Mathias Barthel*, Rosario Tomasello and Mingya Liu

Online comprehension of conditionals in context: A self-paced reading study on *wenn* ('if') versus *nur wenn* ('only if') in German

https://doi.org/10.1515/lingvan-2021-0083
Received May 31, 2021; accepted September 2, 2021; published online June 17, 2022

Abstract: Comprehending conditional statements is fundamental for hypothetical reasoning about situations. However, the online comprehension of conditional statements containing different conditional connectives is still debated. We report two self-paced reading experiments on German conditionals presenting the conditional connectives *wenn* ('if') and *nur wenn* ('only if') in identical discourse contexts. In Experiment 1, participants read a conditional sentence followed by the confirmed antecedent *p* and the confirmed or negated consequent *q*. The final, critical sentence was presented word by word and contained a positive or negative quantifier (*ein/kein* 'one/no'). Reading times of the two quantifiers did not differ between the two conditional connectives. In Experiment 2, presenting a negated antecedent, reading times for the critical positive quantifier (*ein*) did not differ between conditional connectives, while reading times for the negative quantifier (*kein*) were shorter for *nur wenn* than for *wenn*. The results show that comprehenders form distinct predictions about discourse continuations due to differences in the lexical semantics of the tested conditional connectives, shedding light on the role of conditional connectives in the online interpretation of conditionals in general.

Keywords: conditional connectives; conditionals; German; predictive processing; self-paced reading

1 Introduction

Conditional statements are a widely used linguistic tool that is essential for communicating about possible situations and their probabilities. In a simple conditional of the form "If *p*, *q*" (e.g., *If it rains, the streets get wet*), logically represented by the material implication '$p \rightarrow q$', the consequent *q* (*the streets get wet*) is stated to be true in the case that the antecedent *p* (*it rains*) is true. However, the adequacy of the material implication analysis for simple conditionals has been questioned in a number of ways. In psychology, the mental model theory (Johnson-Laird and Byrne 2002) and the suppositional theory (Evans and Over 2004) both assume that in the case of simple conditionals, only a subset of the four conditional possibilities (*p* and *q*; *p* and ¬*q*; ¬*p* and *q*; and ¬*p* and ¬*q*) are activated in the comprehender's mind. In linguistics, a crucial argument against the material implication analysis is that it yields inadequate interpretations for natural language conditionals (von Fintel 2011). Among others, Kratzer (1986) proposes a semantic analysis called the restrictor analysis in which the conditional connective *if* in simple conditionals introduces no conditional meaning in itself and *if*-clauses are used to restrict modal or generic frequency operators. Hence, the sentence *If it rains, the streets get wet* would contain a covert modal operator MUST and means that all the rain-worlds are worlds where the streets get wet. In abstract terms, a simple conditional of "If *p*, *q*" involves universal quantification over worlds or

***Corresponding author: Mathias Barthel**, Department of English and American Studies, Humboldt University of Berlin, Berlin, Germany; and Pragmatics Department, Institute for the German Language, Mannheim, Germany, E-mail: barthel@ids-mannheim.de. https://orcid.org/0000-0003-1747-1290
Rosario Tomasello, Brain Language Laboratory, Department of Philosophy and Humanities, WE4, Freie Universität Berlin, Berlin, Germany; and Cluster of Excellence 'Matters of Activity. Image Space Material', Humboldt Universität zu Berlin, Berlin, Germany, E-mail: tomasello.r@fu-berlin.de
Mingya Liu, Department of English and American Studies, Humboldt University of Berlin, Berlin, Germany, E-mail: mingya.liu@hu-berlin.de. https://orcid.org/0000-0001-9426-3669

cases and it has the semantics that all *p*-cases must be *q*-cases. These different proposals share the assumption that in the presence of a simple conditional "If *p*, *q*", a comprehender will expect the consequent to be true when the antecedent is true. When the antecedent is false, however, matters are more complicated. One factor that contributes to this is the phenomenon called conditional perfection, which we will return to after discussing the meaning of the more complex conditional *only if*.

A compositional semantic derivation of the meaning of *only if* requires a combination of the meanings of *if*, as discussed above, and *only*. Following Horn (2002), a sentence under the focus particle *only* – that is, "Only *p*" – expresses two entailments, the positive entailment *p* and a negative entailment that can be glossed as "Nothing but *p*". The meaning of "Only if *p*, *q*" would then contain the entailments "all *p*-cases are *q*-cases" and "all ¬*p*-cases are ¬*q*-cases", seemingly making *only if* a biconditional connective. However, Herburger (2015, 2019) proposes an alternative analysis of *only if*, in which the foregrounded meaning of "Only if *p*, *q*" is negative and involves universal quantification (i.e., "all ¬*p*-cases are ¬*q*-cases"), while the backgrounded meaning is positive and has existential force ("some *p*-cases are *q*-cases"). Hence, sentences like *Only if you work hard do you succeed* encode the meaning that "lack of hard work results in lack of success" and, crucially, "some hard work results in success" rather than "all hard work results in success".

Although it is clear that *if* and *only if* have different semantics, the pragmatic phenomenon of conditional perfection needs to be taken into consideration to gain the full picture of their lexical meanings in a situational context. For instance, Geis and Zwicky (1971) claim that by uttering the simple conditional sentence *If you mow the lawn, I will give you 5 dollars*, the speaker invites the hearer to additionally draw the inference "If you don't mow the lawn, I won't give you 5 dollars". The semantically weaker conditional sentence is then "perfected" to a semantically stronger biconditional one: "If and only if you mow the lawn will I give you 5 dollars". While there are different accounts for this phenomenon (van der Auwera 1997a, 1997b; Horn 2000; see Noveck 2018: Ch. 4 for a review), all of them treat conditional perfection as a pragmatic inference of a conversational nature, that is, optional and cancelable (see Grice 1989). This means that *if*-sentences encode the meaning "all *p*-cases are *q*-cases" semantically, and "all ¬*p*-cases are ¬*q*-cases" pragmatically. Van Canegem-Ardijns and Van Belle (2008), for instance, provide an extensive discussion of the contextual variability of the invited inference, which involves the strengthening of *if* to *only if* (see also van der Auwera 1997b). However, we have seen above that *only if* differs from *if* in the positive meaning component, and thus is not a biconditional connective but rather a mirror image of *if*. The question remains whether the existential meaning component in the semantics of *only if* can be strengthened as well, so that it receives a biconditional meaning. It seems to us that this is also context-dependent, that is, it is more possible in some cases (e.g., *Only if it rains do the streets get wet*) than in others (e.g., *Only if you work hard do you succeed*).

We take the analysis by Herburger (2015, 2019) as our working hypothesis, as recent evidence indicates that neither *if* nor *only if* are semantically biconditional. In a series of rating studies investigating the semantics of different conditional connectives in German, Liu and Barthel (2021) find that ratings for modus ponens ("If *p*, *q*. *p*. Therefore *q*."), which is a valid rule of inference for simple conditionals and can be validated for biconditionals in first-order logic (if we treat *iff* as a conjunction of *if*s), were higher for *wenn* ('if') than for *nur wenn* ('only if'), indicating that the material implication $p \rightarrow q$ is not part of the lexical semantics of *nur wenn*. Additionally, ratings for affirmation of the consequent ("If *p*, *q*. *q*. Therefore *p*."), which is invalid for simple conditionals but can be validated for biconditionals, were higher for *nur wenn* than for *wenn*, indicating that the mirrored implication $q \rightarrow p$ is not part of the lexical semantics of *wenn*. This means that neither of the two conditional connectives is semantically biconditional, supporting Herburger's analysis.

From a processing perspective, given these differing lexical meanings of *wenn* ('if') and *nur wenn* ('only if'), comprehenders can be assumed to generate different expectations about the probability of the consequent being true when they know about the truth or falsity of the antecedent. Expectation generation, or prediction, is an integral part of language comprehension at sentence and discourse levels (Drenhaus et al. 2014; Grisoni et al. 2020; Köhne and Demberg 2013; Pickering and Garrod 2004, 2013, 2021; Rohde and Horton 2014; Rohde et al. 2011; van Bergen and Bosker 2018; van Berkum et al. 2005; Schwab and Liu 2020; Schwab et al. 2022; Xiang and Kuperberg 2015). Upon encountering a hypothetical conditional statement, the truth probabilities of both antecedent and consequent are unknown. Conditionals are therefore non-veridical contexts, and we assume non-veridical

equilibrium to be the default for conditionals (implying that p and $\neg p$ are equal possibilities; see Giannakidou and Mari [2021]). Thus, without contextual manipulations that would trigger any bias (Liu 2019; Liu et al. 2021), comprehenders would not be expected to commit to a prediction of their truth or falsity. If, in addition to the conditional, the comprehender knows about the antecedent being true, they can expect the truth of the consequent in the case of *wenn* ('if') but to a lesser degree in the case of *nur wenn* ('only if'), since only in *wenn* would the truth of the consequent semantically follow from the truth of the antecedent ($p \rightarrow q$). In cases where the antecedent is known to be false, on the other hand, the truth of the consequent is uncertain under *wenn*, making the scenario less predictable. Under *nur wenn*, however, the consequent can be expected to be false ($\neg p \rightarrow \neg q$), making the scenario more predictable. We test these predictions and the underlying lexical meanings of the conditional connectives, measuring reading times as an indicator of prediction in conditional scenarios.

The present study consists of two self-paced reading experiments presenting natural language scenarios of conditional situations. In Experiment 1, participants read conditionals with *wenn* ('if') or *nur wenn* ('only if') in identical discourse contexts ("Wenn/Nur wenn p, q"), the confirmed antecedent (p), and then word by word the confirmed or negated consequent (q or not-q) that contained a positive or negative quantifier (*ein/kein* 'one/no'). In Experiment 2, participants read "Wenn/Nur wenn p, q. not-p. q/not-q." with the negated antecedent. In both experiments, the critical dependent measure is the reading time of the quantifier in the final sentence that reveals the consequent to be either true or false. In Experiment 1, we hypothesize that comprehenders generate stronger expectations about the truth of the consequent in *wenn* than in *nur wenn*, as the foregrounded meaning of *wenn* is "all p-cases are q-cases", while in *nur wenn* only "some p-cases are q-cases". We therefore expect shorter reading times for positive quantifiers (*ein*) in *wenn* than in *nur wenn*. In Experiment 2, on the other hand, we expect shorter reading times for negative quantifiers (*kein*) in *nur wenn* than in *wenn*, since the foregrounded meaning of *nur wenn is* "all $\neg p$-cases are $\neg q$-cases", while in *wenn* this meaning can only be inferred via conditional perfection.

2 Experiment 1

2.1 Method

2.1.1 Participants

Twenty-four German native speakers (10 female, 14 male, mean age = 33.5 years, SD = 12.8 years) were recruited online via Prolific and paid for their participation. The complete study obtained ethics approval from the ethics committee of the German Linguistic Society.

2.1.2 Materials

For this experiment, 108 critical items containing short four-sentence scenarios were composed, similar to that in example (1). The first sentence (S1) sets the scenario context; the second sentence (S2) contains a conditional (i.e., it has the form "If p, q"), using either *wenn* or *nur wenn*; the third sentence (S3) confirms the antecedent (i.e., p); the fourth sentence (S4) either confirms or negates the consequent (i.e., q or $\neg q$). Note that the critical difference in S4 about the truth or falsity of the consequent is encoded in the quantifier *ein/eine/einen* 'one (ACC.N/F/M)' or *kein/keine/keinen* 'no (ACC.N/F/M)'; henceforth, this will be referred to just as *ein* or *kein*. This ensures that the sentences containing the confirmed or negated consequent are structurally and lexically identical apart from the critical word.

(1) S1: *Leon besuchte seine Eltern und dachte sich:*
'Leon visited his parents and thought:'
S2: ***Wenn/Nur wenn*** *die Blumensträuße hübsch sind, bringe ich einen mit.*
'If the bouquets are pretty, I will take one with me.'/
'Only if the bouquets are pretty will I take one with me.'

S3: *Wie sich zeigte, waren die Blumensträuße hübsch.*
'As became apparent, the bouquets were pretty.'

S4:	*Von*	*denen*	*brachte*	*er*	***einen/keinen***	*mit*	*und*	*ging*	*weiter.*
	of	those	took	he	one/no	with	and	went	further

'Of those he took one/none and went on.'

In an additional control condition, S2 did not contain a conditional statement, but was instead of the format "Vielleicht *q*" ('Maybe *q*'). These control trials served as a baseline condition and sanity check, as reading times for negated consequents should be slower than for confirmed consequents, because processing negation should require increased effort (Farshchi et al. 2021; Kaup and Dudschig 2020; Wason 1965).

In 24 additional filler items, S3 contained a negated antecedent ("As became apparent, $\neg p$"). Filler items were followed by a polar comprehension question about the scenario. Half of the questions targeted the answer "yes" (e.g., 'Did Leon visit his parents?'), half targeted the answer "no". Twelve additional practice items were composed to be run before the experiment proper.

2.1.3 Design and procedure

The present study used a 2 × 2 factorial design, in which all items contained the factors CC (conditional connective) in S2 (*wenn/nur wenn*) and Consequent in S4 (confirmed/negated; see (1) above). Two additional, non-conditional control conditions (of the format "Maybe *q*" in S2) also contained the factor Consequent in S4 (confirmed/negated) but did not contain any CC in S2. Six experimental lists were designed using a Latin square design so that each participant read each of the scenarios only once during the experiment.

The experiment was implemented online using Ibex Farm (Drummond 2013). Participants were instructed to read the presented scenarios carefully at their own pace. Before each scenario, a fixation cross appeared in the middle of the screen, which was replaced by S1 when the subject pressed the space bar. S1 to S3 were presented as whole sentences one by one with each sentence replacing the previous one when the subject pressed the space bar. S4 was presented word by word with each word replacing the previous one when the subject pressed the space bar. Each scenario was followed by a blank screen for 1 s. Comprehension questions after filler trials had to be answered by pressing either "j" for "yes" or "n" for "no". The entire experiment took about 20 min.

2.2 Results

All subjects had a correct response rate of greater than 80% in the comprehension questions (mean = 90%, SD = 6%). After a visual inspection of the density distribution of trial lengths, 159 critical trials (6.1%) were discarded because they had a total duration of less than 3 s or more than 15 s.

The following analyses will focus on the reading times of the critical quantifier *ein/kein* ('one/no') that encodes whether the consequent *q* is true or false in the presented scenarios; that is, we focus on the fifth word in S4. Sixty-six trials (2.1%) were discarded because the reading times of the critical word deviated more than 2.5 SD from the subject mean reading times for the critical word. The critical word in the remaining 2,367 critical trials had a mean reading time of 360 ms (SD = 139 ms).

A Bayesian mixed-effects model on the reading times of the critical word was run in R (R Core Team 2021) using the package *brms* (Bürkner 2017), with CC and Consequent plus their interaction as fixed effects and as random effects by subject and by item plus the centered length of the critical word and the centered trial number as control variables. CC and Consequent were dummy coded, with "wenn" and "negative Consequent" as reference levels. Table 1 shows the model output. Figure 1 shows the modeled reading times per condition. Bayes factors for hypothesis-specific tests were computed with the *hypothesis* function of *brms*. For a guideline to the interpretation of Bayes factors and the labeling categorization of effect strengths, see Andraszewicz et al. (2015).

Table 1: Bayesian mixed-effects model output on reading times of the critical word in Experiment 1. CrI = credible interval; CC_nonCond = non-conditional control condition.

	Estimate	Lower 95% CrI	Upper 95% CrI
Intercept	378.12	330.33	425.62
length_c	0.93	−7.48	9.52
trialNumber_c	−0.46	−0.54	−0.37
CC_nurWenn	6.53	−5.02	17.99
CC_nonCond	−2.15	−13.44	9.73
Consequent_positive	−14.47	−32.78	3.84
CC_nurWenn:Consequent_positive	−9.96	−25.57	5.69
CC_nonCond:Consequent_positive	0.13	−15.15	15.84

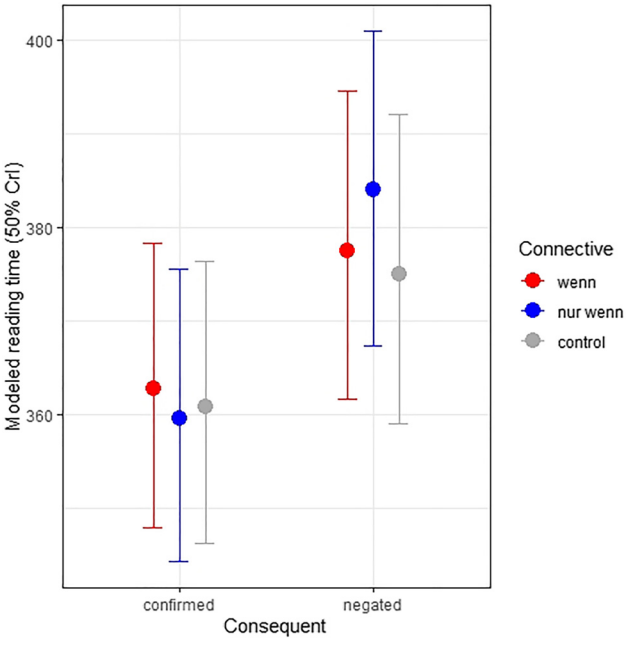

Figure 1: Modeled reading times for the critical word in Experiment 1 (confirmed antecedent), as extracted from the mixed-effects model presented in Table 1. CrI = credible interval.

Very strong evidence was found for a main effect of Consequent, with negative quantifiers being read more slowly than positive quantifiers (β = −19.45, CrI = [−33.9, −4.85], BF_{10} = 69.59). No evidence was found for a main effect of CC (β = 3.1, CrI = [−11.78, 17.87], BF_{10} = 1.79), meaning that reading times for the critical word did not differ between *wenn* and *nur wenn*. Moderate evidence was found for an interaction effect of CC and Consequent (β = 9.96, CrI = [−2.95, 23.25], BF_{10} = 8.58), with the effect of CC being slightly smaller in positive *q* than in negative *q*. Breaking this interaction down into simple effects, it was found that while there is moderate evidence for an effect of CC in negative *q* (β = 6.53, CrI = [−3.13, 16.22], BF_{10} = 6.57), there is no evidence for such an effect in positive *q* (β = −3.43, CrI = [−13.49, 6.63], BF_{10} = 0.4). In these small, moderate effects, Bayes factors are below 10, making the effects unreliable. Therefore, we will not base our interpretations on these effects.

Reading times for the critical word in the critical conditions were overall comparable to reading times of the critical word in the non-conditional control condition (β = −5.73, CrI = [−19.82, 8.53], BF_{10} = 2.88). In the control condition, a strong effect of Consequent was attested as well (β = −14.34, CrI = [−30.65, 1.99], BF_{10} = 12.79). This effect was found to be comparable to the effect of Consequent attested for the conditional scenarios (β = −5.11, CrI = [−17.04, 6.34], BF_{10} = 3.27).

A parallel analysis on the reading times of the post-critical word was conducted to check for potential spill-over effects and yielded the same pattern of results.

2.3 Discussion

This first experiment tested the reading times of the consequent in conditional scenarios containing the CCs *wenn* ('if') or *nur wenn* ('only if') and confirmed antecedents. Assuming that conditionals under *wenn* semantically convey "all *p*-cases are *q*-cases" and that conditionals under *nur wenn* semantically convey "some *p*-cases are *q*-cases", we hypothesized that comprehenders would generate stronger expectations about the truth of the consequent in *wenn* than in *nur wenn*. Therefore, we hypothesized reading times on the positive critical quantifier *ein*, which reveals the consequent of the scenario to be true, to be shorter for *wenn* than for *nur wenn*. This result was not obtained. Instead, reading times for *ein* were found to be statistically identical in conditionals containing either of the two CCs. Furthermore, the existential force in the conditional meaning of *nur wenn* can invite the inference that "not all *p*-cases are *q*-cases", or "some *p*-cases are ¬*q*-cases". While we did not form any hypothesis on this, we checked whether reading times on the negative critical quantifier *kein* after the true consequent were shorter for *nur wenn* than for *wenn*. However, reading times for the negative quantifier *kein* were also found to be statistically non-different between the two CCs. These results therefore yield no support for an observable difference in meaning or processing of conditionals containing either *wenn* or *nur wenn*. Reading times for the negative quantifier *kein* were found to be reliably longer than for the positive quantifier *ein* in both CCs, as well as in the non-conditional control scenarios, indicating that it is most probably due to increased processing costs of negation.

The results of Experiment 1 do not offer direct support of an analysis stating that the meaning component "$p \rightarrow q$" in *nur wenn* involves existential quantification. Instead, reading times of the consequent under *nur wenn* were parallel to those under *wenn*. This pattern of results could indicate that participants' reading of conditionals under *nur wenn* was pragmatically enriched, raising the meaning component "some *p*-cases are *q*-cases" to be universal. We will return to this possibility in Section 4. Alternatively, the absence of an effect of CC could be interpreted as evidence for *nur wenn* carrying a biconditional semantic meaning, that is, expressing a sufficient condition for the consequent (van der Auwera 1997b), which is subject to further contextual manipulations as in the case of *Only if you work hard do you succeed*. Since this conclusion would be based on the absence of an expected effect, we are cautious about jumping to this conclusion, especially in the light of recent findings that contradict this analysis (Herburger 2015; Liu and Barthel 2021). To more thoroughly investigate the possible interpretations of the observed results with respect to the meaning and processing of the contrasted conditional connectives, we tested participants' comprehension of conditional scenarios that contain negated antecedents in Experiment 2.

3 Experiment 2

3.1 Method

3.1.1 Participants

Twenty-nine German native speakers (12 female, 17 male, mean age = 28.5 years, SD = 8.1 years) that did not take part in Experiment 1 were recruited online via Prolific and paid for their participation.

3.1.2 Materials

Materials were the same as in Experiment 1 (see (1) above), except that S3 in the critical items contained a negated antecedent, such as *Wie sich zeigte waren die Blumensträuße nicht hübsch* ('As became apparent, the

bouquets were not pretty'). Explicit negation was used so as to keep S3 otherwise lexically identical between Experiments 1 and 2 in order to avoid potential confounds due to differences in lexical parameters. S3 in the filler items contained a confirmed antecedent.

3.1.3 Design and procedure

Design and procedure were the same as in Experiment 1.

3.2 Results

Analyses of results were parallel to Experiment 1. Five subjects were discarded because they had a correct response rate of less than 80% in the comprehension questions, leaving the data of the 24 remaining participants for analysis, who had a mean correctness rate of 93% (SD = 4%). After visual inspection of the density distribution of trial lengths, 153 critical trials (5.9%) were discarded because they had a total duration of less than 3 s or more than 15 s. A further 61 trials (2.3%) were discarded because the reading times of the critical word deviated more than 2.5 SD from the subject mean reading times for the critical word.

The critical quantifier in the remaining 2,367 critical trials had a mean reading time of 327 ms (SD = 112 ms). A Bayesian mixed-effects model on the reading times of the critical word was run using the R package *brms*,

Table 2: Bayesian mixed-effects model output on reading times of the critical word in Experiment 2. CrI = credible interval; CC_nonCond = non-conditional control condition.

	Estimate	Lower 95% CrI	Upper 95% CrI
Intercept	340.64	300.43	381.01
length_c	7.21	0.52	14.11
trialNumber_c	−0.74	−0.81	−0.67
CC_nurWenn	−12.06	−22.18	−2.09
CC_nonCond	−6.20	−15.89	3.35
Consequent_positive	−4.13	−15.55	7.35
CC_nurWenn:Consequent_positive	12.18	−1.21	25.59
CC_nonCond:Consequent_positive	2.44	−10.83	15.78

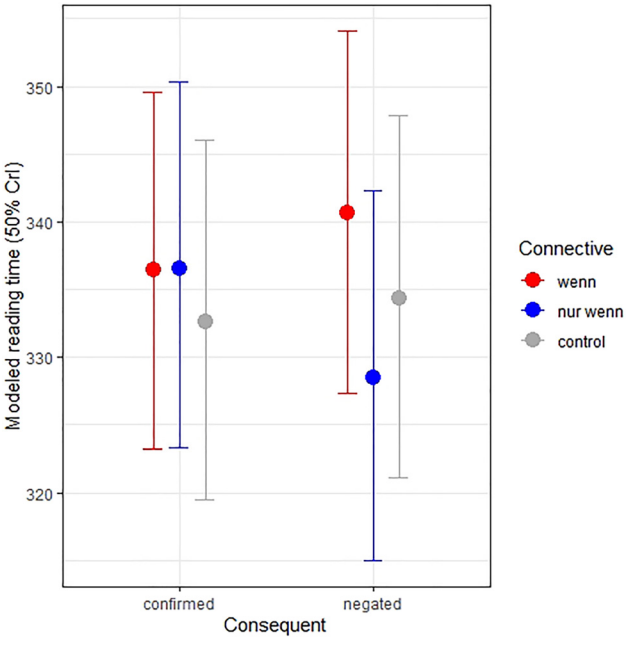

Figure 2: Modeled reading times for the critical word in Experiment 2 (negated antecedent) as extracted from the mixed-effects model presented in Table 2. CrI = credible interval.

with CC and Consequent plus their interaction as fixed effects and as random effects by subject and by item plus the centered length of the critical word and the centered trial number as control variables. CC and Consequent were dummy coded, with "wenn" and "negative Consequent" as reference levels. Table 2 shows the model output. Figure 2 shows the modeled reading times per condition.

Hypothesis-specific tests were conducted using the *hypothesis* function implemented in *brms*. No evidence was found for a main effect of Consequent ($\beta = 1.96$, CrI = [–5.87, 9.93], $BF_{10} = 0.53$). Strong evidence was found for a main effect of CC ($\beta = -11.93$, CrI = [–24.6, 0.84], $BF_{10} = 15.3$), meaning that reading times for the critical word were shorter in *nur wenn* than in *wenn*. Strong evidence was found for an interaction effect of CC and Consequent ($\beta = -12.18$, CrI = [–23.57, –1.07], $BF_{10} = 26.59$), with the effect of CC being smaller in positive Consequent (*ein*) than in negative Consequent (*kein*). Breaking this interaction down into simple effects, it was found that while there is very strong evidence for an effect of CC in negative Consequent ($\beta = -12.06$, CrI = [–20.41, –3.81], $BF_{10} = 121.45$), there is strong evidence for the absence of a CC effect in positive Consequent ($\beta = 0.13$, CrI = [–8.6, 8.86], $BF_{10} = 0.99$). Reading times for the critical word in the critical conditions were overall comparable to reading times of the critical word in the non-conditional control condition ($\beta = -3.99$, CrI = [–15.38, 7.46], $BF_{10} = 2.54$). In the control condition, no effect of Consequent was attested ($\beta = -1.68$, CrI = [–11.47, 8.17], $BF_{10} = 1.57$), and no interaction of control/critical with Consequent was found ($\beta = 3.65$, CrI = [–6.35, 13.56], $BF_{10} = 2.7$).

A parallel analysis on the reading times of the post-critical word was conducted to check for potential spillover effects, yielding no significant effects of interest.

To compare the reading times of Experiments 1 and 2, both data sets were combined and tested for differences in grand mean reading times of the critical quantifier as well as of the pre-critical region preceding the critical quantifier (containing four words). Two Bayesian mixed-effects models were run, with Experiment as a fixed effect and as a random effect by item and a random intercept by subject plus the centered trial number as a control variable. No reliable evidence was found for a main effect of Experiment in the pre-critical region ($\beta = -84.84$, CrI = [–242.27, 75.83], $BF_{10} = 4.24$), nor in the critical quantifier ($\beta = -34.03$, CrI = [–85.35, 17.98], $BF_{10} = 6.03$), indicating that any global differences in reading times between experiments are mainly caused by inter-individual variation between the two groups of subjects.

3.3 Discussion

This second experiment tested the reading times of the consequent in conditional scenarios containing the CCs *wenn* ('if') or *nur wenn* ('only if') and negated antecedents. We assumed that conditionals under *nur wenn* contain the meaning component "all $\neg p$-cases are $\neg q$-cases" in their semantics and that conditionals under *wenn* only do so pragmatically via conditional perfection, which allows some $\neg p$-cases to be q-cases (e.g., in the case of *If it rains, the streets are wet*, even without rain, the streets can be wet due to snow). We thus hypothesized that with negative minor premises ($\neg p$), comprehenders would generate stronger expectations about the negation of the consequent ($\neg q$) in *nur wenn* than in *wenn*, to be indexed by shorter reading times on the negative critical quantifier *kein* for *nur wenn* than for *wenn*. This pattern was indeed observed, yielding evidence for the assumed difference in meaning of conditionals with the CCs *wenn* versus *nur wenn*. This result can be interpreted as evidence corroborating the analysis that "$\neg p \rightarrow \neg q$" is part of the semantics of *nur wenn*, while this meaning is not a component of the semantics of *wenn*. Furthermore, as *nur wenn* is compatible with a minor premise that denies the antecedent, a negated consequent ought to follow and an affirmative one ought to be surprising. However, reading times of the positive Consequent (q) did not differ between the two CCs, yielding no additional evidence that comprehenders generated different predictions about the consequent under *wenn* versus *nur wenn* in these cases.

In the control condition, no effect of negation in the consequent was observed, unlike in Experiment 1. One possible interpretation of this is that while possibility modals such as "Vielleicht q" ('Maybe q') presuppose q and $\neg q$ as equal possibilities (Giannakidou and Mari 2021), they might carry a weak positive speaker bias toward q. The truth of the antecedent p in Experiment 1 either left the bias unimpacted or even strengthened it,

so that q was processed faster than $\neg q$. In Experiment 2, the falsity of p, on the other hand, may have removed the weak positive bias. Thus, the non-veridical equilibrium remains reserved, explaining the lack of differences between q versus $\neg q$ in processing.

Participants in Experiment 2 might have been assumed to process the consequent less deeply than in Experiment 1, as the minor premise on the denial of the antecedent argument potentially renders the consequent less relevant. This hypothesis would predict faster reading times in Experiment 2 as compared to Experiment 1 right from the beginning of the consequent, even before the critical quantifier. This result was not obtained, indicating that the consequent was attended to comparatively in both experiments. Future studies could add comprehension questions targeting the consequent sentence to critical items as well so as to directly test this hypothesis.

4 General discussion and conclusion

In everyday communication, conditional connectives (CCs) are critical cues in predicting possible developments of situations. Presenting short conditional scenarios, this study investigated the meaning contribution of two different CCs, *wenn* ('if') and *nur wenn* ('only if'), and the predictions they trigger in two self-paced reading experiments. In Experiment 1, the conditional was followed by a confirmed antecedent, while in Experiment 2 it was followed by a negated antecedent. The critical word in the ensuing consequent, the quantifier *ein* ('one') or *kein* ('no'), revealed differences in reading times when preceded by a confirmed or negated antecedent under the CCs *wenn* or *nur wenn*. Negative quantifiers were generally read more slowly than positive quantifiers after confirmed antecedents in Experiment 1, with no differential effect of CC. Interestingly, after negated antecedents in Experiment 2, reading times for confirmed consequents were identical for both CCs, while reading times for negated antecedents were significantly faster after *nur wenn* than after *wenn*.

Following a common analysis of the meaning of *wenn*, the central component of its semantics is that in all cases where the antecedent is true, the consequent must also be true ("all p-cases are q-cases"; Kratzer 1986), with the invited inference that "all $\neg p$-cases are $\neg q$-cases" (Geis and Zwicky 1971; Horn 2000; Van Canegem-Ardijns and Van Belle 2008; van der Auwera 1997b). In contrast, according to Herburger (2015, 2019), the semantics of *nur wenn* is that all cases where the antecedent is false are cases where the consequent is also false ("all $\neg p$-cases are $\neg q$-cases"), and that some p-cases are q-cases. Taking these analyses as working hypotheses, comprehenders of a conditional scenario of the form "Wenn p, q. p.", as presented in Experiment 1, were hypothesized to predict the truth of the consequent q. Comprehenders of a conditional scenario of the form "Nur wenn p, q. p.", on the other hand, were hypothesized to generate this prediction to a lesser degree. Thus, we expected reading times of the critical quantifier *ein* (indicating q) after p to be longer for *nur wenn* than for *wenn*. For the negative quantifier *kein* (indicating $\neg q$), the opposite pattern of reading times was predicted, as "some p-cases are $\neg q$-cases" is in line with the meaning of *nur wenn* but not with that of *wenn*. We thus expected reading times on the negative quantifier *kein* (indicating $\neg q$) after p to be shorter for *nur wenn* than for *wenn*. In conditional scenarios with negated antecedents ("Wenn p, q. $\neg p$."), as presented in Experiment 2, *nur wenn* would allow comprehenders to generate the expectation of a negated consequent, while *wenn* should trigger such a prediction not semantically but pragmatically. Therefore, reading times for the negative quantifier *kein* were predicted to be shorter after *nur wenn* than after *wenn*. Reading times for the positive quantifier *ein*, on the other hand, were predicted to be longer after *nur wenn*, because of the violation of comprehenders' expectations.

This predicted pattern of reading times was observed in Experiment 2, with a significant interaction of CC and Consequent. In particular, the results for the negative quantifier *kein* showed a clear difference between connectives, with faster reading times after *nur wenn* than after *wenn*. This shows that the related inference "If $\neg p$, $\neg q$" is weaker or more context-dependent for conditionals under *wenn*, in line with the pragmatic accounts of conditional perfection. Results for the positive quantifier showed no difference in reading times between connectives. While the lack of a difference needs further investigation, it is possible that it results

from conditional perfection of *wenn* so that the affirmation of q given $\neg p$ is equally surprising as in the case of *nur wenn*.

Contrary to these hypotheses, however, no interaction of CC and Consequent was found on the reading times for the critical quantifier in Experiment 1. While the absence of the expected interaction effect needs to be interpreted with care, two potential explanations come to mind. The first would be that *nur wenn* is semantically biconditional. In the light of recent analyses and findings of the meaning and interpretation of *nur wenn*, this possibility seems unlikely. The second, alternative explanation is that *nur wenn* was pragmatically enriched to a biconditional reading via a generalized conversational implicature, which cannot be easily canceled. This explanation appears reasonable for the scenarios used as stimuli in the present study. In (1), for example, if the bouquets are pretty, Leon is highly likely to buy one, especially as he can control the action himself. This contrasts with cases involving less controllable events; for example, success in the case of hard work, as analyzed by Herburger (2019).

Taken together, comprehenders were shown to use the different meaning nuances of distinct conditional connectives to predict the probabilities of conditional situations and the linguistic expressions that encode them. The different semantics of conditional connectives have been shown to be taken into account by comprehenders during the online processing of conditional scenarios, as the differences in processing costs surface already at the earliest point in a given sentence that encodes the falsity of the consequent. This study corroborates accumulating evidence that comprehenders immediately use available linguistic material to make predictions about the probabilities of upcoming material (Xiang and Kuperberg 2015; van Berkum et al. 2005; Schwab and Liu 2020) – in this case, the conditional background of a given situation – to predict the probability of the consequent.

Acknowledgments: We thank Ira Noveck for reviewing an earlier version of this paper and Ezgi Akyar and Matthias Andreß for their support in creating the experimental stimuli.
Research funding: M.B. and M.L.'s work on the project was supported by a DFG (German Research Foundation) grant for the project Semantics and Pragmatics of Conditional Connectives: Cross-linguistic and Experimental Perspectives (project number: 367088975).
Author contributions: M. B. and M. L. conceived and designed this research together. M. B. conducted the experiments and the data analyses. All authors contributed to the interpretation of the data. M. B. and M. L. were responsible for the writing of the manuscript, with M. B. preparing the first draft, and revisions from M. L. and R. T. M. L. provided funding acquisition, project administration, and resources.

References

Andraszewicz, Sandra, Benjamin Scheibehenne, Jörg Rieskamp, Raoul Grasman, Josine Verhagen & Eric-Jan Wagenmakers. 2015. An introduction to Bayesian hypothesis testing for management research. *Journal of Management* 41(2). 521–543.

Bürkner, Paul-Christian. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). https://doi.org/10.18637/jss.v080.i01.

Drenhaus, Heiner, Vera Demberg, Judith Köhne & Francesca Delogu. 2014. Incremental and predictive discourse processing based on causal and concessive discourse markers: ERP studies on German and English. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 36, 403–408.

Drummond, Alex. 2013. *Ibex Farm*. https://ibex.spellout.net/.

Evans, Jonathan St. B. T. & David E. Over. 2004. *If*. Oxford: Oxford University Press.

Farshchi, Sara, Annika Andersson, Joost van de Weijer & Carita Paradis. 2021. Processing sentences with sentential and prefixal negation: An event-related potential study. *Language, Cognition and Neuroscience* 36(1). 84–98.

Geis, Michael L. & Arnold M. Zwicky. 1971. On invited inferences. *Linguistic Inquiry* 2(4). 561–566.

Giannakidou, Anastasia & Alda Mari. 2021. *Truth and veridicality in grammar and thought: Modality, mood, and propositional attitudes*. Chicago: University of Chicago Press.

Grice, Herbert P. 1989. *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Grisoni, Luigi, Rosario Tomasello & Friedmann Pulvermüller. 2020. Correlated brain indexes of semantic prediction and prediction error: Brain localization and category specificity. *Cerebral Cortex* 31(3). 1553–1568.

Herburger, Elena. 2015. Only if: If only we understood it. *Proceedings of Sinn Und Bedeutung* 19. 284–301.
Herburger, Elena. 2019. Bare conditionals in the red. *Linguistics and Philosophy* 42(2). 131–175.
Horn, Laurence R. 2000. From if to iff: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics* 32(3). 289–326.
Horn, Laurence R. 2002. Assertoric inertia and NPI licensing. *Chicago Linguistic Society* 38. 55–82.
Johnson-Laird, Philip N. & Ruth M. J. Byrne. 2002. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review* 109(4). 646–678.
Kaup, Barbara & Carolin Dudschig. 2020. Understanding negation: Issues in the processing of negation. In Viviane Déprez & Maria Teresa Espinal (eds.), *The Oxford handbook of negation*, 635–655. Oxford: Oxford University Press.
Köhne, Judith & Vera Demberg. 2013. The time-course of processing discourse connectives. *Proceedings of the Annual Meeting of the Cognitive Science Society* 35. 2760–2765.
Kratzer, Angelika. 1986. Conditionals. Chicago Linguistics Society 22(2). 1–15. Reprinted as: Kratzer, A. 1991. Conditionals. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 651–656. Berlin: de Gruyter.
Liu, Mingya. 2019. The elastic nonveridicality property of indicative conditionals. *Linguistics Vanguard* 5(s3). https://doi.org/10.1515/lingvan-2019-0007.
Liu, Mingya & Mathias Barthel. 2021. Semantics processing of conditional connectives: German *wenn* 'if' vs. *nur wenn* 'only if. *Journal of Psycholinguistic Research* 50(6). 1337–1368.
Liu, Mingya, Stephanie Rotter & Anastasia Giannakidou. 2021. Bias and modality in conditionals: Experimental evidence and theoretical implications. *Accepted for publication in Journal of Psycholinguistic Research*.
Noveck, Ira. 2018. *Experimental pragmatics: The making of a cognitive science*. Cambridge University Press. https://doi.org/10.1017/9781316027073.
Pickering, Martin J. & Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27(2). 169–190. discussion 190–226.
Pickering, Martin J. & Simon Garrod. 2013. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences* 36(4). 329–347.
Pickering, Martin J. & Simon C. Garrod. 2021. *Understanding dialogue: Language use and social interaction*. Cambridge: Cambridge University Press.
R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Available at: http://cran.r-project.org/.
Rohde, Hannah & William S. Horton. 2014. Anticipatory looks reveal expectations about discourse relations. *Cognition* 133(3). 667–691.
Rohde, Hannah, Roger Levy & Andrew Kehler. 2011. Anticipating explanations in relative clause processing. *Cognition* 118(3). 339–358.
Schwab, Juliane & Mingya Liu. 2020. Lexical and contextual cue effects in discourse expectations: Experimenting with German 'zwar … aber' and English 'true/sure…but. *Dialogue & Discourse* 11(2). 74–109.
Schwab, Juliane, Ming Xiang & Mingya Liu. 2022. Antilocality effect without head-final dependencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 48(3). 446–463.
van Bergen, Geertja & Hans R. Bosker. 2018. Linguistic expectation management in online discourse processing: An investigation of Dutch *inderdaad* 'indeed' and *eigenlijk* 'actually. *Journal of Memory and Language* 103. 191–209.
van Berkum, Jos J. A., Colin M. Brown, Pienie Zwitserlood, Valesca Kooijman & Peter Hagoort. 2005. Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31(3). 443–467.
Van Canegem-Ardijns, Ingrid & William Van Belle. 2008. Conditionals and types of conditional perfection. *Journal of Pragmatics* 40(2). 349–376.
van der Auwera, Johan. 1997a. Pragmatics in the last quarter century: The case of conditional perfection. *Journal of Pragmatics* 27(3). 261–274.
van der Auwera, Johan. 1997b. Conditional perfection. In Angeliki Athanasiadou & René Dirven (eds.), *Conditionality, hypotheticality, counterfactuality*, 169–190. Amsterdam: John Benjamins.
von Fintel, Kai. 2011. Conditionals. In C. Maienborn, K. von Heusinger & P. Portner (eds.), *Semantics: An international handbook of natural language and meaning*, 1515–1538. Berlin: de Gruyter.
Wason, Peter C. 1965. The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior* 4(1). 7–11.
Xiang, Ming & Gina Kuperberg. 2015. Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience* 30(6). 648–672.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/lingvan-2021-0083).