *wh*-questions: Empirical investigations of weak crossover violations and long *wh*-movement. *Linguistic Variation Yearbook* 5(1). 37–63. DOI: *https://doi.org/10.1075/livy.5.03fan*

Felser, Claudia, Harald Clahsen & Thomas F. Münte. 2003. Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and *wh*-movement in German. *Brain and Language* 87(3). 345–354. DOI: *https://doi.org/10.1016/S0093-934X(03)00135-4*

Fiebach, Christian J., Matthias Schlesewsky & Angela D. Friederici. 2001. Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research* 30(3). 321–338. DOI: *https://doi.org/10.1023/A:1010447102554*

Fortmann, Jette. 2017. Modalpartikeln in Nebensätzen. M.A. thesis, University of Konstanz.

Fox, John & Sanford Weisberg. 2011. *An R companion to applied regression*. Thousand Oaks, CA: Sage. 2nd edn. *http://socserv.socsci.mcmaster.ca/jfox/Books/Companion*.

Frazier, Lyn & Giovanni B. Flores d'Arcais. 1989. Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language* 28. 331–344. DOI: *https://doi.org/10.1016/0749-596X(89)90037-5*

Giannakidou, Anastasia. 2006. *Only*, emotive factive verbs, and the dual nature of polarity dependency. *Language,* 575–603. DOI: *https://doi.org/10.1353/lan.2006.0136*

Giannakidou, Anastasia. 2011. Negative and positive polarity items: Variation, licensing, and compositionality. *Semantics: An International Handbook of Natural Language Meaning* 3.

Gotzner, Nicole & Katharina Spalek. 2017. Role of contrastive and noncontrastive associates in the interpretation of focus particles. *Discourse Processes* 54(8). 638–654. DOI: *https://doi.org/10.1080/0163853X.2016.1148981*

Gotzner, Nicole & Katharina Spalek. 2019. The life and times of focus alternatives: Tracing the activation of alternatives to a focused constituent in language comprehension. *Language and Linguistics Compass* 13(2). e12310. DOI: *https://doi.org/10.1111/lnc3.12310*

Grosz, Patrick. 2005. *"Dn" in Viennese German. The syntax of a clitic version of the discourse particle "denn"*. University of Vienna Diplomarbeit.

Guerzoni, Elena & Yael Sharvit. 2014. *Whether or not anything but not Whether anything or not*. In L. Crnic & U. Sauerland (eds.), *The art and craft of semantics: A Festschrift for Irene Heim,* 199–224. Cambridge, MA: MITWPL.

Gutzmann, Daniel. 2015. Continuation-based semantics for modal particles. Deriving syntax from semantics. *MIT Working Papers in Linguistics* 75. 133–150.

Hamblin, Charles L. 1973. Questions in Montague grammar. *Foundations of Language* 10(1). 41–53.

Han, Chung-hye & Maribel Romero. 2004. The syntax of *Whether/Q… Or* questions: Ellipsis combined with movement. *Natural Language and Linguistic Theory* 22(3). 527–564. DOI: *https://doi.org/10.1023/B:NALA.0000027674.87552.71*

Heister, Julian, Kay-Michael Würzner, Johannes Bubenzer, Edmund Pohl, Thomas Hanneforth & Alexander Geyken. 2011. dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung [dlexdb – a lexical data base for psychological and linguistic research]. *Psychologische Rundschau* 32(4). 10–20. DOI: *https://doi.org/10.1026/0033-3042/a000029*

Hope, Ryan. 2013. Rmisc: Ryan miscellaneous. *https://cran.r-project.org/package=Rmisc*.

Horn, Laurence R. 2010. *The expression of negation.* Berlin & New York: Walter de Gruyter. DOI: *https://doi.org/10.1515/9783110219302.39*

Israel, Michael. 2004. The pragmatics of polarity. *The handbook of pragmatics* 701–723. DOI: *https://doi.org/10.1002/9780470756959.ch31*

Kadmon, Nirit & Fred Landman. 1993. Any. *Linguistics and Philosophy* 16(4). 353–422. DOI: *https://doi.org/10.1007/BF00985272*

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25(3–4). 209–257. DOI: *https://doi.org/10.3765/salt.v4i0.2462*

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3–4). 243–276. DOI: *https://doi.org/10.1556/ALing.55.2008.3-4.2*

König, Ekkehard. 1977. Modalpartikeln in Fragesätzen. In Harald Weydt (ed.), *Aspekte der Modalpartikeln*, 115–130. Tübingen: Niemeyer.

Linebarger, Marcia C. 1987. Negative polarity and grammatical representation. *Linguistics and Philosophy* 10(3). 325–387. DOI: *https://doi.org/10.1007/BF00584131*

Liu, Mingya, Peter König & Jutta L. Mueller. 2019. Novel ERP evidence for processing differences between negative and positive polarity items in German. *Frontiers in Psychology* 10. 376. DOI: *https://doi.org/10.3389/fpsyg.2019.00376*

Munaro, Nicola & Cecilia Poletto. 2004. Sentential particles and clausal typing in the Veneto Dialects. In Benjamin Shaer, Werner Frey & Claudia Maienborn (eds.), *Proceedings of the Dislocated Elements workshop*, vol. 35 (ZAS Papers in Linguistics 2), 375–397. Berlin: ZAS.

Onea, Edgar & Markus Steinbach. 2012. Where question, conditionals and topics converge. In *Selected Papers of the Amsterdam Colloquium* 2011, 42–51. Berlin, Heidelberg: Springer. DOI: *https://doi.org/10.1007/978-3-642-31482-7_5*

Parker, Dan & Colin Phillips. 2016. Negative polarity illusions and the format of hierarchical encodings in memory. *Cognition* 157. 321–339. DOI: *https://doi.org/10.1016/j.cognition.2016.08.016*

Paul, Hermann. 1920. *Deutsche Grammatik vol. IV*. Halle: Niemeyer.

Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation,* 262–294. Amsterdam: John Benjamins. DOI: *https://doi.org/10.1075/la.101.14pes*

Petrova, Svetlana. 2017. On the status and the interpretation of the left-peripheral sentence particles *inu* and *ia* in Old High German. In Josef Bayer & Volker Struckmeier (eds.), *Discourse particles: Formal approaches to their syntax and semantics* (Linguistische Arbeiten 564), 304–331. Berlin: Mouton de Gruyter. DOI: *https://doi.org/10.1515/9783110497151-012*

Phillips, Colin, Nina Kazanina & Shani H. Abada. 2005. ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22(3). 407–428. DOI: *https://doi.org/10.1016/j.cogbrainres.2004.09.012*

Phillips, Colin & Dan Parker. 2014. The psycholinguistics of ellipsis. *Lingua* 151. 78–95. DOI: *https://doi.org/10.1016/j.lingua.2013.10.003*

Pickering, Martin & Guy Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes* 6(3). 229–259. DOI: *https://doi.org/10.1080/01690969108406944*

Potts, Christopher. 2005. *The logic of conventional implicatures.* Oxford University Press on Demand. DOI: *https://doi.org/10.1093/acprof:oso/9780199273829.001.0001*

R Development Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria. *http://www.Rproject.org*.

Rapp, Irene. 2018. Differential object marking: Iconicity vs. economy. *Linguistische Berichte* 254. 183–228.

Rawlins, Kyle. 2008. *(Un)conditionals: An investigation in the syntax and semantics of conditional structures:* UCSC doctoral dissertation.

Rohde, Douglas. 2003. *Linger: A exible platform for language processing experiments, version 2.94*.

Romero, Maribel. 2015. High negation in subjunctive conditionals and polar questions. In *Proceedings of Sinn und Bedeutung* 19. 519–536.

Romero, Maribel. 2017. German discourse particles in questions. Invited talk at the workshop "Inquisitiveness below and beyond the sentence boundary" (InqBnB2). December 18th–19th, University of Amsterdam.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116.

Rullmann, Hotze & Sigrid Beck. 1998. Presupposition projection and the interpretation of *which* questions. In D. Strolovitch & A. Lawson (eds.), *Proceedings of Semantics and Linguistic Theory VIII*. 215–232. Ithaca, NY: Cornell University. DOI: *https://doi.org/10.3765/salt.v8i0.2811*

Saddy, Douglas, Heiner Drenhaus & Stefan Frisch. 2004. Processing polarity items: Contrastive licensing costs. *Brain and Language* 90. 495–502. DOI: *https://doi.org/10.1016/S0093-934X(03)00470-X*

Spalek, Katharina & Amir Zeldes. 2017. Converging evidence for the relevance of alternative sets: Data from NPs with focus sensitive particles in German. *Language and Cognition* 9(1). 24–51. *https://doi.org/10.1017/langcog.2015.12*.

Stowe, Laurie. 1986. Parsing *Wh*-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1. 227–245. DOI: *https://doi.org/10.1080/01690968608407062*

Struckmeier, Volker. 2014. 'Ja doch wohl' C? Modal particles in German as C-related elements. *Studia Linguistica* 68(1). 1–48. DOI: *https://doi.org/10.1111/stul.12019*

Theiler, Nadine. 2020. *Denn* as a highlighting-sensitive particle. *Linguistics and Philosophy,* 1–40. DOI: *https://doi.org/10.1007/s10988-019-09290-7*

Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen.* Tübingen: Niemeyer. DOI: *https://doi.org/10.1515/9783111354569*

Thurmair, Maria. 1991. Zum Gebrauch der Modalpartikel 'denn' in Fragesätzen. Eine korpusbasierte Untersuchung. In Eberhard Klein, Françoise Pouradier Duteil & Karl Heinz Wagner (eds.), *Betriebslinguistik und Linguistikbetrieb. Akten des 24. Ling. Koll. Bremen, September* 1989. 377–387. Tübingen: Max Niemeyer.

Tremblay, Antoine & Johannes Ransijn. 2015. *LMERConvenienceFunctions: Model selection and post-hoc analysis for (g)lmer models. https://CRAN.R-project.org/package=LMERConvenienceFunctions*. R package version 2.10.

Turco, Giuseppina, Bettina Braun & Christine Dimroth. 2014. When contrasting polarity, the Dutch use particles, Germans intonation. *Journal of Pragmatics* 62. 94–106. DOI: *https://doi.org/10.1016/j.*pragma.2013.09.020

Vasishth, Shravan, Sven Brüssow, Richard L. Lewis & Heiner Drenhaus. 2008. Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science* 32(4). 685–712. DOI: *https://doi.org/10.1080/03640210802066865*

Wagers, Matthew W. & Colin Phillips. 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics* 45(2). 395–433. DOI: *https://doi.org/10.1017/S0022226709005726*

Wegener, Heide. 2002. The evolution of the German modal particle 'denn'. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization,* 379–393. Amsterdam: Benjamins. DOI: *https://doi.org/10.1075/tsl.49.24weg*

Wickham, Hadley. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* 21(12). *http://www.jstatsoft.org/v21/i12/paper*. DOI: *https://doi.org/10.18637/jss.v021.i12*

Wickham, Hadley. 2009. *ggplot2: Elegant graphics for data analysis*. New York: Springer. *http://ggplot2.org*. DOI: *https://doi.org/10.1007/978-0-387-98141-3*

Wickham, Hadley. 2011. The split-apply-combine strategy for data analysis. *Journal of Statistical Software* 40(1). 1–29. *http://www.jstatsoft.org/v40/i01/*. DOI: *https://doi.org/10.18637/jss.v040.i01*

Xiang, Ming, Brian Dillon & Colin Phillips. 2009. Illusory licensing effects across dependency types: ERP evidence. *Brain and Language* 108(1). 40–55. DOI: *https://doi.org/10.1016/j.bandl.2008.10.002*

Xiang, Ming, Brian Dillon, Matt Wagers, Fengqin Liu & Taomei Guo. 2014. Processing covert dependencies: An SAT study on Mandarin *wh*-in-situ questions. *Journal of East Asian Linguistics* 23(2). 207–232. DOI: *https://doi.org/10.1007/s10831-013-9115-1*

Xiang, Ming, Julian Grove & Anastasia Giannakidou. 2013. Dependency-dependent interference: NPI interference, agreement attraction, and global pragmatic inferences. *Frontiers in Psychology* 4. 708. DOI: *https://doi.org/10.3389/fpsyg.2013.00708*

Xiang, Ming, SuiPing Wang & YanLing Cui. 2015. Constructing covert dependencies – The case of Mandarin *wh*-in-situ dependency. *Journal of Memory and Language* 84. 139–166. DOI: *https://doi.org/10.1016/j.jml.2015.05.006*

Yurchenko, Anna, Dirk-Bart Den Ouden, Jack Hoeksema, Olga Dragoy, John C. J. Hoeks & Laurie A. Stowe. 2013. Processing polarity: ERP evidence for differences between positive and negative polarity. *Neuropsychologia* 51(1). 132–141. DOI: *https://doi.org/10.1016/j.neuropsychologia.2012.10.028*

TO CITE THIS ARTICLE:
Czypionka, Anna, Maribel Romero and Josef Bayer. 2021. Question-sensitive discourse particles at the interfaces of syntax, semantics and pragmatics – an experimental approach. *Glossa: a journal of general linguistics* 6(1): 24. 1–34. DOI: *https://doi.org/10.5334/gjgl.1203*

Submitted: 30 January 2020
Accepted: 12 October 2020
Published: 22 February 2021

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.