



Yatsushiro, Kazuko & Trinh, Tue & Żygis, Marzena & Solt, Stephanie & Benz, Anton & Krifka, Manfred. 2022. *Certainly but not certain: The expression of subjective and objective probability*. *Glossa: a journal of general linguistics* 7(1). pp. 1–48. DOI: <https://doi.org/10.16995/glossa.5847>



Open Library of Humanities

## ***Certainly but not certain: The expression of subjective and objective probability***

**Kazuko Yatsushiro**, Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Schützenstr, Berlin, DE, [yatsushiro@leibniz-zas.de](mailto:yatsushiro@leibniz-zas.de)

**Tue Trinh**, Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Schützenstr, Berlin, DE, [trinh@leibniz-zas.de](mailto:trinh@leibniz-zas.de)

**Marzena Żygis**, Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Schützenstr, Berlin, DE, [zygis@leibniz-zas.de](mailto:zygis@leibniz-zas.de)

**Stephanie Solt**, Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Schützenstr, Berlin, DE, [solt@leibniz-zas.de](mailto:solt@leibniz-zas.de)

**Anton Benz**, Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Schützenstr, Berlin, DE, [benz@leibniz-zas.de](mailto:benz@leibniz-zas.de)

**Manfred Krifka**, Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Schützenstr, Berlin, DE, [krifka@leibniz-zas.de](mailto:krifka@leibniz-zas.de)

This paper investigates the interpretation of the epistemic modal adjectives *possible*, *certain* and their adverbial counterparts *possibly*, *certainly*, taking the perspective that the former express objective modality, whereas the latter express subjective modality. In support of this view, we report on two experiments that assess speakers' acceptance of *possible*, *certain*, *possibly*, *certainly* in different contexts. Our results extend those of Lassiter (2016) in demonstrating a difference in interpretation between the adjectival and adverbial modals, with *possibly* less acceptable than *possible* but conversely *certainly* more acceptable than *certain* in the same situations. Furthermore, the results of our experiments suggest a new insight, namely that these differences depend, to some extent, on the probability of the eventuality in question and polarity of the prejacent sentence.



# 1 Introduction

## 1.1 Force and category of epistemic operators

The empirical focus of this paper is on modal sentences containing one of the following four items: *certain*, *certainly*, *possible*, *possibly*. These sentences can be categorized in terms of the force and the syntactic category of their modal operator: They can be identified by (i) whether the modal is universal or existential, and (ii) whether the modal is adjectival or adverbial.

| (1) | modal       | adjectival                                      | adverbial                             |
|-----|-------------|---|---------------------------------------|
|     | universal   | <i>It is certain that John passed the exam</i>  | <i>John certainly passed the exam</i> |
|     | existential | <i>It is possible that John passed the exam</i> | <i>John possibly passed the exam</i>  |

There is definitely a sense in which the sentences on the top row entail those on the bottom row. Sequences such as *John certainly passed the exam, therefore it is possible that he passed the exam* might sound contrived, but must surely be accepted as valid arguments. Thus, the vertical contrast between universal and existential operators, unsurprisingly, is semantically relevant, resembling that between  $\Box$  and  $\Diamond$  in modal logic. Whether the horizontal contrast between adjectival and adverbial operators, i.e., that between sentences on the same row, has any interpretive effect and, if it does, what precisely this effect is, seems much less clear. It is this latter question which is the object of the present study.

The approach we take is experimental: We confront native speakers with different scenarios and test their reactions to sentences of the form *certain  $\phi$* , *certainly  $\phi$* , *possible  $\phi$* , and *possibly  $\phi$* , as exemplified in (1) under these scenarios. The scenarios are designed to control for properties of the prejacent  $\phi$  regarding its polarity and probability. Specifically, the independent variables are: (A) whether  $\phi$  contains negation, (B) how probable  $\phi$  is, and (C) whether the value of  $\phi$ 's probability is arrived at subjectively, based on guess work, or objectively, based on calculation.

The reasons for choosing (A), (B) and (C) as criteria will be elaborated in sections 3 and 4. It should be noted here that taking these questions, specifically (B) and (C), to be the starting point of our investigation indicates we must already have an intuition that an interpretive difference between an adjectival modal and its adverbial counterpart does in fact exist and, moreover, that this difference is related in some way to the nature of the prejacent's probability. Subsection 1.2 discusses various observations which corroborate this intuition, as well as several theoretical positions which have been taken towards these observations. At the center of this discussion is the notion that epistemic modality may be subjective or objective, and that degrees of assertability may be grammatically represented. The discussion in 1.2 will serve as background for the presentation of our study. In section 2, we list the hypotheses which are to be tested empirically. Sections 3 and 4 describe the experiments we conducted. Section 5 discusses their results. Appendices (Norming study and Comprehension questions) provide, for the interested readers, some more details concerning the design of the experiments.

## 1.2 Background discussion

### 1.2.1 Subjective vs. objective epistemic modality

“Epistemic modality”, as used in linguistics, refers to the expression of “an estimation, ... typically, but not necessarily by the speaker, of the chances or the likelihood that the state of affairs expressed in the clause applies to the world” – such is the general definition in a recent handbook (Nuyts 2016: 38). This characterization appeals to the phenomenon that the role of the judge of the “estimation of likelihood” is not restricted to the speaker, but may be taken by other epistemic authorities, like recognized experts. Lyons (1977: 797 ff) distinguished between “subjective” and “objective” epistemic modality in the readings of modal verbs like *must*, assuming that the former refers to the personal opinion of the speaker, whereas the latter refers to some factual chance that is of an interpersonal nature. Nuyts (1992) proposed that objective epistemic modality corresponds to a setting of the epistemic source to some authority that may be different from the speaker. A formal device that is able to express this dependency of epistemic modal statements was introduced by von Stechow & Gillies (2007), who propose a parameter for epistemic operators that may be set to the speaker or to some other epistemic authority, like a group of experts (cf. also Roberts 2019).

Differentiating between subjective and objective interpretations of epistemic modal verbs like *must* and *may* is a subtle affair. It is far easier to consider epistemic adverbs like *certainly* and *possibly* in contrast to epistemic adjectives like *certain* and *possible*. They clearly differ in their linguistic form, and they differ in their distribution in ways that suggest that epistemic adverbs like (2) are interpreted subjectively, whereas epistemic adjectives like (3) are interpreted objectively (Hengeveld 1988; Nuyts 1992; 1993).

(2) *This rock certainly / probably / possibly is a meteorite.*

(3) *It is certain / probable / possible that this rock is a meteorite.*

We will present the known differences between epistemic adverbs and adjectives and discuss how they relate to the idea that the former are subjective and the latter are objective epistemic operators.

One well-known distributional difference is that epistemic adverbials generally cannot be in the scope of negation, in contrast to epistemic adjectives (cf. Lyons 1977; Bellert 1977; Hengeveld 1988; Ernst 2009; Piñon 2009; Wolf et al. 2016). This holds for syntactic and morphological negation.<sup>1</sup>

---

<sup>1</sup> The claims concerning negation from the literature deserve greater scrutiny. Google NGrams, which tracks published English sources, shows frequent usages of sentence adverbials *not certainly*, *not probably* and *not possibly* in the 19th century. *Uncertainly* appears to occur only as manner adverbial, *improbably* and *impossibly* occurs within DPs as in *impossibly cute*.

- (4) a. *This rock is \*not certainly / probably / possibly a meteorite.*  
 b. *This rock is \*uncertainly / \*improbably / \*impossibly a meteorite.*
- (5) a. *It is not certain / probable / possible that this rock is a meteorite.*  
 b. *It is uncertain / improbable / impossible that this rock is a meteorite.*

There are apparent exceptions to this generalization. Bellert (1977) mentions *undoubtedly*, but this predicate has an overall positive meaning and is not derived from *\*doubtedly* (for *not necessarily*, cf. Papafragou (2006)). The generalization can be related to subjective and objective epistemicity, following Krifka (to appear (a)): When a speaker S asserts a sentence of the form *certainly*  $\phi$ , S wants to communicate the proposition  $\phi$  but indicates a reduced commitment to the proposition that S considers  $\phi$  certain. The commitment to a subjective estimation of likelihood is easier to defend than the commitment to the proposition  $\phi$  itself, hence assertion of *certainly*  $\phi$  appears to be weaker than assertion of  $\phi$  (cf. also Wolf (2015) for a related but different explanation). The assertion of *not certainly*  $\phi$  defeats the pragmatic purpose of communicating  $\phi$ , as S would express doubts about the truth of  $\phi$ . In contrast, when S asserts *it is certain that*  $\phi$ , S communicates that there is a relevant epistemic instance that considers  $\phi$  certain. As the epistemic operator is part of the communicated proposition itself, it can be under the scope of negation.

A second distributional difference is that the strong epistemic adverbial *certainly* does not easily occur in polar questions, in contrast to the epistemic adjective *certain*. This is predicted under the hypothesis that *certainly* is a subjective epistemic operator anchored to the speaker, as the epistemic authority in questions is handed over to the addressee.<sup>2</sup>

- (6) *Is this rock (\*certainly) a meteorite?*
- (7) *Is it certain that this rock is a meteorite?*

The interactions with negation and question formation are distributional characteristics, leading to reduced grammaticality. The third observation refers to a difference in interpretation. Epistemic adjectives can be in the scope of tense, meaning that they make reference to a past

---

<sup>2</sup> The epistemic adverbs *probably* and in particular *possibly* sometimes do occur in questions, cf. also Herbristrit 2020. We think that the reason for this is that in questions, the epistemic authority is flipped to the addressee, and the weaker epistemic adverbs allow and invite an answer even if the addressee is less certain about the issue at hand. Using Google n-gram we could even locate a few examples of *certainly* in questions, typically in 19th century texts. One particularly interesting case can be found in John Henry Newman in a text about St. Mary: “Consider what I have said. Is it, after all, *certainly* irrational? is it *certainly* against the Scriptures? ... is it *certainly* idolatrous? I cannot help smiling as I put the questions” (highlighting in the original). These questions are clearly rhetorical, implying that the answers consist of the negated prejacent of these questions. It should also be noted that the generalization does not hold for declarative questions, which allow for epistemic adverbials, as in *This rock is certainly a meteorite?*

belief state that may have ceased at the time of utterance (cf. Wolf 2015). This is exemplified in (8). Epistemic adverbs, by contrast, cannot be in the scope of tense, making the same continuation infelicitous, per (9). Again, this can be explained under the hypothesis that epistemic adverbials are subjective, anchored to the epistemic authority of the speaker at the time of utterance.

- (8) *It was certain that the Socialists won the election, but then it turned out that some of the results had been reported incorrectly, so in fact, now there is a chance that the Conservatives won.*
- (9) *The Socialists certainly won the election, #but then it turned out that some of the results had been reported incorrectly, so in fact, now there is a chance that the Conservatives won.*

The issue of wide scope of negation and tense, as well as the occurrence in questions and in relation to information structure, have also been discussed with respect to other epistemic modals, especially modal verbs (cf. Lyons 1977, for German Öhlschläger 1989 and the survey in Maché 2019).

Another distributional difference is that epistemic adjectives easily occur in the protasis of conditionals and in the scope of factive predicates, in contrast to epistemic adverbs (cf. Wolf et al. 2016), as evidenced by (10). This can be related to our hypothesis if we assume that the protasis of a conditional has to specify an objective fact, not just a subjective estimation by the speaker.

- (10) *?If it certainly / probably / possibly will snow, we need better shoes.*
- (11) *If it is certain / probable / possible that it will snow, we need better shoes.*
- (12) *?It is surprising that Mary certainly / probably / possibly won the race.*
- (13) *It is surprising that it is certain / probable / possible that Mary won the race.*

Furthermore, while we can combine epistemic adjectives and adverbials, cf. (14) (cf. Moss 2015), the combination of two epistemic adverbials in the same clause seems to be excluded, cf. (15). This can be explained under our hypothesis: We can express a subjective estimation of likelihood for an objective probability, but not two conflicting subjective estimations of likelihood.

- (14) *It is certainly possible that Mary will win the race.*
- (15) *\*Mary will certainly possibly win the race.*

Another property that distinguishes epistemic adverbs from adjectives is that the latter, but not the former, can be modified by expressions such as *ninety-five percent*, as evidenced by

the following contrast.<sup>3</sup> This is consistent with our hypothesis, as the subjective estimation of likelihood is plausibly not quantifiable in the same way as objective probabilities, which can be based on frequencies.

(16) *It is ninety-five percent certain that it will snow.*

(17) *\*It will ninety-five percent certainly snow.*

It has been claimed that epistemic adjectives and adverbs can give conflicting information without leading to contradiction, cf. (18), an example from Nilsen (2004), in contrast to (19) and (20).<sup>4</sup>

(18) *It is possible that Le Pen will win, even though he certainly won't.*

(19) *#Le Pen will possibly win, even though he certainly won't.*

(20) *#It is possible that Le Pen will win, even though it is certain that he won't.*

For a combination with similar semantic effect, cf. examples like (21), which are non-contradictory even though the second clause is understood as 'I believe not'.

(21) *It is possible that Le Pen will win but I am certain that he won't.*

For Nilsen, cf. also Piñon (2009), *certainly* is a speaker-oriented adverb, hence has a subjective interpretation. (18) says that it is coherent that a speaker deviates in his or her subjective estimation of likelihood (expressed by the adverb *certainly*) from the objective estimation (expressed by the adjective *possible*). The direct confrontation of subjective estimations as in (19) and of objective estimations as in (20) leads to infelicity.

Subjective epistemic predicates, at least in root clauses,<sup>5</sup> are subjective with respect to the speaker, whereas objective epistemic predicates may involve other epistemic authorities (Nuyts 2001). This is supported by the following examples (cf. Wolf et al. 2016; Herbstritt 2020). The continuation in (22) is pragmatically infelicitous because the epistemic adverb has speaker A as its epistemic source.

---

<sup>3</sup> Note that both *certain* and *certainly* can in principle be modified, say by *very*. See note 20.

<sup>4</sup> However, Wolf et al. (2016) showed experimentally that speakers rate (18) and (19) as equally incoherent, similar to a plain contradiction (cases like (20) were not tested). But the reaction times for (18) were significantly longer, indicating a delayed recognition of the inconsistency. The late reaction can be explained by saying that the two sub-clauses are not immediately felt to be contradictory, but that it is pragmatically odd in the experimental context to apply different standards for subjective and objective epistemic statements.

<sup>5</sup> In complement clauses, the speaker-oriented epistemicity may be shifted to the subject of the embedded predicate, as in *John thinks that Mary certainly will win the race*.

(22) A: *Trump will probably lose.*  
 B: *#Who says so?*

(23) A: *It is probable that Trump will lose.*  
 B: *Who says so?*

However, this contrast has never been experimentally shown, and while we agree that there is a difference, (22) does not appear very odd. We think that even if the first sentence expresses a speaker-related epistemic proposition, the addressee can ask for the source of information.

Examples (22) and (23) clearly differ in another respect: While (B) in (22) is understood as asking *Who says that Trump will lose?*, (B) in (23) can also be understood as *Who says that it is probable that Trump will lose?* That is, the anaphoric options for the sentential anaphor *so* are different. Similar differences have been observed with assertive antecedents and the sentential anaphor *that*, as in the following cases (cf. Papafragou 2006; Wolf 2015; Herbstritt 2020).<sup>6</sup>

(24) A: *Mary is certainly / probably / possibly at home.*  
 B: *That's not true.*  
 'It is not true that Mary is at home.'

(25) A: *It is certain / probable / possible that Mary won the race.*  
 B: *That's not true.*  
 'It is not true that it is certain / probable / possible that Mary won the race.'

The judgements in (24) and (25) show that epistemic adverbs are not part of the at-issue-meaning, in contrast to epistemic adjectives (cf. Beaver et al. 2017). Additional evidence for this claim is adduced by Herbstritt (2020), cf. also Roberts (2017). A question that asks for the likelihood of a proposition cannot be answered by an assertion with an epistemic adverbial, only by one with an epistemic adjective. Notice, also, that there is no appropriate *wh*-form for epistemic adverbials (or other sentence adverbials, for that matter).

(26) A: *How likely is it that Mary will win the race?*  
 B: *?Mary will certainly / probably / possibly win.*  
 B': *It is certain / probable / possible that Mary will win.*

The differences between subjective adverbs and objective adjectives is correlated with the information-structural status of the epistemic operator. One common description is that the meaning contribution of epistemic modality is outside of the proposition that is conveyed as an

---

<sup>6</sup> An anonymous reviewer points out that such observations as that made in (24) and (25) can be described in terms of "inferred QUD". In particular, *Mary is certainly at home* seems to answer the question whether Mary is at home, "with a side comment about speaker confidence about this answer," while *it is certain that Mary is at home* seems to answer the question whether it is certain that Mary is at home, or how likely it is that Mary is at home. We agree with the reviewer's remark.



expression of an attitude of the speaker towards that proposition (cf. von Fintel & Gillies 2007, with references going back to Kant and Frege).<sup>7</sup> As has been observed, this appears to hold only for subjective epistemicity, as expressed by epistemic adverbs (cf. Hengeveld 1988; Nuyts 1993; Verstraete 2001). When we see subjective epistemicity as an expression of the attitude of the speaker to the asserted proposition at the time of the utterance, we can explain why it cannot be negated, temporally shifted, directly responded to, or questioned. This is different for objective epistemicity, as expressed by epistemic adjectives, which are part of the proposition asserted. This does not mean that epistemic adjectives are necessarily part of the at-issue-information, but epistemic adverbs, at any rate, seem to be excluded from it.

The interpretation of epistemic adverbs as expressing an epistemic stance of the speaker towards the prejacent proposition and of epistemic adjectives as being part of that proposition itself, suggests that epistemic adverbs are situated outside of the syntactic projection that identifies that proposition (cf. Hengeveld 1988; Krifka to appear (a)). However, there are uses of epistemic adverbs that share properties with epistemic adjectives, e.g., they can be negated, as in *not necessarily*, cf. Papafragou 2006, and German *unmöglich* ‘impossibly’, cf. Krifka (to appear (b)). This is evidence that epistemic adverbs can also be interpreted as part of the communicated proposition. We suggest that these epistemic adverbs are lexically specified as having an objective interpretation, which allows for a realization within the communicated proposition.

### 1.2.2 Degrees of assertability

In section 1.2.1 we have presented evidence that epistemic adverbs are interpreted as subjective epistemic modals, and as outside of the asserted proposition, in contrast to epistemic adjectives. We now turn to the issue: Do these two types of epistemic operators have different interpretations? That is, given a certain situation that leads to a certain estimation of likelihood for a proposition, would speakers apply an epistemic adjective differently from its corresponding epistemic adverb? As there is no *prima facie* reason why, e.g., *certainly* and *certain*, or *possibly* and *possible*, should target different likelihood values for their proposition, any difference in their use should be attributable to their hypothesized subjective/objective dimension and/or their hypothesized information status as being part of the asserted proposition or not. In this section we will review previous findings on the use of epistemic operators in particular situations.

There is substantial research on how linguistic expressions can be related to numerical estimations of probability in psychology, cf. e.g., Beyth-Marom (1982), Wallsten et al. (1986), the survey article of Clark (1990), and the discussion in Herbstritt (2020). However, these authors

---

<sup>7</sup> Note that von Fintel & Gillies (2007) presents this view as one most common in the “descriptive” literature. Thus, it is a description, not an analysis. The analysis provided by von Fintel & Gillies (2007) actually regards the whole modalized sentence as expressing one complex proposition. We thank an anonymous reviewer for pointing out the need for this clarification.



generally only discuss adjectival concepts like *it is certain/probable/possible*, not their adverbial counterparts. This is not surprising, as it should be easier to “translate” objective epistemic operators into objective probabilities. For our current purpose, the literature from psychology is not particularly relevant as it deals only with objective epistemic expressions.

Wolf (2015) and Wolf et al. (2016) assume that epistemic adverbs modify the illocutionary act of assertion, whereas epistemic adjectives modify the asserted proposition. More specifically, epistemic adverbs are said to reduce the strength of the assertion. This helps to explain the distribution of sentences with epistemic adverbs and epistemic adjectives. It also leads to different justifications for asserting propositions under different levels of likelihood estimation on the one hand, and for asserting propositions about different levels of likelihood estimation on the other. Focusing on *certain* and *possible*, we can render Wolf’s analysis as follows:

- (27) ASSERT<sub>possibly</sub>( $\phi$ ): Assertion of  $\phi$  with weak strength ( $> 0$ )  
 ASSERT( $\phi$ <sub>possible</sub>): Assertion that ( $\phi$ )  $> 0$  with regular strength ( $\geq$  high)
- (28) ASSERT<sub>certainly</sub>( $\phi$ ): Assertion of  $\phi$  with high strength ( $= 1$ )  
 ASSERT( $\phi$ <sub>certain</sub>): Assertion that  $P(\phi) = 1$  with regular strength ( $\geq$  high)

The implementation is problematic, as assertions without epistemic modifications, as in *It will snow*, are generally stronger than assertions with subjective epistemic modifiers, as in *It will certainly snow*, cf. Karttunen (1972) – as expressed in Halliday’s dictum, “we only say we are certain when we are not” (Halliday & Matthiessen 2004: 625). This weakening effect of modals is contrary to Wolf’s analysis, as ASSERT has strength  $\geq$  high, whereas ASSERT<sub>certainly</sub> is said to have strength  $= 1$ .

Wolf et al. (2016) predict differences in the applicability of sentences with adverbial and adjectival epistemic modifiers. In particular, Nilsen’s example (18) is explained by recognizing that there might be situations in which there are epistemic authorities in the common ground that assign the proposition a probability  $> 0$ , whereas the speaker assigns it a probability  $= 0$ .

Lassiter (2016) is an experimental investigation into the strength of different epistemic operators. The main objective of this paper is an inquiry into the strength of the modal verb *must*, in view of corpus evidence that shows that it is weaker than *know*, cf. (29).

- (29) *The speedometer shows 38,000 miles and it must be 138,000, but I don’t know for sure.*

For this purpose, Lassiter designed an experiment in which the participants were asked whether they agreed or disagreed to a sentence, after reading the following scenario:

- (30) *Yesterday, Bill bought a single ticket in a raffle with 1000 total tickets. There were also 999 other people who bought one ticket each. [...] The drawing was held last night, and the winner will be announced this evening.*

The eight sentences included the following four containing epistemic adjectives and adverbs, in addition to sentences with the epistemic predicates *must*, *might*, *know*, the sentence without any modifier, and its negation.

- (31) a. *It is possible that Bill won the raffle.*  
 b. *Bill possibly won the raffle.*  
 c. *It is certain that Bill did not win the raffle.*  
 d. *Bill certainly did not win the raffle.*

The experiment was carried out online. Each participant had to judge just one sentence, to prevent that judgements were influenced by judgments of other items. The result was that significantly more participants agreed to (31a) *possible* than to (31b) *possibly*, and that significantly fewer participants agreed to (31c) *certain not* than to (31d) *certainly not*. We can summarize this result as in (32), where  $a \ll b$  is to mean  $a$  is less “assertable” than  $b$ , skipping over the fact that *certain* and *certainly* were tested with negated prejacent clauses  $\phi$ .

- (32) a. *possibly*  $\phi \ll$  *possible*  $\phi$   
 b. *certain*  $\phi \ll$  *certainly*  $\phi$

For the purpose of this paper, let us define “assertability” as follows.

- (33) Assertability  
 Sentence  $a$  is less assertable than sentence  $b$ , i.e.  $a \ll b$ , iff in the same context, the number of people who accept  $a$  as true is smaller than the number of people who accept  $b$  as true

We avoid appealing to the notion of logical strength in our definition of assertability. While  $a \ll b$ , as defined above, might well be due in part to  $a$  being logically stronger than  $b$ , thus  $a$  being true in a subset of circumstances where  $b$  is true, we suspect that there are pragmatic factors beyond mere truth conditions that lead to the difference.

The items *must*  $\phi$  and *know*  $\phi$ , on which Lassiter focuses most, are roughly of the same assertability as *certainly*  $\phi$  and *certain*  $\phi$ , respectively. The unmodified bare sentence  $\phi$  is the least assertable.

Ricciardi et al. (2020) took up Lassiter’s discussion of epistemic *must* and carried out a similar experiment, concentrating on *know not*, *must not* and *it is certain that not*. They do not investigate the adverbial counterpart *certainly not*, but as in Lassiter’s experiment *must not* and *certainly not* showed similar results, their findings concerning the assertability of *must*  $\phi$  vs. *certain*  $\phi$  should be of interest for the current issue. However, one should be cautious about equating the use conditions for *must*  $\phi$  and *certainly*  $\phi$ , as they appear not to be completely interchangeable in contexts like (29).

Ricciardi et al. (2020) first replicate Lassiter’s findings for the three selected items. In a second experiment, they make sure by clear baselines that participants understand the task as about what they call the “literal” meaning of these expressions. This involves that each participants had to judge nine sentences with respect to the lottery scenario above, including the three target sentences involving *know not*, *must not* and *it is certain that not*. Now, the three items were rated equally, i.e., there was no difference in assertability between *must not* and *certain not*. Consequently, under the assumption that *certainly not* behaves like *must not*, we would predict that the stated fact (32b) might well not obtain under the second experimental technique.

In a third experiment, Ricciardi et al. (2020) change the scenario (30) minimally so that the last sentence is *Today, you meet Bill and he looks a bit disappointed*. This has the effect that the sentence to be rated, like *Bill must not have won the raffle*, can be interpreted as an explanation. In this situation, the difference between *must not* and *certain not* reappears. Ricciardi et al. (2020) conclude that epistemic *must* is a marker that indicates that the speaker is drawing an inference. It is an open question whether this would also hold for *certainly*, in contrast to *certain*.

We are aware of one other investigation of assertability that contrasted sentence adverbs with their adjectival variants, Müller (2019), which is on German, and does not refer to Lassiter’s work. In Müller’s Experiment 2, participants rated sentences in contexts that involve the epistemic operators *sicher* ‘certain(ly)’, *möglich(erweise)* ‘possible/possibly’ and the evidential *offensichtlich* ‘evident(ly)’. Participants showed a significantly higher agreement to sentences with the strong operators *sicher* and *offensichtlich* when they were used as sentence adverbs compared to when they were used as adjectives, corresponding to (32b). There was no significant difference with the weak operator *möglich(erweise)*, but the participants agreed slightly more to the adjectival use, which is consonant with (32a).

After having reviewed the existing experimental evidence concerning the issue of strength, or assertability, of epistemic adverbials and adjectives (and in general, of subjective and objective epistemics), we turn to explanations of these differences. As stated above, Lassiter’s main interest concerned the status of *must* (which turned out to be of a strength comparable to *certainly*). The distinction between epistemic adjectives and adverbials he describes as follows:

[Epistemic adverbials] have often been claimed to be more “subjective” than their adjectival counterparts, which have an “intersubjective” or “objective” flavor (e.g., Lyons 1977; Nuyts 2001). I conjectured that participants would be less strict in their use of subjective *possibly* and *certainly*, on the ground that a privately held certainty, or denial of possibility, may be felt to be less subject to public scrutiny and approbation if it turns out to be incorrect. If this is right, then participants should be more willing in the lottery scenario to endorse “Bill certainly did not win” than its adjectival counterpart “It is certain that Bill did not win.” Similarly, they should

be more willing to reject the subjective adverbial “Bill possibly won” on the basis of very low probability than the adjectival “It is possible that Bill won.” (p. 132)

The motivation Lassiter gives for assertability differences between *certainly* and *certain* corresponds to the discussion of the communicative role of subjective epistemics in Krifka (to appear (a)): It is easier to defend *certainly*  $\phi$ , a commitment to one’s own high subjective estimation of likelihood of  $\phi$  than a commitment to  $\phi$  itself, than the commitment *certain*  $\phi$ , that there is an objective high probability for  $\phi$ . However, it is unclear how this argument should work in the case of *possibly* and *possible*: By the same argument, it should also hold that it is easier to defend *possibly*  $\phi$  than *possible*  $\phi$ . Lassiter switches from *endorse* in the case of *certainly*  $\phi$  to *reject* in the case of *certain*  $\phi$  in his argument, and it is unclear why.

We have argued, following Krifka (to appear (a)), that epistemic adverbs and adjectives differ not only in the former being subjective, and the latter objective epistemics, but that in addition, the former are outside of the communicated proposition, whereas the latter are part of it. Hence, with *possibly*  $\phi$  the speaker wants to put  $\phi$  as a relevant option into the common ground. In contrast, with *possible*  $\phi$  the speaker may just concede that it cannot be excluded that the proposition  $\phi$  is true. Under this additional assumption, the observations of Lassiter follow: *possibly*  $\phi$  is less assertable under a low likelihood because in addition to saying that  $\phi$  is possible, the speaker would also communicate  $\phi$  as an option to be considered.

Before we turn to our experiments, let us return to the finding of Ricciardi et al. (2020), that the assertability of *must* (which was of similar assertability as *certainly* in Lassiter’s experiment) aligns with *certain* (and *know*). This can be explained by the assumption that it is difficult for speakers to publicly commit to differences between subjective and objective likelihood estimations, that is, between one’s own estimation of likelihood and the one by an “expert” epistemic source. In the second experiment of Ricciardi et al. (2020), as well as in the tests of conjunctions like (18) by Wolf et al. (2016), participants were directly confronted with such statements and adapted their interpretation of subjective epistemics to the interpretation of objective epistemics. Our conclusion, with respect to experimental technique, is opposite to Ricciardi et al. (2020): In order to investigate how epistemic adverbials and epistemic adjectives differ from each other, it is important to *prevent* participants in experiments from aligning their judgements between these categories.<sup>8</sup>

---

<sup>8</sup> The second finding of Ricciardi et al. (2020), that *must*  $\phi$  is more assertable than *certain*  $\phi$  in contexts that invite the drawing of inferences in order to provide explanations, can be explained by the fact that *must* is outside of the at-issue part of the sentence and hence does not address the question under discussion, cf. Herbsttritt (2020). The sentence, *Today, you meet Bill and he looks a bit disappointed* plausibly gives rise to the question, ‘Why does Bill look disappointed?’, for which the proposition ‘Bill did not win the raffle’, which is the prejacent proposition of (31d), provides a better answer than the proposition ‘It is certain that Bill did not win the raffle’, which is expressed by (31c). It remains to be seen whether *certainly* behaves similarly to *must* in this respect.

## 2 Research questions, hypotheses, and experiments

After having provided an overview of the known differences between epistemic adjectives and adverbials and theoretical assumptions to explain these differences, we want to explore these expressions further. In particular, we want to investigate the assertability differences (32) further.

Our research question is inspired by Lassiter (2016), and so are our methods. However, in the first experiment, we diverge from his original set-up in two ways for the following reasons.

First, Lassiter's experimental scenario made it necessary to test the assertability of *certain* and *certainly* with negated preadjacent clauses, cf. (31c) and (31d).

Negation adds complexity, and might have an influence on the experimental results.<sup>9</sup> We therefore changed the scenario in order to circumvent the use of negation by manipulating the number of raffle tickets purchased by the main protagonist.

Our second goal was to investigate epistemic adverbs and adjectives in a scenario in which the probability is not given in a quantitative, measurable setting, like the lottery scenario of Lassiter. It is true that scenarios that use quantitative probabilities have the advantage of providing objective information about a state of affairs that can be intersubjectively verified, and they offer an easy way to vary the objective likelihood. For this reason, they make excellent stimuli for experiments about the use of epistemic and probability expressions. Without doubt, this is the reason why the overwhelming majority of experimental work in psychology and linguistics has relied on them. However, when investigating subjective epistemic operators like *certainly* and *possibly*, it is important to see how these operators are used in situations in which the underlying probabilities cannot be objectively measured, and whether they differ from objective epistemic operators like *certain* and *possible*. We hypothesized that the differences between subjective epistemic adverbials and objective epistemic adjectives should be more pronounced in the scenarios with non-measurable probabilities, as the expression of subjective probabilities is less constrained by objectively given information. Specifically our hypotheses predict an interaction in assertability between two types of measurability and operator types (adjectival vs. adverbial).

Note that we do not subscribe to any particular theory of *certain*, *certainly*, *possible*, and *possibly*. Our empirical study is not aimed at adjudicating between specific semantic analyses of these items. As can be seen below, such an analysis is not needed for the formulation of our research hypotheses, nor is it needed for the interpretation and evaluation of the experimental

---

<sup>9</sup> Negative polarity has been found to cause processing difficulties (cf. Wason 1959; 1961; Just & Carpenter 1971; Deschamps et al. 2015).

results. We will therefore abstain from proposing formal semantic representations of the modal adjectives and adverbs in question.

### 3 Experiment 1: mathematically measurable vs. non-measurable

As discussed in the previous section, our experimental goals are as follows: (i) investigate whether there is a difference between the adjectival and adverbial operators in terms of the threshold of probability compatible with the expression, and (ii) whether mathematical measurability of probabilities has an effect on this threshold.

Our research hypotheses are shown in (34) and (35).<sup>10</sup>

- (34) Low probabilities
- a. The adverb *possibly* leads to a lower level of agree responses than the adjective *possible* in the low probability conditions.
  - b. The assertability difference between *possible* vs. *possibly* is greater in the non-measurable low condition than in the measurable low condition.
- (35) High probabilities
- a. The adverb *certainly* leads to a higher level of agree responses than the adjective *certain* in the high probability conditions.
  - b. The assertability difference between *certain* and *certainly* is greater in the non-measurable high than measurable high condition.

Note that the hypothesis (34a) corresponds to (32a) and the hypothesis (34a) correspond to (32b). Hypotheses (34b) and (35b) are original to this paper.

Let us say a few words about these hypotheses before moving on. What (34b) and (35b) say is that subjectivity of probability is directly related to assertability difference: The more difficult it is to calculate the probability of  $\phi$ , the larger the gap is between *certain*  $\phi$  and *certainly*  $\phi$  as well as between *possible*  $\phi$  and *possibly*  $\phi$ , with respect to the number of agree responses. Now, why should this and not the opposite be our hypothesis? Why do we not say that subjectivity is *inversely* related to assertability difference, i.e., that the more subjective the probability of  $\phi$  is, the less the adjectives and the adverbs differ? It should be noted that the point of the experiment is to test whether there is a difference, and if there is one, what it is. Thus, it really makes little difference whether our hypothesis is the “directly related” or the “inversely related” one. However, we decided on “directly related” because of the following consideration: It seems very intuitive to us that what is objective must be subjective as well. Thus, if one is confronted with an objective result of a mathematical calculation, it would be impossible not to subjectively accept

---

<sup>10</sup> The experiment was pre-registered at OSF (<https://osf.io/sxchg>). The norming study and a pilot study for the present study were also pre-registered at OSF (<https://osf.io/bw83q>).

it also.<sup>11</sup> However, it is quite natural for a subjective estimation not to be objective. This means that in the measurable conditions, i.e., those scenarios where the relevant probability can be determined mathematically, there is a tendency for the subjective and the objective to collapse into one. Thus, we expect the difference between the adverbs, which express subjectivity, and the adjectives, which express objectivity, to be smaller in these scenarios.

### 3.1 Procedure

The experiment was conducted in the following order:<sup>12</sup>

- (36)
- a. Participants were first told that they would read a short text and answer some questions about it.
  - b. They then read the context story.
  - c. After the context story was provided, one of the test sentences was shown below the context story.
  - d. Participants were asked whether they agree or disagree with the statement by clicking on a button for Agree or Disagree, placed side-by-side, while the context story and the prompt for a response were still visible. After the response, the page cleared.
  - e. Comprehension questions and demographic questions were then asked.

Each participant saw one of the continuations, followed by only one of the test sentences, either with an adjective or an adverb operator. For example, a participant saw only the high probability continuation and read only the test sentence with an adjective operator.

There were three comprehension questions that each participant was required to respond to check their attention. The comprehension questions are shown in Appendix B.

### 3.2 Material

We created two conditions, one with measurable and the other with non-measurable probabilities, which we discuss below. In order to compare the results with non-measurable and measurable conditions directly, we needed the perceived probability for the non-measurable condition to be around the same value as the probability for the measurable condition. As it is not possible to measure the probability for the non-measurable condition, we conducted a norming study in which the task for the participant was to provide a numerical probability/likelihood (between 0 and 100%) that the proposition (test sentence for the main study) is true (see Appendix A for

---

<sup>11</sup> Note that this intuition is in line with Lewis' "Principal Principle" which states, essentially, that subjective probability is subjective expectation of objective chance (Lewis 1980: 266–267). We thank two anonymous reviewers for making us aware of this connection.

<sup>12</sup> The experiments reported in this paper have been approved by Ethics committee at the German linguistic society (Deutsche Gesellschaft für Sprachwissenschaft) in January, 2020 (#2018-05-200124).



details). Based on the norming study, we chose 95% and 10% as probabilities for the measurable, objective condition, diverging from Lassiter's 99.9% and 0.1%.<sup>13</sup>

### 3.2.1 Mathematically non-measurable conditions

For the mathematically non-measurable conditions, we created the following short sketch. Two lists of evidence followed the set-up story, one set of evidence suggesting that the main protagonist of the story is highly likely to be guilty ((38)), while the other suggests that he is not likely to be guilty ((39)).

(37) *Set-up*

A murder took place on a yacht in the middle of the Atlantic Ocean. The victim was stabbed. The police concluded that the murderer must have been one of the passengers on the yacht. Jay is one of the passengers. The amount of information the police has gotten so far is as follows:

(38) *High probability continuation:*

- a. Jay was seen having a heated conversation with the victim on the yacht shortly before the murder.
- b. The murder was most likely committed by a left-hander. Jay is left-handed.
- c. The murder weapon was found in Jay's cabin.
- d. Jay's fingerprints were the only ones found at the crime scene.
- e. Jay stands to inherit a large sum of money from the victim.

(39) *Low probability continuation:*

- a. Jay was seen having a friendly conversation with the victim on the yacht shortly before the murder.
- b. The murder was most likely committed by a left-hander. Jay is right-handed.
- c. The murder weapon was found in the cabin of another passenger, not Jay's.
- d. Finger prints from several passengers, including Jay, were found at the crime scene.
- e. Jay does not stand to benefit from the victim's death.

**Prompt for a response:**

(40) Please indicate whether you agree or disagree with the following statement:

---

<sup>13</sup> An anonymous reviewer raises the issue that Lassiter's 99.9% and 0.1% are closer to the literal meaning of the universal and the existential modal than our 95% and 10%, and therefore, that our norming study might be "not enough." We agree that 99.9% is, intuitively, closer to 'all', but we are not sure in what sense 0.1% is closer to 'some' than 10%. It would of course be interesting if we could reach 99.9% and 0.01% in our norming study, so we have an exact non-measurable counterpart of Lassiter's measurable scenario. However, this was not feasible. Nevertheless, we do not think the norming study result is of no use. With it, we can compare measurable and non-measurable scenarios at non-extreme values, and compare measurable scenarios with non-extreme values to those with extreme ones. This is what we did with our experiments.

**Test sentences:**

- (41) *The high probability continuation:*
- a. It is certain that Jay committed the murder.
  - b. Jay certainly committed the murder.
- (42) *The low probability continuation:*
- a. It is possible that Jay committed the murder.
  - b. Jay possibly committed the murder.

**3.2.2 Mathematically measurable conditions**

For the mathematically measurable probabilities, we adapted the experimental design in Lassiter (2016) for the reasons described in the previous section. We manipulated the number of raffle tickets that the main protagonist purchases in order to (i) avoid having a negation in the prejacant and (ii) change the numerical probabilities from 99.9% to 95%, and 0.1% to 10%, as shown below.

- (43) *Set-up*  
The elementary school held a raffle to raise money for student activities. A total of 1000 tickets were sold. Of those, ...
- (44) a. *95% probability continuation*  
950 tickets were purchased by Jay, a wealthy local businessperson. 50 tickets were purchased by other members of the community.
- b. *10% probability continuation*  
100 tickets were purchased by Jay, a wealthy local business person. 900 tickets were purchased by other members of the community.

*Prompt for a response:*

- (45) The drawing was held last night, and the winner is about to be announced.  
Please indicate whether you agree or disagree with the following statement:

**Test sentences:**

- (46) *95% probability:*
- a. It is certain that Jay won the raffle.
  - b. Jay certainly won the raffle.
- (47) *10% probability:*
- a. It is possible that Jay won the raffle.
  - b. Jay possibly won the raffle.

### 3.2.3 Participants

1659 participants were recruited on the Amazon MechanicalTurk platform. There were eight lists (2 (probability)  $\times$  2 (operator type)  $\times$  2 (measurability)) for this experiment.

Approximately 200 participants took part in each list. Each participant participated in only one list, and hence, read only one of the experimental sentences in one of the experimental conditions. When the participant took part in more than one list, we kept only the response from the first participation by the participant, excluding the subsequent responses on different lists.

We used the following exclusion criteria:

- (48)
- a. Data from non-native speakers of English were excluded;
  - b. One participation per participant: When they took part in more than one list, only one response per participant (the first response) was included in the analysis
  - c. Responses for the three comprehension questions: Participants who responded incorrectly two or more times were excluded.

The numbers of remaining participants for each list is shown in **Table 1**.

|          | measurable      |                 |                |                  | non-measurable  |                 |                |                  |
|----------|-----------------|-----------------|----------------|------------------|-----------------|-----------------|----------------|------------------|
| operator | <i>possible</i> | <i>possibly</i> | <i>certain</i> | <i>certainly</i> | <i>possible</i> | <i>possibly</i> | <i>certain</i> | <i>certainly</i> |
| N        | 140             | 150             | 107            | 118              | 187             | 115             | 155            | 154              |

**Table 1:** Number of participants for each list.

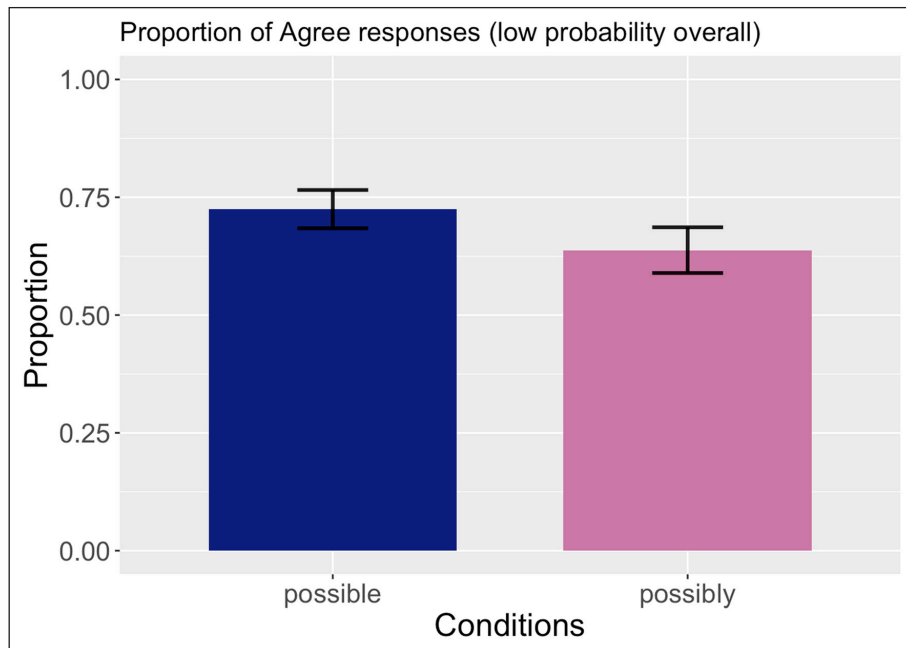
### 3.3 Results

We discuss the results in relation to the research hypothesis (34) and (35). Let us start with a comparison between *possible* and *possibly*. Throughout this paper, we analyze the data from the experiment by fitting a generalized linear model (glm) using the lme4 package in R (Bates et al. 2015).<sup>14</sup> We will specify which factors we used as the fixed effects for each model. The dependent variable was always the response type (*agree* or *disagree*). The analysis plan was pre-registered at OSF before data collection (see footnote 10).

We first tested whether there is a difference in overall proportions of **agree** responses between *possible* and *possibly*. We fitted a generalized linear model with a response type (*agree* or *disagree*) as the dependent variable and operator type (*possible* or *possibly*) as the fixed predictor, corresponding to **Figure 1**. The descriptive statistics are shown in **Table 2**, with the confidence

---

<sup>14</sup> We had pre-registered our analysis plan at osf as using  $\chi^2$  for comparisons without interaction, and using glm to check for interactions. We however used glm for all the comparisons, including those that do not involve interactions. The results of the analysis, however, do not differ.



**Figure 1:** Both measurabilities combined.

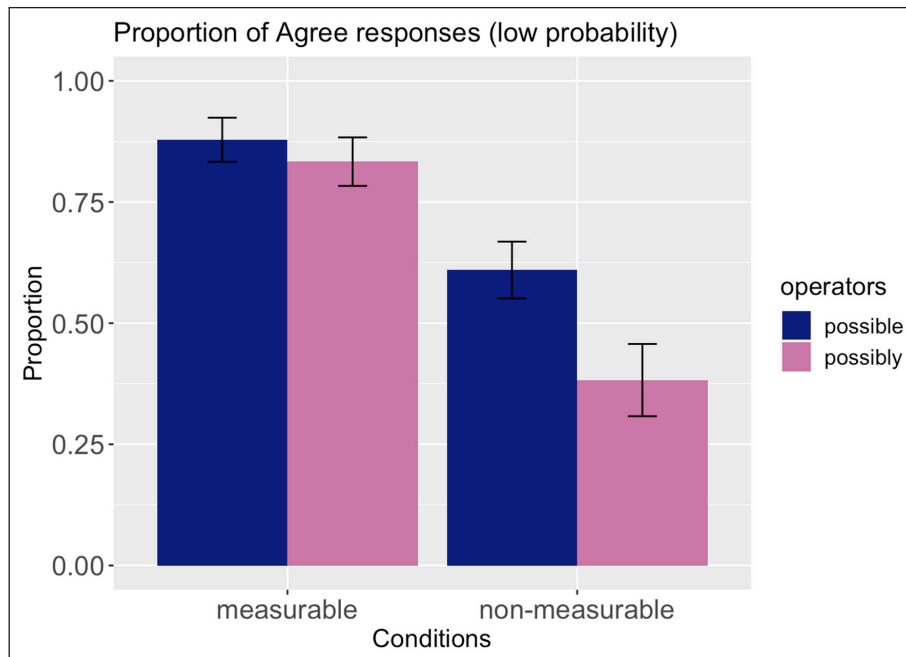
| operator                    | possible | possibly |
|-----------------------------|----------|----------|
| <i>N</i> of participants    | 327      | 265      |
| <i>N</i> of agree responses | 237      | 169      |
| % of agree                  | 72.5     | 63.8     |
| upper confidence interval   | 76.5     | 68.6     |
| lower confidence interval   | 68.4     | 58.9     |

**Table 2:** Comparison between *possible* and *possibly*.

interval set at 95%.<sup>15</sup> The result shows that the hypothesis in (34a) – *possible* >> *possibly* – is supported (z-value: -2.263,  $p < .05$ ).

Next we separate the data by measurability. The descriptive statistics are shown in **Table 3**, and the **Figure 2** visualizes the proportion of *agree* responses for each list. We again fitted generalized linear models, this time with the operator type (*possible* or *possibly*) as the fixed effect, separately for two measurabilities. We found a significant effect of operator type for the non-measurable context (z-value: -3.796,  $p < .01$ ) but not for the measurable context (z-value: -1.090,  $p = 0.276$ ).

<sup>15</sup> In all tables in this paper, when the confidence interval is indicated, it is set at 95%.



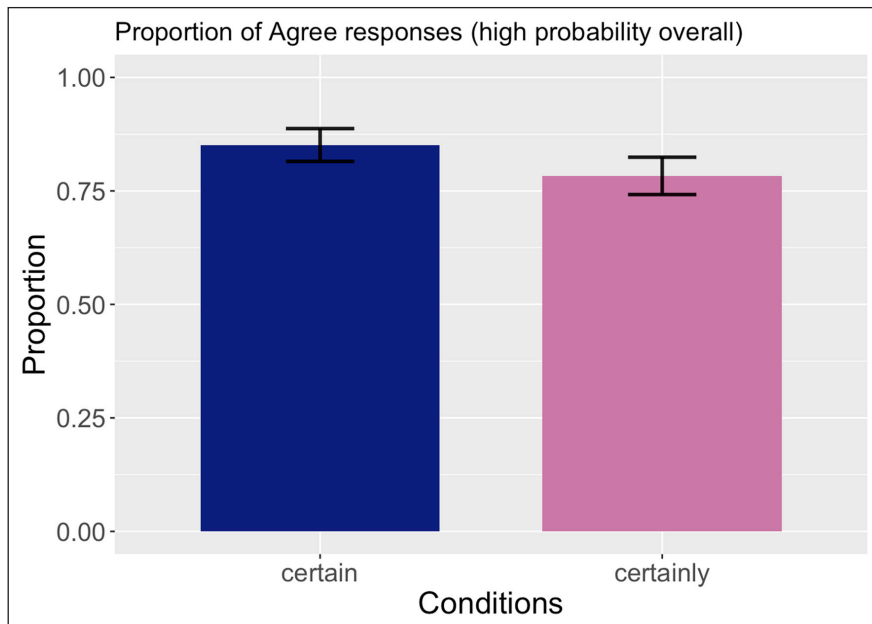
**Figure 2:** Conditions separated.

| Measurability               | measurable |          | non-measurable |          |
|-----------------------------|------------|----------|----------------|----------|
|                             | possible   | possibly | possible       | possibly |
| <i>N</i> of Participants    | 140        | 150      | 187            | 115      |
| <i>N</i> of agree responses | 123        | 125      | 114            | 44       |
| % of agree                  | 87.9       | 83.3     | 61.0           | 38.3     |
| upper confidence interval   | 92.4       | 88.3     | 66.8           | 45.7     |
| lower confidence interval   | 83.3       | 78.3     | 55.1           | 30.8     |

**Table 3:** Comparison between *possible* and *possibly*.

In a model with both operator type and measurability as fixed effects and their interaction, we found that the measurability ( $z: -5.127, p < .01$ ) was significant predictors, but there was no significant interaction ( $z: -1.329, p = .184$ ). However, our results point towards the direction of (34b): The difference in assertability between *possible* and *possibly* is greater in the non-measurable context (22.7%) than in the measurable context (4.6%).

Let us now turn to comparisons between *certain* and *certainly*. The overall prediction was that we obtain more agree responses with the adverbial operator *certainly* than with the adjectival operator *certain*, (35a). As can be seen on **Table 4**, visualized in **Figure 3**, this prediction was not supported. The result is contrary to what we expected: Slightly, but significantly more



**Figure 3:** Both conditions together.

| operator                  | certain | certainly |
|---------------------------|---------|-----------|
| N of participants         | 262     | 272       |
| N of agree responses      | 223     | 213       |
| % of agree                | 85.1    | 78.3      |
| upper confidence interval | 88.7    | 82.4      |
| lower confidence interval | 81.5    | 74.2      |

**Table 4:** Comparison between *certain* and *certainly*.

participants agreed with the test sentence with *certain* than the one with *certainly* (with the operator type (*certain* vs. *certainly*) as the fixed predictor for the data from both measurabilities combined: z-value:  $-2.021$ ,  $p < .05$ ).

We have an even more striking result when we separate the measurable and non-measurable conditions. The descriptive statics are shown in **Table 5**, and the proportions of *agree* responses are visualized in **Figure 4**. In the measurable condition, we observe that *certain* received significantly higher proportion of agree responses than *certainly*, the opposite to what was predicted in (35a). The generalized linear models fitted with the operator type (*certain* or *certainly*) as the fixed effect for each measurability condition separately show that there is an effect of operator type for the measurable condition (z-value:  $-3.046$ ,  $p < .01$ ) but not for the non-measurable condition (z:  $0.125$ ,  $p = 0.9$ ).

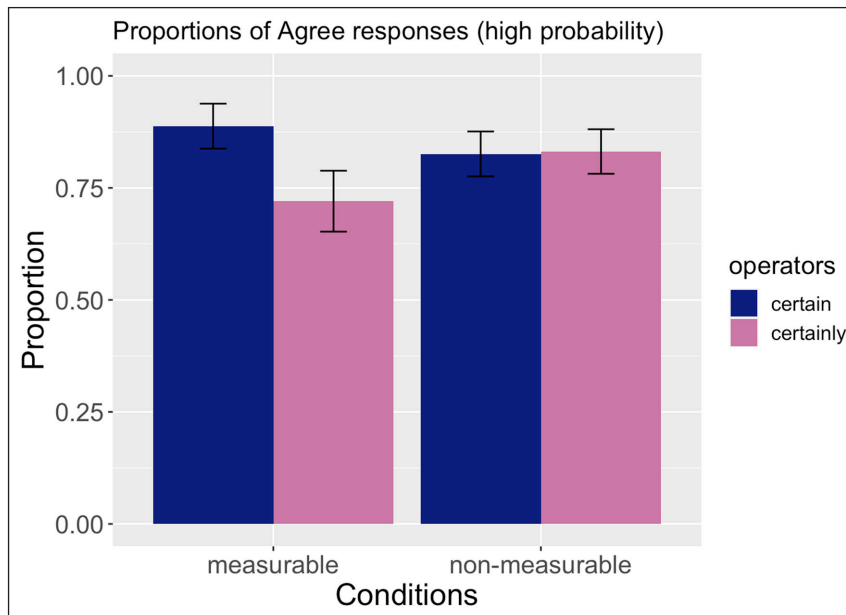


Figure 4: Conditions separated.

| Measurability               | Measurable |           | Non-measurable |           |
|-----------------------------|------------|-----------|----------------|-----------|
|                             | certain    | certainly | certain        | certainly |
| <i>N</i> of participants    | 107        | 118       | 155            | 154       |
| <i>N</i> of agree responses | 95         | 85        | 128            | 128       |
| % of agree                  | 88.8       | 72.0      | 82.6           | 83.1      |
| upper confidence interval   | 94.8       | 80.1      | 88.6           | 89.0      |
| lower confidence interval   | 82.8       | 63.9      | 76.6           | 77.2      |

Table 5: Comparison between conditions.

A model with both the operator type (*certain* or *certainly*) and the measurability type (*measurable* or *non-measurable*) as the fixed effects revealed that the operator type is a significant predictor (z-value:  $-2.002$ ,  $p < .05$ ). There was an interaction between the measurability and the operator type (z-value:  $2.436$ ,  $p < .05$ ) due to the unexpectedly high agree responses for *certain* compared to *certainly* in the measurable condition. The prediction of hypothesis (35b) is clearly not supported either: The difference in assertability between *certain* and *certainly* is greater in the measurable condition than in the non-measurable condition.

### 3.4 Interim discussion

Summarizing our findings, we observed that, overall, *possible* is more assertable than *possibly*, supporting hypothesis (34a). When we consider measurable and non-measurable conditions



separately, however, we see a significant difference only in the non-measurable condition. While statistical analysis, using generalized linear models, did not show interaction between the operator type and the measurability, our results show the pattern predicted by our hypothesis: A larger difference between the two operators in the non-measurable condition compared to that in the measurable condition.

The results from the comparison between *certain* and *certainly*, on the other hand, were unexpected. We received more agree responses with *certain* than *certainly* in the measurable probability condition, contrary to (35a), and also contrary to Lassiter’s result. Also, (35b) was rejected, as the difference in assertability between *certain* and *certainly* was actually smaller in the non-measurable condition than in the measurable condition.

The predicted and obtained differences in assertabilities for each adjective/adverb pairs are schematized in **Table 6**. As explained above in Section 1.2.2,  $a \gg b$  indicates the direction of assertability, where the assertability is defined by the number of people who accept the use of the operator:  $a \gg b$ , if  $a$  as true is bigger than the number of people who accept  $b$  as true. Recall that we predicted the difference to be larger in the non-measurable condition than in the measurable condition.

| Predicted direction               | possible $\gg$ possibly | certain $\ll$ certainly |
|-----------------------------------|-------------------------|-------------------------|
| Actual direction (measurable)     | possible = possibly     | certain $\gg$ certainly |
| Actual direction (non-measurable) | possible $\gg$ possibly | certain = certainly     |

**Table 6:** Summary of predicted & observed assertability differences.

We are left with the following questions:

- (49) a. Why could we not replicate Lassiter’s results for the measurable conditions for both high and low probabilities?  
 b. Why was the acceptability difference between *certain* vs. *certainly* not greater in the non-measurable high condition than in the measurable high condition?

We will comment on the second question in the General Discussion but will leave it for future research. We address the first question in Section 4, concentrating on the measurable probability conditions.

## 4 Experiment 2

The results from experiment 1 contrasted with Lassiter’s results in that we did not find the effect of operator type for *possible* vs. *possibly*, and found the effect of operator type but the direction was opposite for *certain* vs. *certainly*. This is interesting because with the low probability conditions overall and with the non-measurable probability condition, in particular, we find the expected

effect of the operator type. The goals of this section is to discuss our second experiment, which extends our Experiment 1 to investigate what affected the assertability in the Experiment 1.

Our experimental conditions and test sentences for Experiment 1 diverged from that of Lassiter's in two domains, namely, the use of less extreme probabilities for both high and low probability conditions, and the removal of negation from the experimental sentence for the high probability conditions.

We assume that the reason we did not obtain the expected effect of operators, which was obtained in Lassiter's original experiment, was due to the changes we have made. To verify this, we conducted the second experiment with the following three goals: (i) the replication of Lassiter's experiment; (ii) an investigation of the role of more extreme probabilities; and (iii) an investigation of the role of negation.

Below, we will make comparisons between conditions where probabilities are kept constant while polarity of the test sentences are contrasted (positive vs. negative), and conditions where polarities are kept constant while the probabilities are contrasted (99.9% vs. 95%, and 0.1% vs. 10%).

In order to examine the effect of negation, on the one hand, and the role of extreme probabilities, on the other, we created five conditions, shown in **Table 7**.

| probability | polarity                                   |
|-------------|--|
| 99.9%       | negative (replication of Lassiter's study) |
| 99.9%       | positive                                   |
| 95%         | negative                                   |
| 0.1%        | positive (replication of Lassiter's study) |
| 10%         | negative                                   |

**Table 7:** Conditions examined.

Assuming that the factors that we consider affected the experimental results for experiment 1, not the specific experimental set-up we used, we investigated the following experimental hypotheses with the Experiment 2:

- (50) High probabilities:
- a. When the proposition does not contain a negation, we obtain fewer agree answers for *certain* than *certainly* in the 99.9% condition, with this effect not present in 95% condition (results from Experiment 1).
  - b. We obtain fewer agree answers for *certain* than *certainly* at the 95% negative condition, with this effect not present in the 95% positive condition.
  - c. At 99.9%, we obtain fewer agree answers for *certain* than *certainly* with a negative proposition (replication of Lassiter's result).

(51) Low probabilities:

- a. When the proposition does not contain negation, we obtain fewer agree answers for *possibly* than *possible* in the 0.1% condition, with this effect not present in the 10% condition (results from Experiment 1).
- b. When the proposition contains a negation, we obtain fewer agree answers for *possibly* than *possible* in the 10%, with this effect not present in the 10% positive condition.

## 4.1 Participants

2049 participants in total were recruited on the Amazon MechanicalTurk platform. There were 10 lists in total for this experiment.

We used the same exclusion criteria as the first experiment: (i) self-declared 52 non-native speakers of English were excluded from the analysis; (ii) when they took part in more than one list, only one response per participant (the first response) was included in the analysis; (iii) at least two out of three comprehension questions needed to be answered correctly.

The numbers of remaining participants for each list is shown in **Table 8** and **Table 9**.

|                   | 99.9%          |                  | 95%            |                  | 99.9%   |           |
|-------------------|----------------|------------------|----------------|------------------|---------|-----------|
|                   | certain $\neg$ | certainly $\neg$ | certain $\neg$ | certainly $\neg$ | certain | certainly |
| N of participants | 125            | 131              | 100            | 91               | 99      | 149       |

**Table 8:** Number of participants: high probabilities.

|                   | 10%             |                 | 0.1%     |          |
|-------------------|-----------------|-----------------|----------|----------|
|                   | possible $\neg$ | possibly $\neg$ | possible | possibly |
| N of participants | 98              | 110             | 147      | 109      |

**Table 9:** Number of participants: low probabilities.

## 4.2 Material

The set-up story, the prompt for a response, the test sentences, and the participants' task were identical to those for the measurable conditions from Experiment 1. In order to get the target probability, we manipulated the number of tickets the protagonist bought, as shown below.<sup>16</sup>

(52) a. *99.9% probability with negation*

1 ticket was purchased by Jay, a wealthy local businessperson. 999 tickets were purchased by other members of the community.

<sup>16</sup> As with Experiment 1, the experimental design and the analysis plan was pre-registered at OSF, viewable at <https://osf.io/6w9na>.

- b. *99.9% probability positive*  
999 tickets were purchased by Jay, a wealthy local businessperson. 1 ticket was purchased by a member of the community.
  - c. *95% probability with negation*  
50 tickets were purchased by Jay, a wealthy local businessperson. 950 tickets were purchased by other members of the community.
- (53) a. *10% probability with negation*  
900 tickets were purchased by Jay, a wealthy local businessperson. 100 tickets were purchased by other members of the community.
- b. *0.1% probability positive*  
1 ticket was purchased by Jay, a wealthy local businessperson. 999 tickets were purchased by other members of the community.

Test sentences were as follows:

- (54) High probability, positive conditions:
- a. It is certain that Jay won the raffle.
  - b. Jay certainly won the raffle.
- (55) High probability, negative conditions:
- a. It is certain that Jay didn't win the raffle.
  - b. Jay certainly didn't win the raffle.
- (56) Low probability, positive conditions:
- a. It is possible that Jay won the raffle.
  - b. Jay possibly won the raffle.
- (57) Low probability, negative conditions:
- a. It is possible that Jay didn't win the raffle.
  - b. Jay possibly didn't win the raffle.

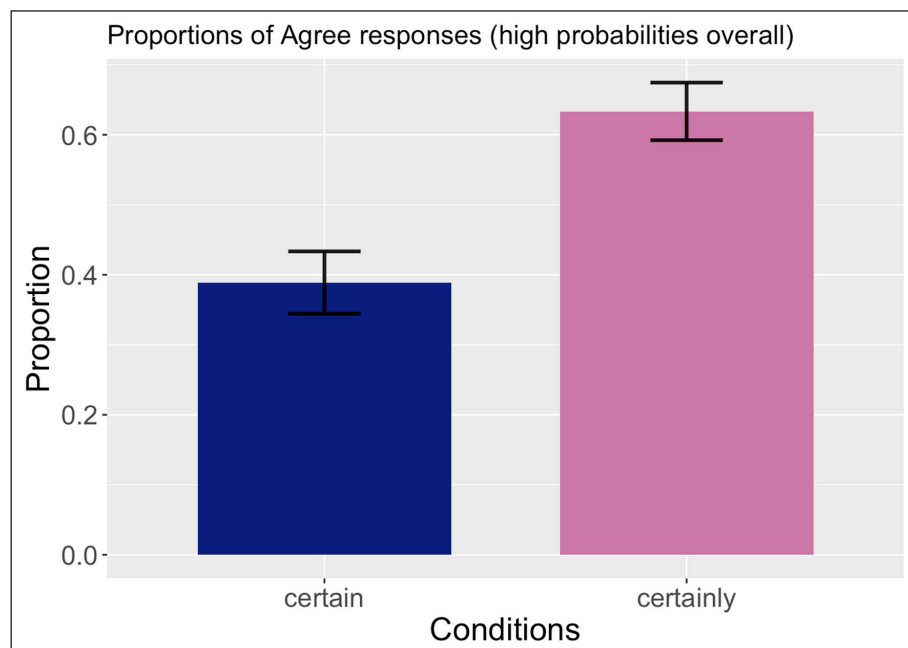
### 4.3 Certain vs. certainly

We first discuss comparisons for the high probability conditions, which resulted in a surprising pattern in the first experiment. We test the hypotheses in (50), which are repeated here as (58).

- (58) a. When the proposition does not contain a negation, we obtain fewer agree answers for *certain* than *certainly* in the 99.9% condition, with this effect not present in 95% condition (results from Experiment 1).
- b. We obtain fewer agree answers for *certain* than *certainly* at the 95% negative condition, with this effect not present in the 95% positive condition.

- c. At 99.9%, we obtain fewer agree answers for *certain* than *certainly* with a negative proposition (replication of Lassiter’s result).

In stark contrast to the results from Experiment 1, we do find a significant effect of operator type across Experiment 2 conditions, in the direction our hypothesis predicts (*certain* < < *certainly*), as can be seen in **Figure 5**. The descriptive statistics are shown in **Table 10**.



**Figure 5:** *certain* vs. *certainly* overall.

| operator                    | certain | certainly |
|-----------------------------|---------|-----------|
| <i>N</i> of participants    | 324     | 371       |
| <i>N</i> of agree responses | 126     | 235       |
| % of agree                  | 38.9    | 63.3      |
| upper confidence interval   | 43.3    | 67.5      |
| lower confidence interval   | 34.4    | 59.2      |

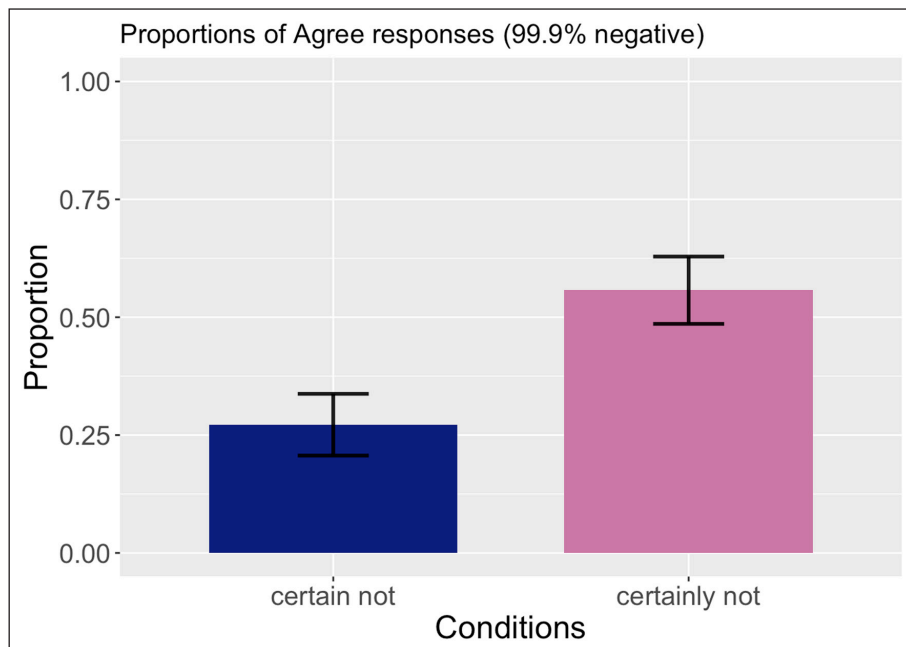
**Table 10:** Comparison between *certain* and *certainly*.

Generalized linear models with operator type (adjective vs. adverb) as the fixed predictor finds the operator type to be a significant effect (z-value: 6.369,  $p < .01$ ).

Let us now examine each condition.

### 4.3.1 Results: 99.9% probability with negation

We first discuss the replication of Lassiter’s conditions with the operators *certain* and *certainly*, at 99.9% probability, and with a negation in the prejacent (hypothesis in (50c)/(58c)). As the descriptive statistics in **Table 11** and **Figure 6** show, we observe that the proportion of *Agree* responses are higher when the test sentence contains *certainly not*, rather than *certain not*, as expected by our hypothesis. Our results from this condition closely mirror those from Lassiter’s experiment, where the proportion of agree responses was 25% for *certain not* with confidence interval 15%–35%, and 54% for *certainly not* with confidence interval 42%–67%. We fitted a model with an operator type as the random effect, which shows that the operator type is a significant predictor ( $z: 4.547, p < .01$ ).



**Figure 6:** 99.9%: *certain not* vs. *certainly not*.

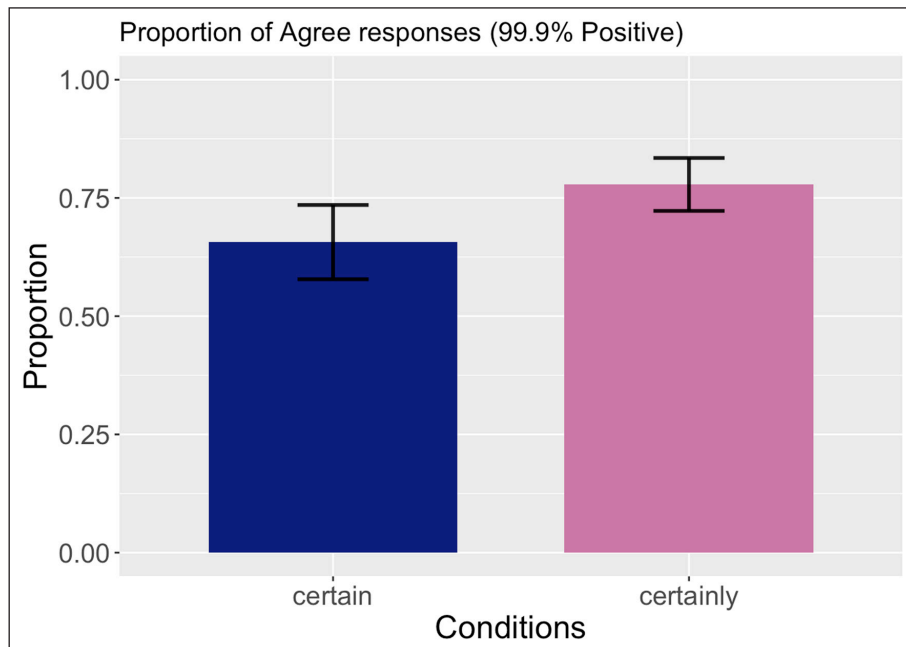
| operator                  | <i>certain</i> $\neg$ | <i>certainly</i> $\neg$ |
|---------------------------|-----------------------|-------------------------|
| N of participants         | 125                   | 131                     |
| N of agree responses      | 34                    | 73                      |
| % of agree                | 27.2                  | 55.7                    |
| upper confidence interval | 33.7                  | 62.9                    |
| lower confidence interval | 20.7                  | 48.6                    |

**Table 11:** Results of 99.9% negative condition.

Results from this condition show that what we found in Experiment 1 – that *certain* was more assertable than *certainly* – must be due to the changes we have made in comparison to Lassiter’s design.

#### 4.3.2 Results: 99.9% positive condition

Let us next consider the condition in which the polarity is changed from the previous condition. As we see in **Table 12** and **Figure 7**, we obtained a higher proportion of *agree* responses, even when the sentences did not contain negation. The generalized linear models, fitted with the operator type (adjective vs. adverb) confirms that the operator type is a significant predictor (z-value: 2.105,  $p < 0.05$ ).



**Figure 7:** 99.9% positive: *certain* vs. *certainly*.

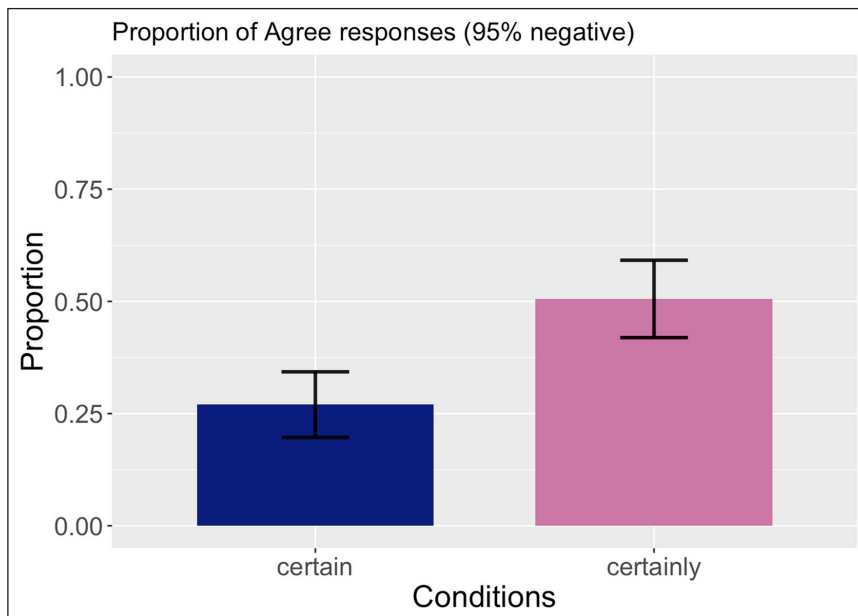
| operator                    | <i>certain</i> | <i>certainly</i> |
|-----------------------------|----------------|------------------|
| <i>N</i> of participants    | 99             | 149              |
| <i>N</i> of agree responses | 65             | 116              |
| % of agree                  | 65.7           | 77.9             |
| upper confidence interval   | 73.6           | 83.4             |
| lower confidence interval   | 57.8           | 72.3             |

**Table 12:** Results of 99.9% positive condition.



### 4.3.3 Results: 95% negative condition

The third new condition lowers the probability from Lassiter’s test sentences from 99.9% to 95%, keeping the polarity constant (negative). As we see on **Table 13**, visualized in **Figure 8**, the effect of operator type is visible, and as expected, there were higher proportion of agree responses with test sentence containing the adverbial operator. The generalized linear models, fitted with the operator type as the fixed effect predictor, confirms that the operator type has a significant effect on participants’ response patterns (z-value: 3.304,  $p < .01$ ).



**Figure 8:** 95% negative: *certain not* vs. *certainly not*.

| operator                  | <i>certain</i> $\neg$ | <i>certainly</i> $\neg$ |
|---------------------------|-----------------------|-------------------------|
| N of participants         | 100                   | 91                      |
| N of agree responses      | 27                    | 46                      |
| % of agree                | 27.0                  | 50.5                    |
| upper confidence interval | 34.3                  | 59.2                    |
| lower confidence interval | 19.7                  | 41.9                    |

**Table 13:** Results of 99.9% negative condition.

The summary of the predicted and observed differences in assertability is shown in **Table 14**. As before,  $a << b$  indicates the direction of assertability: b is more assertable than a, where more assertable means more “agree” responses.

| Predicted: <i>certain</i> < < <i>certainly</i> |          |                                     |
|--|----------|-------------------------------------|
| probability                                    | polarity | direction                           |
| 99.9%  | negative | <i>certain</i> < < <i>certainly</i> |
| 99.9%  | positive | <i>certain</i> < < <i>certainly</i> |
| 95%  | negative | <i>certain</i> < < <i>certainly</i> |

**Table 14:** Summary of predicted & observed assertabilities (*certain* vs. *certainly*).

As can be seen, now the direction of observed assertability difference aligns with that of predicted one for all conditions.

#### 4.3.4 Comparisons: effect of negation and level of probabilities

Now we are in the position to discuss the two remaining hypotheses ((50a) and (50b)), repeated in (59).

- (59) a. When the proposition does not contain negation, we obtain fewer agree answers for *certain* than *certainly* in 99.9% condition, with this effect not present in 95% condition (Experiment 1).
- b. We obtain fewer agree answers for *certain* than *certainly* at 95% negative condition, with this effect not present in 95% positive condition (Experiment 1).

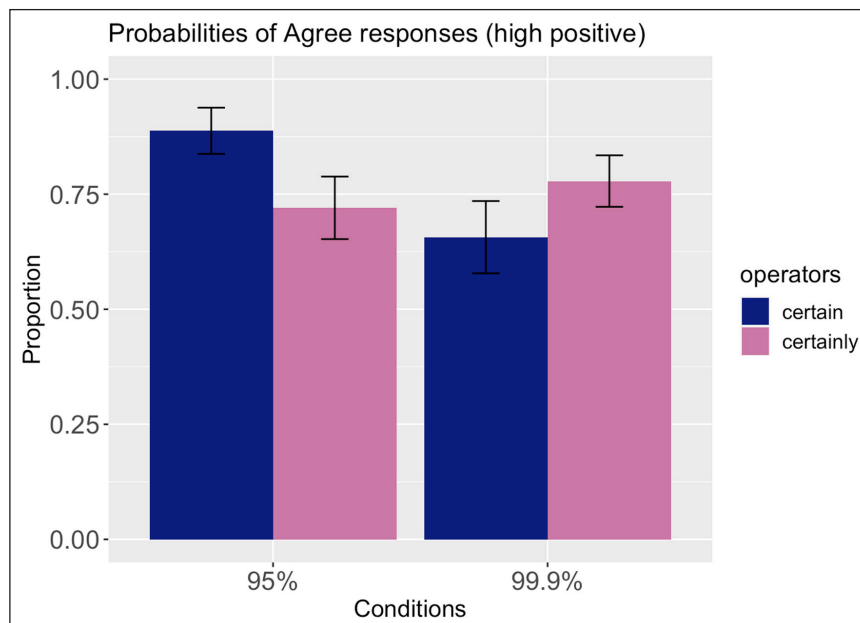
In order to investigate the potential effect of polarity and that of extreme probabilities, we perform the following seven comparisons.

- (60) a. positive: 99.9% vs. 95%
- b. negative: 99.9% vs. 95%
- c. 99.9%: positive vs. negative
- d. 95%: positive vs. negative
- e. total (99.9% + 95%): positive vs. negative
- f. total (positive + negative): 99.9% vs. 95%
- g. total: polarity and extremeness

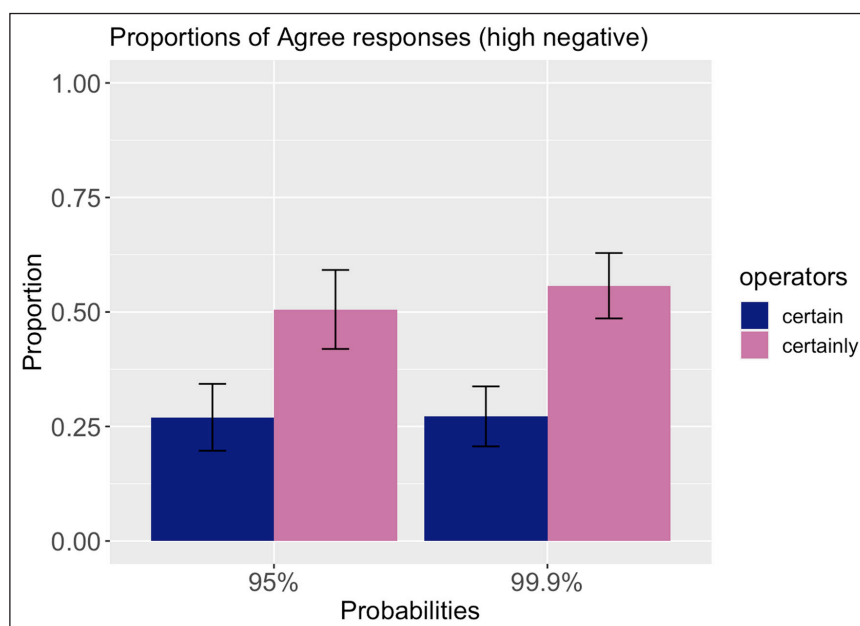
We combine the data from Experiment 1 (95% positive condition) for some comparisons.

Let us start with the effect of extremeness of the probability, (60a) and (60b). The comparisons are visualized in **Figures 9** and **10**. We fitted generalized linear models with the operator type (*certain* or *certainly*) and extremeness of the probabilities (99.9% or 95%) as the fixed effects and their interaction, for each polarity separately.

For the condition with positive sentences (**Figure 9**), both operator type (z-value: 2.105,  $p < .05$ ) and extremeness of the probability (z-value: 3.816,  $p < .01$ ) are significant effects, and



**Figure 9: High probability positive.**



**Figure 10: High probability negative.**

there is an interaction between the operator type and probabilities (z-value:  $-3.695$ ,  $p < .01$ ), supporting the first experimental hypothesis in (59a).

With the negative sentences, on the other hand, the results from the two levels of probabilities do not seem to differ, as visualized in **Figure 10**. A model using the generalized linear models

with operator type (*certain* or *certainly*) and extremeness of the probability (99.9% or 95%) as the fixed effects with the data from four conditions with negative test sentences reveals that operator type is a significant predictor (z-value: 4.547,  $p < .01$ ), as expected, but has no effect of extremeness of probability (z-value:  $-0.034$ ,  $p = .973$ ). There was also no interaction (z-value:  $-0.486$ ,  $p = .627$ ).

To see the effect of polarity at different probability levels ((60c) and (60d)), we used the generalized linear models with the operator type (*certain* or *certainly*) and polarity (*positive* or *negative*) as the fixed effects for different levels of probability, separately. For 99.9% conditions, there is an effect of operator type (z-value: 4.547,  $p < .01$ ) and polarity (z-value: 5.593,  $p < .01$ ), but no interaction (z-value:  $-1.468$ ,  $p = .142$ ). For the 95% conditions, we obtained an effect of operator type (z-value: 3.304,  $p < .01$ ) and polarity (z-value: 8.057,  $p < .01$ ), and an interaction between operator type and polarity (z-value:  $-4.455$ ,  $p < .01$ ).

Let us now combine the whole dataset, consisting of four conditions. First, we fitted a model with operator type (*certain* or *certainly*) and polarity (*positive* or *negative*) and their interaction as the fixed effect for the whole data set (60e). We found an effect of operator type (z-value: 5.625,  $p < .01$ ) and polarity (z-value: 9.950,  $p < .01$ ), as well as interaction (z-value:  $-4.251$ ,  $p < .01$ ).

Lastly, we check whether polarity and the extremeness of the probability interact. We fitted a model with the polarity (*positive* or *negative*) and the extremeness of the probability (99.9% or 95%) as the fixed predictors, and see if there is an interaction ((60g)). We found a significant effect of polarity (z-value: 6.934,  $p < .01$ ), but not of the probability (z-value:  $-0.763$ ,  $p = 0.44570$ ), and there was no interaction (z-value: 1.842,  $p = 0.06549$ ). This model shows that the effect of polarity is observed at all comparisons, whereas the effect of the extremeness of probability is not.

The following overall picture emerged:

- (61) a. The polarity of the test sentence had an effect on the proportion of agree responses. This was observed both at 95% and 99.9% probabilities.
- b. Except for the results from the 95% probability with the adjectival operator *certain* without negation, the extremeness of the probability did not seem to have an effect on the proportion of agree responses, compared to those from the 99.9% probability conditions.<sup>17</sup>

The effect in (61a) is obtained because, overall, we have obtained lower *agree* responses when participants read negative sentences, whether it was at 99.9% or 95%. The effect in (61b) is due to the unexpectedly high proportion of *agree* responses in the 95% positive condition with the experimental sentence with the adjectival operator, *certain*.

---

<sup>17</sup> This observation was additionally supported by fitting the generalized linear models for operator type separately, and see whether the extremeness (moderate or extreme) was a significant predictor. We found that the extremeness of the probability affects participants' responses when the sentence contains *certain* (z-value: 3.816,  $p < .01$ ), but not when the sentence contains *certainly* (z-value:  $-1.093$ ,  $p = .275$ ).

Relating back to the main experimental hypothesis, namely, that the *certainly* is more assertable than *certain*, we found the effect of operator type in every experimental condition, except for the 95% positive condition.

#### 4.4 Possible vs. possibly

Let us next examine the low probability conditions. Recall that we did not find the expected effect of operator type at the 10% measurable probability in Experiment 1, whereas at the 0.1% probability, Lassiter's data show that the operator type is a significant predictor. Therefore, we created two additional conditions for the lower range probabilities (10% negative and 0.1% positive) to test the following hypotheses, repeated from (51):

- (62) a. When the proposition does not contain negation, we obtain fewer agree answers for *possibly* than *possible* in 0.1% condition, with this effect not present in 10% condition (Experiment 1).  
 b. When the proposition contains negation, we obtain fewer agree answers for *possibly* than *possible* in 10%, with this effect not present in 10% positive condition (Experiment 1).

In order to test these hypotheses, we check the following two comparisons.

- (63) a. positive: 0.1% vs. 10%  
 b. 10%: positive vs. negative

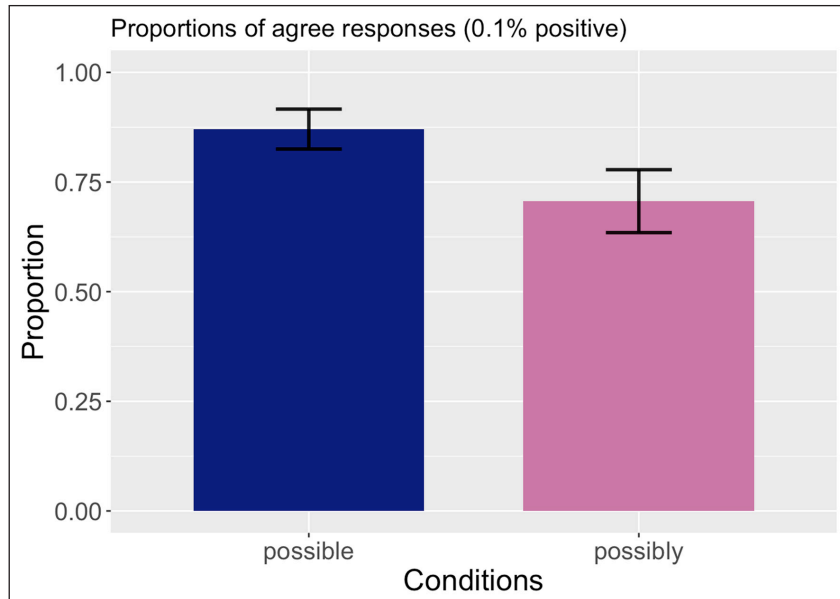
##### 4.4.1 Results

As can be seen in **Figures 11** and **12**, both manipulations resulted in a larger difference in the proportion of agree responses between *possible* and *possibly*.

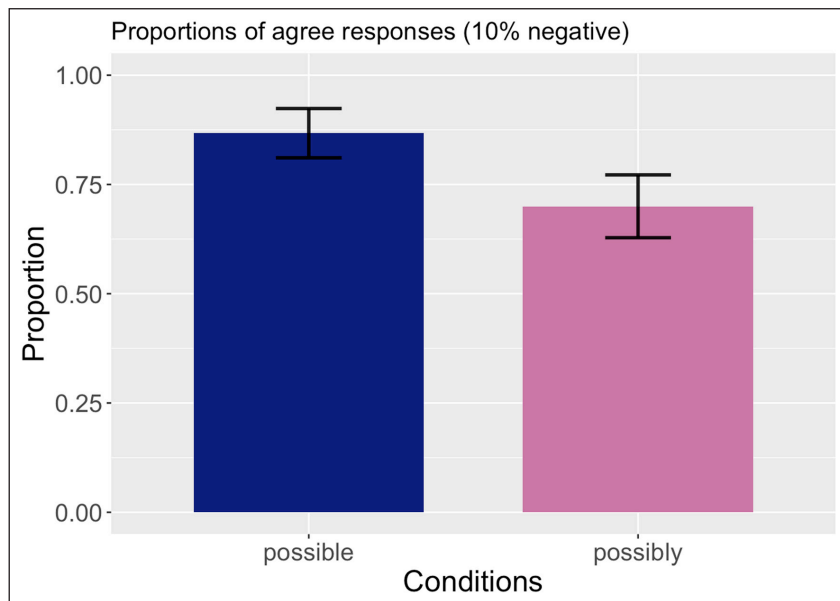
The first condition with the probability being 0.1% is a replication of Lassiter's study. We observe that our results closely mirror those from Lassiter's experiment: In his study, the proportion of *accept* responses for *possible* was 92% (with confidence interval 85% to 98%) and for *possibly* it was 74% (with confidence interval 60% to 84%), which is comparable to the numbers in **Table 15**. A generalized linear model with the operator type (*possible* or *possibly*) as the fixed predictor found a significant effect of the operator type (z-value:  $-3.182$ ,  $p < .01$ ).

Next, consider the condition where we kept the probability at 10%, as in Experiment 1, but changed the polarity of the test sentence to negative. Recall that with the operators *certain* and *certainly*, we observed an overall lower proportion of agree responses when the proposition contained negation. As can be seen in **Figures 11** and **12**, however, this did not hold true for the operators *possible* and *possibly*.

The generalized linear model with the operator type as the fixed predictor shows that operator type (*possible* or *possibly*) is a significant factor (z-value:  $-2.836$ ,  $p < .01$ ).



**Figure 11:** 0.1% probability positive.



**Figure 12:** 10% probability negative.

The summary of predicted and observed differences in asseertability between *possible* and *possibly* is shown in **Table 16**.

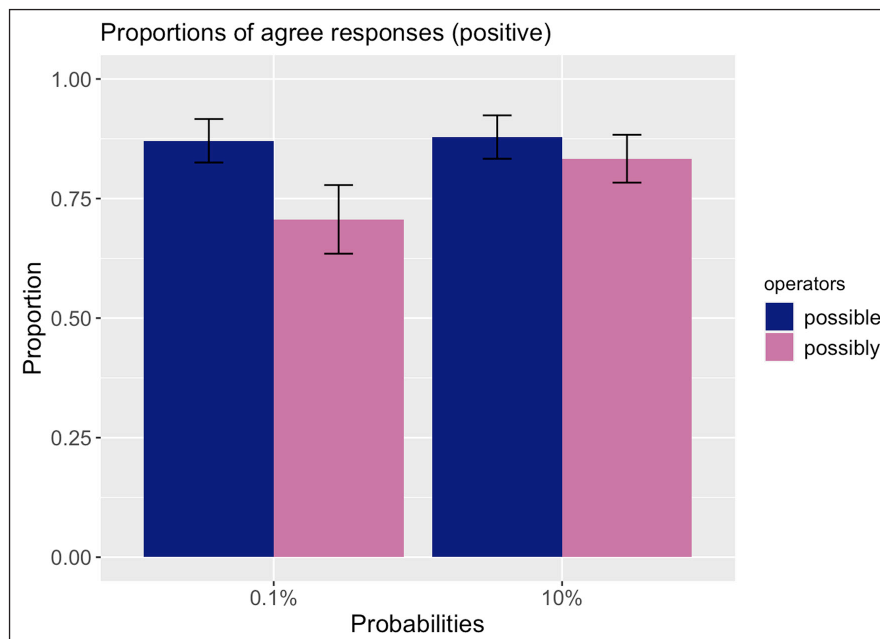
In order to check the effect of extremeness of the probability (**Figure 13**), we compare the results from Experiment 1 (10% positive condition) and 0.1% positive condition from Experiment 2. The statistical model with the operator type (*possible* or *possibly*) and extremeness of the probability

| operator                  | <i>possible</i> 0.1% | <i>possibly</i> 0.1% | <i>possible</i> ~ 10% | <i>possibly</i> ~ 10% |
|---------------------------|----------------------|----------------------|-----------------------|-----------------------|
| N of participants         | 147                  | 109                  | 98                    | 110                   |
| N of agree                | 128                  | 77                   | 85                    | 77                    |
| % of agree                | 87.1                 | 70.6                 | 86.7                  | 70.0                  |
| upper confidence interval | 91.6                 | 77.8                 | 92.4                  | 77.2                  |
| lower confidence interval | 82.5                 | 63.5                 | 81.1                  | 62.8                  |

**Table 15:** *possible* vs. *possibly*.

| Predicted: <i>possible</i> >> <i>possibly</i> |          |                                    |
|---|----------|------------------------------------|
| probability                                   | polarity | direction                          |
| 10%   | negative | <i>possible</i> >> <i>possibly</i> |
| 0.1%  | positive | <i>possible</i> >> <i>possibly</i> |

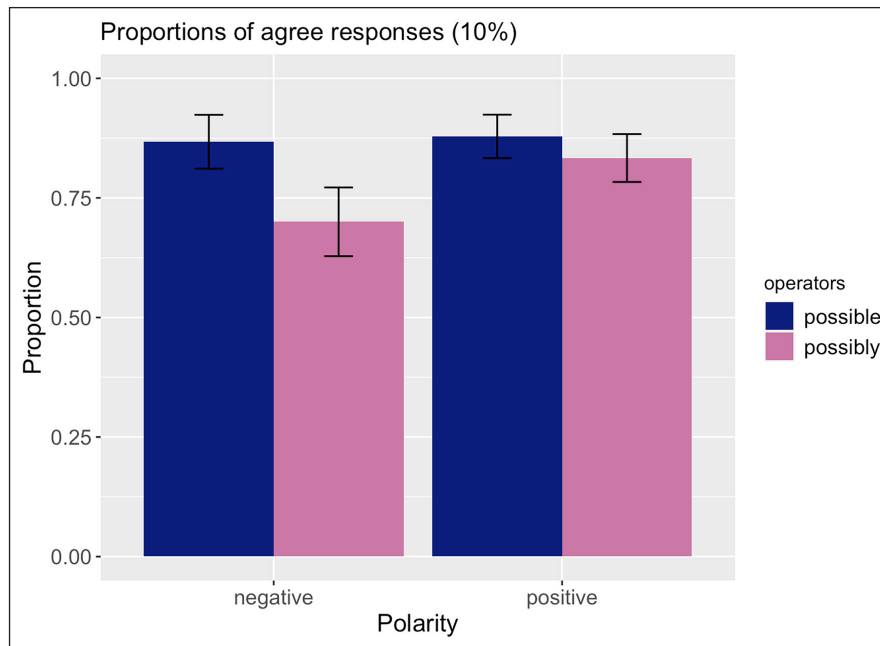
**Table 16:** Summary of predicted & observed assertabilities (*possible* vs. *possibly*).



**Figure 13:** 0.1% vs. 10%.

(0.1% or 10%) as the fixed effects reveal the effect of operator type (z-value:  $-3.182$   $p < .01$ ), but not the extremeness of the probability (z-value:  $0.200$ ,  $p = 0.84149$ ) and no interaction between operator type and extremeness of the probability (z-value:  $1.408$ ,  $p = .15905$ ).

Finally, we check the effect of polarity, as visualized in **Figure 14**. We compare, again, the results from the 10% positive condition of Experiment 1 and 10% negative condition of



**Figure 14:** negative vs. positive (10%).

Experiment 2. The generalized linear models with the operator type (*possible* or *possibly*) and polarity (*positive* or *negative*) as the fixed effect show that there is an effect of operator type (z-value:  $-2.836$ ,  $p < .01$ ) but not of polarity (z-value:  $0.257$ ,  $p = .79741$ ), and no interaction between the operator type and polarity (z-value:  $1.330$ ,  $p = .18354$ ).

#### 4.5 Summary of Experiments

Summarizing the findings from Experiment 2, we have observed the pattern we expected to obtain in all 5 conditions, including the conditions that mirror Lassiter's experimental conditions.

Below, we summarize the comparison between the results from Experiment 1 (measurable probability conditions) and 2.

- (64) a. Overall, the negative sentences elicited lower proportions of agree responses, compared to the positive counterpart, for higher probability conditions that contrasted *certain* and *certainly*.
- b. Except for the results from the 95% probability with the adjectival operator *certain* without negation, the extremeness of the probability did not seem to have an effect on the proportion of agree responses, compared to those from the 99.9% probability conditions, for the high probability conditions.<sup>18</sup>

<sup>18</sup> Note, however, that we did not test 0.1% with *possible* and *possibly*, and therefore, we could not construct a similar model as the data from *certain* and *certainly* conditions. This was because Lassiter's original experimental sentence for *possible* and *possibly* did not include negation, and hence, there was no reason to expect an interaction between polarity and probability.



- c. The changes we have made for Experiment 2, the use of negation and the use of more extreme probability, had enhanced the difference between *possible* and *possibly*.<sup>19</sup>

If we put aside the puzzling data from 95% probability with the adjective operator without a negation, our data support our research hypotheses in (34a) and (35a), repeated here in (65).

- (65) a. The adverb *possibly* leads to a lower level of agree responses than the adjective *possible* in the low probability conditions.  
 b. The adverb *certainly* leads to a higher level of agree responses than the adjective *certain* in the high probability conditions.

The overall lowering effect of negation on the assertability when the operators are *certain* and *certainly*, but not with *possible* and *possibly* remains a puzzle. This is not predicted by our hypotheses, and we leave this for future research.

## 5 General discussion

In this paper we investigated the interpretation of two epistemic adverbials, *certainly* and *possibly*, and their cognate adjectives, *certain* and *possible*, in predicative position, *it is certain/possible that*. A review of the literature revealed evidence for the hypothesis that the adverbials express subjective epistemic modality, whereas adjectives express objective epistemic modality. Also, there is evidence that only epistemic adjectives belong to the at-issue part of a sentence, whereas epistemic adverbials are not-at-issue. From the subjective/objective distinction we derived the following predictions concerning the use of these epistemic operators:

- (66) a. In situations of low probability, *possibly*  $\phi$  is less assertable than *possible*  $\phi$   
 b. In situations of high probability, *certain*  $\phi$  is less assertable than *certainly*  $\phi$
- (67) In situations involving a non-measurable estimation of likelihood, the difference between epistemic adverbs and epistemic adjectives should be greater compared to situations involving a measurable estimation of likelihood.

The motivation for (66a) is that in the case of *possibly*  $\phi$  the speaker, in addition to considering  $\phi$  merely possible, must have an interest to put  $\phi$  forward as a proposition that should be considered for the current conversation. For (66b), we argued that the assignment of a high subjective probability expressed by *certainly*  $\phi$  is easier to defend than the assignment of an objectively high probability. The motivation for (67) is that objective epistemic operators *possible*  $\phi$  and *certain*

---

<sup>19</sup> An anonymous reviewer suggested that the measurable probability condition in the Experiment 1 with low probability at 10% may be due to that the threshold for both *possible* and *possibly* may be below 10%, and as a result, we did not observe a difference in assertability. While this is tenable, given that we obtained a significant effect of the operator type with 10% if the test sentences contained negation, however, this cannot be the sole cause of the result in Experiment 1. We leave this issue for future study.

$\phi$  come with higher verification standards that are more easily met in situations in which the likelihood of a proposition can be stated explicitly.

The hypotheses in (66) was already supported by Lassiter (2016), a study that investigated eight epistemic expressions including the four items investigated here. We replicated this experiment as a part of our Experiment 2, which uses a quantitatively specified scenario involving a lottery with a 99.9% chance vs. a 0.1% chance of winning. Hence, hypotheses (66) has gained further support from our experiment.

Our first experiment diverged from that of Lassiter's in two respects. First, in order to be able to work with just one scenario, Lassiter tested the high probability epistemics with negated sentences, *certainly not  $\phi$*  and *certain not  $\phi$* . We suspected a possible influence of negation and therefore constructed a second scenario in which high probability epistemics could be tested with a non-negated sentence.

Second, in order to be able to test (67), which involves a comparison between a scenario with an objectively given "measurable" probability and one which only allows for a "non-measurable", subjective estimation, we constructed a scenario involving a 95% vs. a 10% chance for the measurable condition. This was based on the norming study we conducted with the set-up story and two lists of evidence for the non-measurable contexts that we used for the main experiment.

With these changed scenarios, our experiment (part of Experiment 1) failed to support (66a): Participants gave *agree* responses more often for *certain  $\phi$*  than for *certainly  $\phi$* . While this is surprising, especially because Lassiter's study had supported (66a) with a negative sentence in conditions with more extreme probabilities, this provides us with a new insight into what might affect the assertability of these operators. Below, we provide an analysis toward explaining our findings.

As discussed in section 4, a comparison between the 99.9% scenario (Experiment 2) and the 95% scenario revealed that the change from 95% to 99.9% did not change the proportion of *agree* responses for *certainly*, but it had a significant effect for *certain*.

This initially very surprising result can be explained as follows. It is known that rounder numerals lead to a more approximate interpretation than numerals that are less round (Krifka 2009; Solt 2014). For this reason, the 99.9% scenario evokes a high degree of precision (it involves nine hundred ninety-nine out of one thousand lottery tickets), whereas the 95% scenario evokes a lesser degree of precision (nine hundred fifty out of one thousand lottery tickets). We think that under the high precision standard evoked in the 99.9% case, the participants were *less* willing to apply the objective epistemic *certain*, compared to the lower precision standard evoked in the 95% case. The influence of context is similar as in (68), where the second clause is likely interpreted in a more approximate way in (68a), and in a precise way in (68b) (see also Aparicio Terasa 2017; Beltrama & Schwarz 2021 for experimental evidence of contextual effects on the level precision at which a numerical expression is interpreted).

- (68) a. Yesterday we had 950 people in the Zoom conference, and today there were 1000 people.  
 b. Yesterday there were 999 people in the Zoom conference, and today there were 1000 people.

A possible test whether it is the granularity level that is responsible for the interpretative differences of the adjective *certain*, and not the difference in probabilities, would be to test the assertability of *certain* at a lower probability level that is given with higher precision standards, e.g., in a scenario in which 951 tickets out of 1000 tickets were bought by a specified person. However, it is imaginable that such proportions are translated to a coarser granularity level, something that is not possible for 999 tickets out of 1000 tickets.

Interestingly, the raising of the precision standard by the scenario did not affect the subjective epistemic *certainly*. We see this as confirming our understanding of the nature of epistemic adjectives vs. epistemic adverbs: Epistemic adjectives make objective claims about the estimation of likelihood, which depend on the level of granularity or precision that is salient in the context. With epistemic adverbs, on the other hand, a speaker expresses his or her own estimation of likelihood, which is less dependent on contextual expectations depending on the precision level. We think that this difference can be related to a difference in the structure of the scales that such adjectives and adverbs tap into. Only epistemic adjectives make use of a numerical scale of probabilities on which the notion of granularity is meaningful, while their adverbial counterparts are interpreted with respect to a coarser or tolerant scale structure, perhaps parallel to the approximate representation of numerosity (cf. Feigenson et al. 2004). A similar distinction is invoked by Solt (2016) in accounting for a range of distributional and interpretative differences between the quantifiers *most* and *more than half*. In this respect, the epistemic adverbs *possibly*, *probably* and *certainly* seem to correspond to *most*, while the epistemic adjectives *possible*, *probable* and *certain* correspond to *more than half*. This is supported by the fact that only epistemic adjectives can be modified numerically, as in *it is ninety-five percent certain that it will rain*, not *\*it will ninety-five percent certainly rain*.<sup>20</sup>

Recall that Lassiter (2016) tested the reactions to high probabilities with negated sentences, *certainly not  $\phi$*  and *certain not  $\phi$* . We also investigated the use of negated sentences in the 99.9 percent condition and in the 95 percent condition, and found no difference between these conditions. In both conditions, (66a) was supported. This means that the explanation for the surprising judgements in the non-negated cases that involved the precision level evoked by the

---

<sup>20</sup> An anonymous reviewer raises the issue whether the argument made with the contrast between *it is ninety-five percent certain that it will rain* and *\*it will ninety-five percent certainly rain* has any force, because the second sentence “is probably just ungrammatical.” Our interpretation of this criticism is this: The second sentence may sound odd for purely *syntactic* reasons, hence its unacceptability should not be used as evidence of *interpretive* differences between adverbs and adjectives. We find this criticism legitimate, and would address it as follows. First, note that both *certain* and *certainly* can in principle be modified. Thus, replacing *nine-five percent* with, say, *very* in the two sentences above would result in both being acceptable. What we observe, then, is that substitution of *very* by *ninety-five percent* results in unacceptability when the modified expression is an adverb, but not when it is an adjective. This observation becomes less mysterious in light of the hypothesis that adjectives makes use of a scale that is more precise than adverbs, hence constitutes supporting evidence for the hypothesis.

scenario does not apply in the negated case. This may be attributable to the influence of negation itself; currently we do not see an obvious reason for such an influence of negation.<sup>21</sup>

Alternatively, it may be an artifact of the scenario. Notice that the non-negated case invoked a rich person that bought 999 or 950 tickets from a raffles with 1000 tickets. This is unusual and may have alerted the participants to apply higher precision standard, which affected their use of epistemic operators. In the negated case, the scenario invoked a person that bought one or fifty tickets from a raffles with 1000 tickets, which is a much more usual scenario that might have evoked more stereotypical reasoning patterns. One way to investigate this issue further would be to design an experiment in which very high probabilities, like 99.9% or 95%, are not unusual.

As for the use of *possibly* and *possible*, we found that these expressions showed no significant difference in a scenario with a 10% probability. Hence (66b) is supported in a 0.1% probability scenario by Lassiter's original experiment and by our replication, but the experimental evidence is largely indifferent in the case of 10% probability. Following the motivation for our hypothesis, the use of *possibly*  $\phi$  implies that the speaker does not only admit that  $\phi$  is possible, but also wants to introduce this proposition as an option into the discussion. We think that this latter motivation is more plausible for the participants in case of a 10% possibility than in case of a 0.1% possibility. Consequently, the difference between *possible* and *possibly* vanishes with greater underlying probabilities. It would be interesting to see at which probability level the difference between *probably* and *probable* becomes insignificant.

Our second goal was to investigate the use of epistemic expressions in situations in which the likelihood of a proposition cannot be objectively measured. We proposed hypotheses (67), which predict that the differences between epistemic adverbs and adjectives should then be greater. We met with severe experimental obstacles here because it was difficult to construct scenarios with non-measurable probability that were judged to be as likely as the lottery scenarios with measurable probability. As mentioned above, we had to adjust the scenario with measurable probability from 99.9%/0.1% to 95%/10% to ensure comparability with a suitable scenario of non-measurable probability. Even after settling on less extreme probabilities we could not achieve optimal experimental conditions, as the participants varied quite widely in their probability estimation. We suspect that it will be difficult to achieve more uniform estimations of likelihood with scenarios specified by texts, as texts are understood as narratives, and the point of narratives often is to tell surprising events. Hence, even if one does one's best to construct a story which makes it extremely likely or unlikely that a given proposition is true, some readers will be suspicious that this is done in order to depict the eventual outcome as even more surprising, and adjust their estimation of likelihood for the proposition accordingly. This might have been more

---

<sup>21</sup> A reviewer pointed out to us that negation has been observed to interact with politeness (Yoon et al. 2020 and Tessler & Franke 2019), for example. While this may be the case, it is not clear to us why negation should affect the truth-value judgments, however, from semantics of negation. We leave this issue for future research.

prominent with the murder scenario that we used, which evokes the genre of detective novels that live on surprising developments, but we think that it could also affect other types of text.

The phenomenon that it is difficult to express extreme probabilities in language without resorting to specialized numerical terminology has been noticed before. As mentioned in 1.2.2, there is a substantial body of psychological work on the question how verbal descriptions of probabilities translate into numerical ones. Dhimi & Mandel (2020) review this literature and discuss the benefits and drawbacks of verbal probability descriptions in intelligence reports. They note that interpretations of verbally stated probabilities show high variability, and that very small and very high probabilities cannot be adequately expressed.

The predictions of hypothesis (67) in the condition of non-measurable probabilities were met in the low probability condition in Experiment 1, cf. **Figure 2**. While there was no significant difference between *possible*  $\phi$  and *possibly*  $\phi$  in the measurable case, in the non-measurable case participants agreed to *possibly*  $\phi$  significantly less often than to *possible*  $\phi$ , in accordance to (66a): The difference between the items was greater than in the measurable case, and in the predicted direction. However, the level of this interaction did not reach significance.

In addition, assertability was significantly lower in the non-measurable condition than in the measurable condition, an observation that we did not predict. This can be explained as follows: In the measurable condition, participants were given an objective reason to assume that the proposition is possible (as the named person bought lottery tickets, and the text did not give any indication that these tickets were prevented from winning). In the non-measurable condition, the scenario gave five strong reasons that made it unlikely that the named person committed the crime. This difference has likely triggered the lower acceptance rate of both *possible*  $\phi$  and *possibly*  $\phi$  under this condition.

The predictions of hypothesis (67) were not met in the the high probability condition (cf. **Figure 2**). Agreement to the items *certain*  $\phi$  and *certainly*  $\phi$  did not differ in the non-measurable case (recall that in the measurable case, they showed a significant difference but the direction opposite from what was predicted by (66b)). One plausible post-hoc explanation for this finding is that in the non-measurable condition there simply is no plausible epistemic judge that would be different from the speaker, hence the interpretation of objective *certain* and subjective *certainly* are identical. This reasoning would also apply to *possible* and *possibly*, of course. But in this case, there is the additional assumed difference that *possibly* has the function of introducing the prejacent proposition into the conversation as an option to be considered, which explains why these epistemic operators show difference in the direction predicted by (66a).

## 6 Conclusion

In this paper we investigated epistemic adverbs like *certainly* and *possibly*, contrasting them with epistemic adjectives like *certain* and *possible*, in sentences like *It is certainly raining* and *It is certain*

*that it is raining*. We have reviewed the existing evidence that the former are subjective epistemic operators, expressing a subjective estimation of likelihood, whereas the latter are objective epistemic operators, reporting on a probability estimation by a relevant epistemic source.

We were particularly interested in the consequences that this difference of epistemic sources have when asserting sentences in scenarios that suggest certain probabilities. Following Lassiter (2016), we hypothesized that assertions of subjective epistemic propositions of the form *certainly*  $\phi$  are easier to defend than assertions of objective epistemic propositions of the form *certain*  $\phi$ , and hence should be more easily assertable in the high frequency condition. We could replicate this finding in our experiment. We also could replicate Lassiter's finding that assertions of the form *possibly*  $\phi$  are *less* easily assertable than assertions of the form *possible*  $\phi$  in the low frequency condition. But we found a problem with Lassiter's motivation for this finding. We proposed a different explanation: With subjective epistemic assertions, the speaker intends to communicate the prejacent proposition  $\phi$  as an option to be considered. This makes the assertion of the form *possibly*  $\phi$  less assertable than sentences of the form *possible*  $\phi$ .

In our experiment, we used similar scenarios as Lassiter, in which the probabilities were objectively specified by given frequencies in a lottery. We intended to test the difference between subjective and objective epistemic operators in scenarios that do not involve measurable probabilities. Our hypothesis was that subjective operators are more suitable, and better assertable, in such scenarios. They are important, as such uses are presumably more frequent when epistemic operators are used in everyday conversation. However, we could not find the expected result: That subjective epistemic sentences are more assertable than objective epistemic sentences in the high likelihood condition. We also faced experimental hurdles in defining scenarios that reliably evoke specific likelihoods.

Furthermore, we discovered a surprising interpretation of the objective operator *certain* for extremely high vs. high probabilities, which we suggest is due to the granularity level at which the probability is specified, in the sense that the standards for *certain* are lower in scenarios of coarser granularity. We also detected an influence of negation which was surprising; we discussed the possibility that it is due to a raising of the granularity level in the previous section. We hope that the factors we identified as relevant for the interpretation of subjective and objective epistemic operators are useful to guide further experimental research.

In particular, we would like to see research on the role of granularity levels in epistemic operators and on their use in cases of non-measurable probabilities. In addition, it would be interesting to see whether, and how, the interactions uncovered here also affect other epistemic expressions such as *must*, *might*, *likely*, among others.

---



## Appendix A: Norming Study

Because it is desirable for the perceived probabilities of the mathematically non-measurable probability scenarios to closely track the mathematically measurable probabilities (1% and 99%), we have conducted a pre-study for the present investigation to find out what the probability that a participant is likely to assign for each scenario would be. The design of the pre-study is as follows: A participant (i) reads the set-up of the story, (ii) reads the evidence for the proposition, and (iii) is asked the probability of the proposition being true, choosing a number between 0% and 100%. Crucially, the sentences used in (i) and (ii) will be used for the present study, which are described below. 104 participants took part in the pre-study through the Mechanical Turk platform. 40 participants read the high probability scenario, and 32 participants read the low probability scenario. 32 participants read a scenario that targeted a middle range probability, for future research. Some participants that provided a number above 100 and/or have responded too quickly (less than 10 seconds) were excluded, leaving 37 participants for the high probability scenario and 30 participants for the low probability scenario. Mean percentage, median, and standard deviation are shown in **Table 17**.

|                  | mean  | median | SD   |
|------------------|-------|--------|------|
| High probability | 90.1% | 95     | 13.0 |
| Low probability  | 14.7% | 10     | 17.2 |

**Table 17:** Norming study results.

The range of numbers that the participants provided was from 50 to 100 for the high probability scenario. 32 out of 37 participants provided a number more than or equal to 90. As for the low probability scenario, the range of numbers that were provided varied from 0 to 100. 19 out of 30 participants provided a number less than or equal to 10. Based on these results, we decided to set the mathematically measurable condition at 10% and 95%. These probabilities were used to compare the effect of measurability on the acceptability of each sentence. Our prediction was that there would be an interaction between the measurability and acceptability.

## Appendix B: Comprehension questions

*Please indicate whether the following is true or false:*

- Measurable probability:
  - i. A political party held a raffle to raise money for their candidate. (false)
  - ii. Most tickets were purchased by the mayor. (false)
  - iii. A total of 1000 tickets were sold. (true)

- Non-measurable probability:
  - i. The murder took place at an airport. (false)
  - ii. The victim was shot. (false)
  - iii. The murder weapon was found. (true)

## Data availability

The scientific data associated with the experiments presented in this paper and R-code are available at osf (DOI: [10.17605/OSF.IO/U3ZE2](https://doi.org/10.17605/OSF.IO/U3ZE2)).

## Ethics and consent

The experiments reported in this paper have been approved by Ethics committee at the German linguistic society (Deutsche Gesellschaft für Sprachwissenschaft) in January, 2020 (#2018-05-200124).

All participants of this study were informed that there is no physical/emotional risk associated with the study, and that they could stop at any point during the experiment, and consented to participate in the study.

## Acknowledgements

For very helpful questions and comments, we would like to thank the participants at Sinn und Bedeutung 26 at University of Cologne, and three anonymous reviewers for this journal. This research was supported by the ERC Advanced Grant 73929. Data and scripts for analyses are available at OSF (DOI: [10.17605/OSF.IO/U3ZE2](https://doi.org/10.17605/OSF.IO/U3ZE2)).

## Funding information

Research for this article was funded by the European Research Council under Horizon 2020, ERC Advanced Grant 787929 (SPAGAD: Speech Acts in Grammar and Discourse).

## Competing interests

The authors have no competing interests to declare.

---

## References

- Aparicio Terasa, Helena. 2017. *Processing context-sensitive expressions: The case of gradable adjectives and numerals*. University of Chicago dissertation.
- Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, Articles* 67(1). 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>



- Beaver, David & Roberts, Craig & Simons, Mandy & Tonhauser, Judith. 2017. Questions Under Discussion: Where information structure meets projective content. *Annual Review of Linguistics* 3. 265–284. DOI: <https://doi.org/10.1146/annurev-linguistics-011516-033952>
- Bellert, Irena. 1977. On semantic and distributional properties of sentential adverbs. *Linguistic Inquiry* 8(2). 337–351.
- Beltrama, Andrea & Schwarz, Florian. 2021. Imprecision and speaker identity: How social meaning affects pragmatic reasoning. In *Semantics and linguistic theory (SALT)* 31. DOI: <https://doi.org/10.3765/salt.v31i0.5107>
- Beyth-Marom, Ruth. 1982. How probable is probable? A numerical translation of verbal probability expressions. *Journal of Forecasting* 1. DOI: <https://doi.org/10.1002/for.3980010305>
- Clark, Dominic. 1990. Verbal uncertainty expressions: A critical review of two decades of research. *Current Psychology* 9. 203–235. DOI: <https://doi.org/10.1007/BF02686861>
- Deschamps, Isabelle & Agmon, Galit & Loewenstein, Yonatan & Grodzinsky, Yosef. 2015. The processing of polar quantifiers, and numerosity perception. *Cognition* 143. 115–128. DOI: <https://doi.org/10.1016/j.cognition.2015.06.006>
- Dhami, Mandeep K. & Mandel, David R. 2020. Words or numbers? communicating probability in intelligence analysis. *American Psychologist*. Advance online publication. DOI: <https://doi.org/10.1037/amp0000637>
- Ernst, Thomas. 2009. Speaker-oriented adverbs. *Natural Language and Linguistic Theory* 27(3). 497–544. DOI: <https://doi.org/10.1007/s11049-009-9069-1>
- Feigenson, Lisa & Dehaene, Stanislas & Spelke, Elizabeth. 2004. Core systems of number. *Trends in cognitive sciences* 8(7). 307–314. DOI: <https://doi.org/10.1016/j.tics.2004.05.002>
- Haliday, M. A. K. & Matthiessen, Christian M. I. M. 2004. *Introduction to Functional Grammar*. London: Hodder Arnold 3rd edn. DOI: <https://doi.org/10.4324/9780203783771>
- Hengeveld, Kees. 1988. Illocution, mood and modality in a functional grammar of Spanish. *Journal of semantics* 6(1). 227–269. DOI: <https://doi.org/10.1093/jos/6.1.227>
- Herbstritt, Michelle. 2020. *Investigating the Language of Uncertainty. Experimental Data, Formal Semantics and Probabilistic Pragmatics*. University of Tübingen dissertation.
- Just, Marcel Adam & Carpenter, Patricia Ann. 1971. Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior* 10. 244–253. DOI: [https://doi.org/10.1016/S0022-5371\(71\)80051-8](https://doi.org/10.1016/S0022-5371(71)80051-8)
- Karttunen, Lauri. 1972. Possible and must. In Kimball, John (ed.), *Syntax and semantics volume 1*, 1–20. Seminar Press. DOI: [https://doi.org/10.1163/9789004372986\\_002](https://doi.org/10.1163/9789004372986_002)
- Krifka, Manfred. 2009. Approximate interpretations of number words: A case for strategic communication. In Hinrichs, Erhard W. & Nerbonne, John (eds.), *Theory and Evidence in Semantics*, 109–132. Stanford: CSLI Publications. DOI: <https://doi.org/10.18452/9508>
- Krifka, Manfred. to appear (a). Layers of assertive clauses: Propositions, judgments, commitments, acts. In Hartmann, Julia & Wöllstein, Angelika (eds.), *Propositionale argumente im sprachvergleich: Theorie und empirie* (Studien zur Deutschen Sprache), Gunter Narr Verlag.

- Krifka, Manfred. to appear (b). Zur negierbarkeit von epistemischen modalen. In Neuhaus, Laura (ed.), *Grammatik und pragmatik der negation im deutschen*. Walter de Gruyter.
- Lassiter, Daniel. 2016. *Must*, knowledge, and (in)directness. *Natural Language Semantics* 24. 117–163. DOI: <https://doi.org/10.1007/s11050-016-9121-8>
- Lewis, David. 1980. A subjectivist's guide to objective chance. *Studies in inductive logic and probability* 2. 263–293. DOI: [https://doi.org/10.1007/978-94-009-9117-0\\_14](https://doi.org/10.1007/978-94-009-9117-0_14)
- Lyons, John. 1977. *Semantics – Volume 2*. Cambridge University Press.
- Müller, Kalle. 2019. *Satsadverbien, evidentialität und non-at-issueness*. Universität Tübingen dissertation.
- Maché, Jakob. 2019. *How Epistemic Modifiers Emerge*. Berlin: Walter de Gruyter. DOI: <https://doi.org/10.1515/9783110411027>
- Moss, Sarah. 2015. On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics* 8. 1–81. DOI: <https://doi.org/10.3765/sp.8.5>
- Nilsen, Øystein. 2004. Domains for adverbs. *Lingua* 114. 809–847. DOI: [https://doi.org/10.1016/S0024-3841\(03\)00052-4](https://doi.org/10.1016/S0024-3841(03)00052-4)
- Nuyts, Jan. 1992. Subjective vs. objective modality: What is the difference. In Fortescue, Michael & Harder, Peter & Kristoffersen, Lars (eds.), *Layered structure and reference in a functional perspective: Papers from the functional grammar conference, copenhagen, 1990*, 73–97. Benjamins Amsterdam.
- Nuyts, Jan. 1993. Epistemic modal adverbs and adjectives and the layered representation of conceptual and linguistic structure. *Linguistics* 31. 933–969. DOI: <https://doi.org/10.1515/ling.1993.31.5.933>
- Nuyts, Jan. 2001. Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics* 33. 383–400. DOI: [https://doi.org/10.1016/S0378-2166\(00\)00009-6](https://doi.org/10.1016/S0378-2166(00)00009-6)
- Nuyts, Jan. 2016. Analyses of the modal meanings. In *The Oxford Handbook of Mood and Modality*. Oxford University Press. DOI: <https://doi.org/10.1093/OXFORDHB/9780199591435.013.1>
- Öhlschläger, Günter. 1989. *Zur Syntax und Semantik der Modalverben*. Tübingen: Niemayer. DOI: <https://doi.org/10.1163/25890859-00104002>
- Papafragou, Anna. 2006. Epistemic modality and truth conditions. *Lingua* 116. 1688–1702. DOI: <https://doi.org/10.1016/j.lingua.2005.05.009>
- Piñon, Christopher. 2009. The thing about modal adverbs. *Snippets* 10.
- Ricciardi, Giuseppe & Ryskin, Rachel A. & Gibson, Edward. 2020. Epistemic ‘must’ is an inferential evidential. *Experiments in Linguistic Meaning* 1.
- Roberts, Craige. 2017. Agreeing and assessing. Oslo Workshop in Non-At-Issue-Meaning and Information Structure. Ms, The Ohio State University.
- Roberts, Craige. 2019. The character of epistemic modality: Evidential indexicals. Ms, The Ohio State University.
- Solt, Stephanie. 2014. An alternative theory of imprecision. In Snider, Todd & D’Antonio, Sarah & Weigand, Mia (eds.), *Proceedings of the 24th semantics and linguistic theory conference (SALT24)*, 514–533. DOI: <https://doi.org/10.3765/salt.v24i0.2446>

- Solt, Stephanie. 2016. On measurement and quantification: the case of *most* and *more than half*. *Language* 92(1). 65–100.
- Tessler, Michael Henry & Franke, Michael. 2019. Not unreasonable: Why two negatives don't make a positive, 1–13. <https://psyarxiv.com/tqjr2>.
- Verstraete, Jean-Christophe. 2001. Subjective and objective modality: Interpersonal and ideational functions in the english modal auxiliary system. *Journal of pragmatics* 33(10). 1505–1528. DOI: [https://doi.org/10.1016/S0378-2166\(01\)00029-7](https://doi.org/10.1016/S0378-2166(01)00029-7)
- von Fintel, Kai & Gillies, Anthony. 2007. An opinionated guide to epistemic modality. *Oxford Studies in Epistemology* 2. 32–62.
- Wallsten, Thomas S. & Budescu, David V. & Rapoport, Amnon & Zwick, Rami & Forsyth, Barbara. 1986. Measuring the vague meanings of probability terms. *Journal of Experimental Psychology: General* 115. 348–365. DOI: <https://doi.org/10.1037/0096-3445.115.4.348>
- Wason, Peter Carthcart. 1959. The processing of positive and negative information. *Quarterly Journal of Experimental Psychology* 11. 92–107. DOI: <https://doi.org/10.1080/17470215908416296>
- Wason, Peter Carthcart. 1961. Response to affirmative and negative binary statements. *British Journal of Psychology* 52. 133–142. DOI: <https://doi.org/10.1111/j.2044-8295.1961.tb00775.x>
- Wolf, Lavi & Cohen, Ariel & Simchon, Almog. 2016. An experimental investigation of epistemic modal adverbs and adjectives. In *Proceedings of sinn und bedeutung*, vol. 20. 798–814.
- Wolf, Lavi. 2015. *Degree of Assertion*: Ben-Gurion University of the Negev dissertation.
- Yoon, Erica J. & Tessler, Michael Henry & Goodman, Noah D. & Frank, Michael C. 2020. Polite speech emerges from competing social goals. *Open Mind* 4. 71–87. DOI: [https://doi.org/10.1162/opmi\\_a\\_00035](https://doi.org/10.1162/opmi_a_00035)

