



Göbel, Alexander & Wagner, Michael. 2023. How the position of *at least* affects its interpretation: experimental data. *Glossa: a journal of general linguistics* 8(1), pp. 1–40. DOI: <https://doi.org/10.16995/glossa.9273>



Open Library of Humanities

## How the position of *at least* affects its interpretation: experimental data

Alexander Göbel, Princeton University, US, [alandergoebel@gmail.com](mailto:alandergoebel@gmail.com)

Michael Wagner, McGill University, Canada, [chael@mcgill.ca](mailto:chael@mcgill.ca)

---

The Focus-particle *at least* is known to be ambiguous between two interpretations: an *epistemic* one conveying uncertainty/ignorance, and a *concessive/evaluative* one that conveys a desirability ranking. Prior literature has argued that the position of *at least* determines what interpretation is available: *at least* is concessive/evaluative ad-sententially and epistemic ad-nominally. We present three experiments that investigate how three properties which the two interpretations have been taken to differ on are restricted by the syntactic position of *at least*, namely entailment of the prejacent, truth of higher alternatives, and desirability. The results overall support the view that the interpretation of *at least* is restricted by its position as previously claimed by some accounts, but syntactic position seems to dissociate the three properties in ways incompatible with previous assumptions. We discuss the implications of these results for formal accounts of *at least*.

---

*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by the Open Library of Humanities. © 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

OPEN ACCESS



## 1 The Ambiguity of *at least*

The Focus-particle *at least* has been studied primarily in the context of modified numerals (Krifka 1999; Geurts & Nouwen 2007; Büring 2008; Schwarz 2016; Alexandropoulou 2018; Mendia 2022). Its characteristic contribution is what has been called an ignorance inference: in (1a), *at least* conveys that the speaker does not know exactly how much Grover ate, essentially suspending the implicature possible in the absence of *at least* that Grover did not eat more than some of his dinner. However, *at least* allows an additional interpretation that is not concerned with knowledge but contributes an evaluation: in (1b), Grover eating some of his dinner is presented as a positive outcome compared to a conceivable worse one, albeit a less than ideal one compared to a conceivable better one. Additionally, the implicature that Grover ate some but not all of his dinner is not necessarily suspended. Prior literature on this ambiguity has labeled the two interpretations ‘scalar’ and ‘evaluative’ (Kay 1992), or ‘epistemic’ and ‘concessive’ (Nakanishi & Rullmann 2009; Biezma 2013; Chen 2018). Here we adopt the labels ‘epistemic’ and ‘evaluative’ for the distinction, borrowing from both dichotomies according to what we take to be descriptively most informative.<sup>1</sup>

- (1) a. Grover ate **at least** [SOME]<sub>F</sub> of his dinner (... he might’ve even eaten all of it).  
 b. **At least** Grover ate [SOME]<sub>F</sub> of his dinner (... he could’ve eaten nothing at all).

A factor that has been identified to disambiguate between the two interpretations is the syntactic position of *at least*, as used in (1). However, accounts differ on some of the specifics: Nakanishi & Rullmann (2009) propose — relative to a configuration like (1) when the object is modified — that the evaluative interpretation is strongly preferred over the epistemic one when *at least* occurs sentence-initially/ad-sententially, while only the epistemic interpretation is available in ad-nominal position (see also Chen 2018). In contrast, Biezma (2013) argues that ad-nominal *at least* allows an evaluative interpretation when supported by the context (see also Cohen & Krifka 2014 for a less restrictive view). The primary goal of this paper is to investigate the claims concerning the mapping between *at least*’s syntactic position and its interpretation experimentally.

A second point where accounts diverge concerns how this apparent ambiguity is formally implemented. According to Nakanishi & Rullmann (2009), *at least* constitutes a case of lexical ambiguity, with distinct and unrelated lexical entries for two interpretations. Tentative support for this view comes from the observation that many languages use separate lexical items for the two interpretations (e.g. German, Grosz 2011). In contrast, Biezma (2013) and Chen (2018) propose a unified analysis of *at least* that leverages pragmatics to disambiguate toward the

---

<sup>1</sup> On the one hand, ‘scalar’ seems inadequate given that both interpretations may be construed as relating to a scale; on the other hand, ‘concessive’ has a negative connotation that in our view misrepresents the discourse function of the respective interpretation (see Section 5). Moreover, to concede something suggests accepting it as true, such that the label ‘concessive’ would presuppose one of the properties the paper aims to investigate.

intended interpretation by using the context. While a unified approach is generally preferable for economy reasons, it has to maintain empirical coverage.

To address this issue, the current paper looks at three properties where epistemic and evaluative *at least* have been argued to differ and discusses the compatibility of the results with each approach. The three properties we look at are (i) whether a higher alternative has to be left open, (ii) whether the truth of the prejacent is entailed, and (iii) whether alternatives are evaluated in terms of desirability. To preview the main findings, we find strong support for the syntactic position of *at least* affecting its interpretation along the lines of Nakanishi & Rullmann (2009), but fail to provide evidence that the prejacent is entailed in any position. The results raise new issues for all existing accounts, since these three properties are predicted to pattern together as bundles.

The paper is structured as follows: Section 2 introduces relevant background from previous proposals. Section 3 presents the three experiments investigating the syntax-semantics mapping of *at least* and Section 4 provides a discussion of the results. Section 5 concludes the paper and adds some more speculative remarks.

## 2 Background

This section first adds more detail to the above mentioned claims about the relation between the possible interpretation of *at least* and its syntactic position (2.1), then presents the most relevant differences of the two interpretations regarding their meaning (2.2), and concludes with a discussion of how these differences are implemented in formal accounts (2.3).

### 2.1 Prior claims on the syntax-semantics mapping of *at least*

As mentioned in the introduction, it has been argued that the syntactic position of *at least* determines what interpretations are available for it, although the exact nature of this mapping has been debated. Nakanishi & Rullmann (2009) assume the symmetrical mapping illustrated in (2), with the symmetry only being weakened by treating ad-sentential position as being strongly biased toward an evaluative interpretation rather than excluding an epistemic one completely.<sup>2</sup> Chen (2018) adopts the same view but does fully exclude epistemic *at least* ad-sententially.

- (2) Syntax-semantics mapping from Nakanishi & Rullmann (2009)
- |                         |  |                                   |
|-------------------------|--|-----------------------------------|
| a. <i>Ad-sentential</i> | <b>At least</b> Grover ate [TUNA] <sub>F</sub> . | strongly biased toward EVALUATIVE |
| b. <i>Ad-verbal</i>     | Grover <b>at least</b> ate [TUNA] <sub>F</sub> . | EPISTEMIC or EVALUATIVE           |

---

<sup>2</sup> Nakanishi & Rullmann furthermore assume that sentence-final position allows both interpretations but we will put this position aside here given it is less central to the disagreement.

c. *Ad-nominal*Grover ate **at least** [TUNA]<sub>F</sub>.

only EPISTEMIC

Contrasting with this view, other researchers have assumed an asymmetrical mapping or directly argued against Nakanishi & Rullmann (2009). Kay (1992) notes – although using different terminology – that for the epistemic interpretation *at least* has to be sister to the Focused constituent and cannot associate at a distance, whereas the evaluative interpretation is available anywhere and has “essentially no role in the structure of the sentence”. This latter point has also been assumed by Cohen & Krifka (2014). Furthermore, Biezma (2013) explicitly argues that evaluative *at least* is possible in ad-nominal position, providing the example in (3) as support for this view.

- (3) The track and field coaches are looking at the statistics and discussing the results of the last competition.

*Coach 1:* The competition was awful.

*Coach 2:* Yes, but Mary won **at least** that gold medal [pointing at the data in the statistics] (Biezma 2013: (12))

However, Chen (2018) raises doubt about the reported judgment regarding (3) and mentions consulted native speakers considering it infelicitous. The main goal of the experiments presented in Section 3 is to resolve this tension and clarify what the facts are with respect to the syntax-semantics mapping of *at least*. An overview of the central accounts and respective claims is shown in **Table 1**, with the main points of divergence highlighted.<sup>3</sup> For ease of reference, Nakanishi & Rullmann (2009) and Chen (2018) will be labeled the *symmetric view* and Biezma (2013) the *asymmetric view*.

	Nakanishi & Rullmann/Chen	Biezma
<i>ad-sentential</i>	EPI (✗), EVAL ✓	EPI ✗, EVAL ✓
<i>ad-verbal</i>	EPI ✓, EVAL ✓	EPI ✓, EVAL ✓
<i>ad-nominal</i>	EPI ✓, EVAL ✗	EPI ✓, EVAL ✓

**Table 1:** Overview of claims about syntax-semantics mapping of *at least*.

The next subsection discusses the semantic and pragmatic properties that the interpretations of *at least* have been associated with.

<sup>3</sup> Although Nakanishi & Rullmann and Chen differ in whether epistemic *at least* competes with a strong bias for evaluative *at least* ad-sententially or is ruled out completely – hence the parentheses – they are put together here insofar as it is difficult to distinguish between these options experimentally. For instance, the cognitive effort to override a strong bias may be reflected in lower acceptability, much like choosing an interpretation that is consistent with the bias but incompatible with the context. While there are ways to address this issue, we focus on categorically opposed claims here and leave further investigations for future research.

## 2.2 Semantic and pragmatic properties of epistemic and evaluative *at least*

There are a number of meaning aspects that the two interpretations of *at least* share and differ on beyond the informal characterizations as conveying uncertainty or desirability, respectively (see Nakanishi & Rullmann 2009 and Chen 2018). One similarity concerns the inability to associate with elements at the bottom or top of a scale, as shown in (4).

- (4) a. #Grover **at least** ate [NOTHING/EVERYTHING]<sub>F</sub>, although I don't know exactly how much.  
 b. #Grover **at least** ate [NOTHING/EVERYTHING]<sub>F</sub>, but it could've been worse.

Regarding differences between the two interpretations, there are three properties most relevant here, which will be leveraged in the experiments presented later. The first property concerns the truth of higher alternatives. For epistemic *at least*, a higher alternative must be an open possibility as part of its essential uncertainty contribution. If all higher alternatives are ruled out, either as true or false, the epistemic interpretation becomes unacceptable, as shown in (5a) (see also Nouwen's (2015) anti-specificity requirement). In contrast, evaluative *at least* is compatible with the truth or falsity of higher alternatives being known, as shown by the acceptability of (5b).<sup>4</sup>

- (5) a. [I'm not exactly sure how Emma did in the race. I know she didn't win a gold medal...]  
 #but she won **at least** [SILVER]<sub>F</sub>.  
 b. [It's too bad Emma didn't win a gold medal...]  
 but **at least** she won [SILVER]<sub>F</sub>.

The second difference related to uncertainty concerns the entailment status of the prejacent. By virtue of the requirement to leave a higher alternative open, epistemic *at least* is compatible with the prejacent being false depending on the type of alternatives involved. For alternatives that are mutually compatible with each other, like *some* relative to *all* in (1), the prejacent will be true by virtue of being entailed by all higher alternatives. In contrast, for alternatives that are mutually exclusive, as in (5), the truth-value of the prejacent remains in principle open: If it turned out that Emma won gold, it would entail the falsity of *Emma won silver*. Evaluative *at least*, on the other hand, has been argued to always entail the truth of its prejacent.<sup>5</sup>

Finally, the third difference between the two interpretations is the type of ranking involved with respect to the alternatives featured in their computation. What might be considered the

---

<sup>4</sup> In fact, Biezma (2013) argues that the evaluative interpretation requires higher alternatives to be false to be available, which is adopted by Chen (2018). We will come back to this issue below.

<sup>5</sup> Note that Nakanishi & Rullmann (2009) argue for this property by relying on what inferences seem to be available intuitively rather than constructing examples where the inference would affect acceptability. Experiments 1 and 3 will directly address this issue.

essential component of evaluative *at least* is that alternatives are ranked according to desirability. In contrast, the epistemic interpretation is neutral in this respect. As illustration, consider the (slightly modified) example set from Kay (1992) in (6), where only (6d) leads to oddness due to the higher alternatives, i.e. *many/all people were injured*, being widely considered less desirable.<sup>6</sup>

- (6) a. In that big train wreck **at least** [several people]<sub>F</sub> were saved.  
 b. In that big train wreck **at least** [several people]<sub>F</sub> were injured.  
 c. **At least** in that big train wreck [several people]<sub>F</sub> were saved.  
 d. **#At least** in that big train wreck [several people]<sub>F</sub> were injured.

(Kay 1992: (22))

A related effect of evaluative *at least* is what has been labeled “settling for less” by Nakanishi & Rullmann (2009), illustrated in (7). The informal intuition with respect to (7b) here is that the outcome of winning eight gold medals is already such a good outcome that evoking a desirability scale seems asking for too much: while winning nine gold medals might have been a better outcome and winning seven a worse one in some sense, they would be considered fully satisfactory on most accounts.<sup>7</sup> Note, however, that this judgment – and the interpretation of *at least* more generally – depends on context, both in terms of what counts as desirable and the alternatives for the ranking. For instance, if Phelps had the unusual goal to win eight gold medals in addition to at least one silver medal, but failed to win any silver medals, (7b) would be acceptable.

- (7) a. Phelps won **at least** [eight gold medals]<sub>F</sub>.  
 b. **#At least** Phelps won [eight gold medals]<sub>F</sub>.

(Nakanishi & Rullmann 2009: (9))

A summary of the relevant properties in relation to the interpretation of *at least* are shown in **Table 2**. Experiment 1 investigates (i) and (ii), Experiment 2 tests (iii), and Experiment 3 features (ii) and (iii).

The next subsection briefly goes over the formal analyses that have been proposed for the two interpretations of *at least* and how they account for the three properties discussed, as well as how they relate to the question of their syntactic distribution.

---

<sup>6</sup> This judgment minimally assumes that the addressee does not root for people getting injured. In a conversation between malicious people, (6d) may be perfectly acceptable. More generally, we assume that it is by default the speaker that is committed to the desirability ranking but that it has to be shared by the addressee to not result in infelicity.

<sup>7</sup> Although Nakanishi & Rullmann (2009) bring up this example and argue that their analysis accounts for it, it is not completely clear whether the explanation works when further scrutinized. We will come back to this issue below.

	<b>EPI</b>	<b>EVAL</b>	
(i) truth of higher alternative	open	agnostic	Exp1
(ii) truth of prejacent	can be entailed	entailed	Exp1&3
(iii) ranking	neutral	desirability	Exp2&3

**Table 2:** Overview of proposed properties of epistemic and evaluative *at least*.

### 2.3 Formal Accounts of epistemic and evaluative *at least*

As mentioned in Section 1, there is a large body of work on epistemic *at least* but comparatively little on its ambiguity and the evaluative interpretation. To keep the background comprehensive, we thus focus on formal accounts that discuss both interpretations of *at least*, beginning with Nakanishi & Rullmann (2009). Their account treats *at least* as a case of lexical ambiguity with distinct lexical entries for each interpretation. Both interpretations are separated into an asserted and a conventionally implicated component without an obvious common core, as shown in (8).

- (8) a. Epistemic *at least*  
**Assertion:**  $\exists q \in C[q \geq p \wedge q(w) = 1]$   
‘there is a proposition  $q$  which ranks higher than or as high as the target proposition  $p$ , and which is true’  
**Conventional Implicature:**  $\exists w'[\text{Epist}(w, w') \wedge \exists q \in C[q > p \wedge q(w') = 1]]$   
‘it is epistemically possible that some proposition  $q$  that ranks higher than  $p$  is true’
- b. Evaluative (/concessive) *at least*  
**Assertion:**  $p(w) = 1$   
‘the target proposition  $p$  [=the prejacent] is true’  
**Conventional Implicatures:**  
(i)  $\forall r, r' \in C[r > r' \leftrightarrow r' \text{ is preferred to } r]$   
‘The scalar ranking reflects a preference ranking’  
(ii)  $\exists q \in C[q > p]$   
‘There is a proposition  $q$  that ranks higher than  $p$ ’  
(iii)  $\exists q \in C[q < p]$   
‘There is a proposition  $q$  that ranks lower than  $p$ ’

The entries capture the three properties discussed in the previous subsection straightforwardly. First, the requirement of epistemic *at least* for higher alternatives to be left open is contributed by its conventional implicature. Second, the prejacent of evaluative *at least* being entailed is directly asserted, and third, its desirability (or preference) ranking is conventionally implicated.

Contrasting with this lexical ambiguity view, Biezma (2013) and Chen (2018) argue for a unified account of *at least* with a single entry that encompasses both readings, which will be discussed in turn. Biezma’s entry is shown in (9). First, the alternatives *at least* interacts

with are related to the Question under Discussion (Roberts 2012) to allow orderings that are determined contextually rather than arising from a lexical scale. Second, it is presupposed that there exists an alternative  $\gamma$  that is ranked below the prejacent  $\alpha$  and an alternative  $\beta$  that is ranked above, with either  $\alpha$  or  $\beta$  being true (see Büring 2008 for the idea of using disjunction for epistemic *at least*). Finally, all alternatives  $\mu$  ranked below the prejacent that are not entailed by it are false.

- (9) Let  $\alpha$  be a proposition, and  $[\alpha]_{A,i}$  the set of alternatives of  $\alpha$  ordered according to  $\leq_i$ , where  $\leq_i$  is a contextually salient order of alternatives and  $\forall \gamma \in [\alpha], \gamma \in \text{QuD}$ .
- $$\llbracket \text{at least } \alpha \rrbracket = \lambda w. \exists \beta, \gamma \in [\alpha]_{A,i} \text{ s.t. } \gamma <_i \alpha <_i \beta \ \& \\ [\alpha(w) \vee \beta(w)] \ \& \\ \forall \mu \in [\alpha]_{A,i}, \mu <_i \alpha, [\neg \mu(w) \vee \alpha \text{ entails } \mu]$$

This analysis can account for the relevant three properties as follows. Starting with the third property, the type of ranking is taken to directly follow from the QuD inferred from the context. That is, alternatives are ordered according to desirability for evaluative *at least* by virtue of taking into account relevant goals of the speaker in the conversation. An additional assumption necessary to derive the two other properties is that an evaluative interpretation can only arise when higher alternatives are known to be false. In this case, the disjunction of the prejacent and the higher alternative will entail the truth of the prejacent since the prejacent is the only remaining possibility, capturing the second property. The requirement of epistemic *at least* for higher alternatives to be left open additionally follows from this assumption, although without being derived through the lexical entry.

Moving on to Chen (2018), his analysis in (10) adopts various aspects from Biezma (2013) but in a different format. First, *at least* again relates to a set of alternatives, represented as  $C$  here, that has to include a true alternative, labeled  $\gamma$ . However, rather than directly using disjunction to limit the true alternative to something as high as the prejacent or above and specifying all non-entailed lower alternatives as false, Chen restricts the alternative set to only those alternatives  $\beta$  that are higher than and not entailed by the prejacent  $\alpha$ . This implementation is illustrated in (11), where the sentence in (11a) gives rise to the alternative set in (11b) with the alternative below the prejacent being excluded. Secondly, the analysis again relies on contextual factors to determine the relevant ranking, this time through a measure function  $\mu_c$ . The use of a measure function allows the analysis to more directly do justice to the *-est* superlative morphology argued to be present in *at least* and similar expressions across languages, which will be put aside here however. Finally, Chen follows Biezma in assuming that an evaluative interpretation only arises when higher alternatives are known to be false.

- (10)  $\llbracket \text{at least } (C) \rrbracket^{w,c} = \lambda \alpha_{<s,t>}. \exists \gamma [\gamma \in C \wedge \gamma_w \wedge \forall \beta [\beta \in C \wedge \beta \neq \alpha \rightarrow \mu_c(\alpha) < \mu_c(\beta)]]$



- (11) a. Emma **at least** won [SILVER]<sub>F</sub>.
- b.  $C = \left\{ \begin{array}{l} \text{Emma won gold} \\ \text{Emma won silver} \\ \text{Emma won bronze} \end{array} \right\}$

Turning to the explanation for the three properties, the analysis captures the requirement of the epistemic interpretation for higher alternative to be open as a case of semantic vacuity: in a case where the truth of higher alternatives is known, epistemic *at least* would no longer contribute anything relative to a sentence without *at least* and appear redundant, analogous to the case of *only* in (12). Secondly, the assumed prejacent entailment of evaluative *at least* is explained similarly to Biezma's account, where – by virtue of the evaluative interpretation only arising when higher alternatives are false – the prejacent is the only remaining alternative in *C* that could be true. Lastly, the difference in ranking is derived through the measure function choosing the appropriate type contextually.

- (12) #Of Lois and Hal, only Lois and Hal came.

A question left open by these accounts, however, is how the two interpretations relate to the syntactic position of *at least*. While Nakanishi & Rullmann (2009) do not provide any commentary on this issue, Biezma (2013) suggests an explanation rooted in processing, since the patterns are mere biases on her view. She proposes that epistemic *at least* is more prevalent in ad-nominal position because the proximity to a scalar item facilitates semantic computation and makes an epistemic interpretation quickly available. In contrast, evaluative *at least* requires access to contextual information that is less readily available such that the additional distance between the particle and its associate provides extra time for the comprehension process to be completed. Although this idea seems interesting, it also raises further questions. For instance, it would seem necessary to stipulate and specify some competition mechanism insofar as it is unclear why the epistemic interpretation should not be available from a distance if it is actually easier to process. Moreover, it is unclear how to treat cases like (13) where *at least* associates with the subject. On the processing account described by Biezma, we might expect only the epistemic interpretation to be available here or at least strongly preferred, despite the intuition that an evaluative interpretation seems quite readily available.

- (13) **At least** [Grover]<sub>F</sub> ate tuna.

Chen (2018) takes a different perspective on the distributional issue, with distinct approaches for each interpretation. For evaluative *at least*, he proposes the hypothesis in (14), which relates the preference for ad-sentential position to a need to scope over a semantic object of sufficient size. However, while intuitively appealing, many treatments of other Focus-particles like *only* and *even* treat them as propositional operators that move at LF (but see Erlewine 2014 for relevant

differences between *only* and *even*). The question then is why evaluative *at least* does not have this movement option available.

- (14) *The Quantificational Domain Hypothesis* (Chen 2018: (104))  
 The quantificational domain of concessive *at least* must be (minimally) propositional (i.e., a set of propositional alternatives).

For the epistemic interpretation, Chen hints at the possibility of evaluative *at least* associating with functional heads above CP that are not accessible to epistemic *at least*. Tentative evidence for this view comes from the observation that fragment answers only allow an epistemic interpretation but not an evaluative one (15).

- (15) *Context*: There are three individuals in the discourse: Adam, Bill and Chris.  
 Emily: Who did John invite?  
 Frank: **At least** [Adam and Bill]<sub>F</sub> ✓EPI, #EVAL  
 cf. **At least** John invited [Adam and Bill]<sub>F</sub> #EPI, ✓EVAL

However, what complicates an explanation for any syntactic restriction for an epistemic interpretation is that English provides the alternative expression *at the very least*, which seems to convey the same meaning as epistemic *at least* but has been taken to allow for association at a distance, illustrated in (16) using the difference in the type of ranking as diagnostic. If the two expressions do in fact share the same meaning, it would become more difficult to rule out the epistemic interpretation ad-sententially by appealing to an interaction of its meaning with parts of the syntactic structure.

- (16) **At the very least** the accident injured [some people]<sub>F</sub>.  
 (cf. #**At least** the accident injured [some people]<sub>F</sub>.)

To sum up, while the analyses in Biezma (2013) and Chen (2018) provide interesting starting points for an explanation for the syntactic mapping of interpretations of *at least*, various issues remain unsolved. While providing a more satisfactory explanation goes beyond the scope of this paper, we will provide further discussion in light of the experimental results in Section 4.

### 3 Experiments

The main goal of the following three experiments is to test which of the three meaning components discussed in the previous subsection that the two interpretations have been argued to differ on – the truth of higher alternatives, the truth of the prejacent, and the type of ranking – are part of the interpretation of *at least* in different syntactic positions. That is, rather than assuming epistemic and evaluative *at least* as a monolithic distinction, we take a neutral stance that tests what individual properties are available in what position. Experiment 1 investigates the truth

of higher alternatives and the truth of the prejacent, Experiment 2 the type of ranking, and Experiment 3 the truth of the prejacent and the type of ranking.

### 3.1 Experiment 1: Truth of Higher Alternatives & Truth of Prejacent

#### 3.1.1 Design & Materials

Stimuli consisted of short dialogues comprised of a context sentence and a target sentence containing *at least* with the intention to render the properties of interest compatible with one dialogue and incompatible with the other:

(17) Sample Item Experiment 1

- a. *Evaluative* CONTEXT: Higher alternatives false  
 A: It's too bad Yvette didn't win a gold medal at the school olympics.  
 B: True, but (**at least**) she (**at least**) won (**at least**) silver.
- b. *Epistemic* CONTEXT: Alternatives open + speaker uncertainty  
 A: Do you know whether Yvette won a gold medal at the school olympics?  
 B: Not sure, but (**at least**) she (**at least**) won (**at least**) silver.

First, in the *evaluative* CONTEXT illustrated in (17a), the higher alternative *win gold* is negated in A's statement and hence no longer open. As a consequence, the use of *at least* in a given position should become redundant if it receives an interpretation that requires a higher alternative to be open, resulting in infelicity. Second, in the *epistemic* CONTEXT given in (17b), B initially conveys uncertainty regarding A's question, leaving alternatives open. However, if *at least* is interpreted to entail the truth of the prejacent, it should result in a conflict with B's asserted uncertainty, since the prejacent being true would constitute a complete answer for mutually exclusive alternatives.<sup>8</sup> Crucially, while the *evaluative* CONTEXT is incompatible with an interpretation that requires higher alternatives to be open, it is compatible with taking the prejacent to be entailed, and vice versa for the *epistemic* CONTEXT.

The second factor SYNTAX was manipulated by placing *at least* in ad-sentential, ad-verbal or ad-nominal position. The idea then is that, given a context, any decrease in acceptability for a particular syntactic position can be attributed to the position being restricted to an interpretation that has the property incompatible with the context.<sup>9</sup> Overall, the experiment was thus a  $2 \times 3$  design.

---

<sup>8</sup> In order to adequately test the assumptions for this manipulation, half of the items contained alternatives that were mutually exclusive, as in (17), and half were mutually compatible (e.g. a target sentence with the predicate *stay until the intermission* would be compatible with the truth of *stay for both acts of the play* given in the context), where any conflict between prejacent entailment and speaker uncertainty should arise for the former type of alternatives but not the latter. However, this manipulation did not significantly affect the overall results such that it will be omitted in the following discussion, but see the OSF repository for additional details.

<sup>9</sup> Notably, the CONTEXTS differ in A's speech act – (17a) being an assertion and (17b) a question – which may independently affect the acceptability of the dialogues. However, given that the main focus is on differences between the positions of *at least* within CONTEXT, this issue is negligible.

We created 24 item sets like (17). To avoid spillover and priming effects, every participant saw only one condition from each item set, for a total of 24 dialogues, and each participant saw an equal number of dialogues from each of the 6 conditions ('Latin Square Design'). For each participant, the 24 dialogues were presented in pseudo-random order, such that a given condition could not occur more than twice in a row. There were an additional 24 filler sentences from a different experiment on *only*, which were interspersed with the test trials. All main items including filler items can be accessed at the OSF repository associated with this paper, see Supplementary Files for the link.

### 3.1.2 Procedure

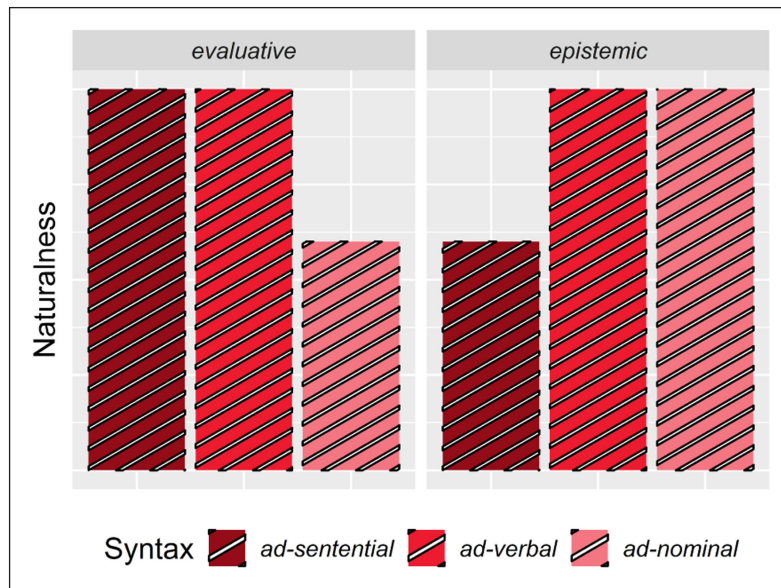
The experiment was implemented through prosodyExperimenter (<https://github.com/prosodylab/prosodylabExperimenter>) and run online on Prolific.ac. Participants first saw a welcome screen, followed by an online consent form, and a language background questionnaire. For the main part of the experiment, participants were presented with a dialogue for them to read, and then had to provide a naturalness rating on a scale from 1 (completely unnatural) to 6 (completely natural) based on how they thought the response sounded given the context. There were three practice trials after receiving instructions, followed by the main experiment with 24 test dialogues and 24 interspersed filler dialogues. The experiment concluded with a chance to provide feedback. The full experiment took 8–10 minutes.

### 3.1.3 Participants

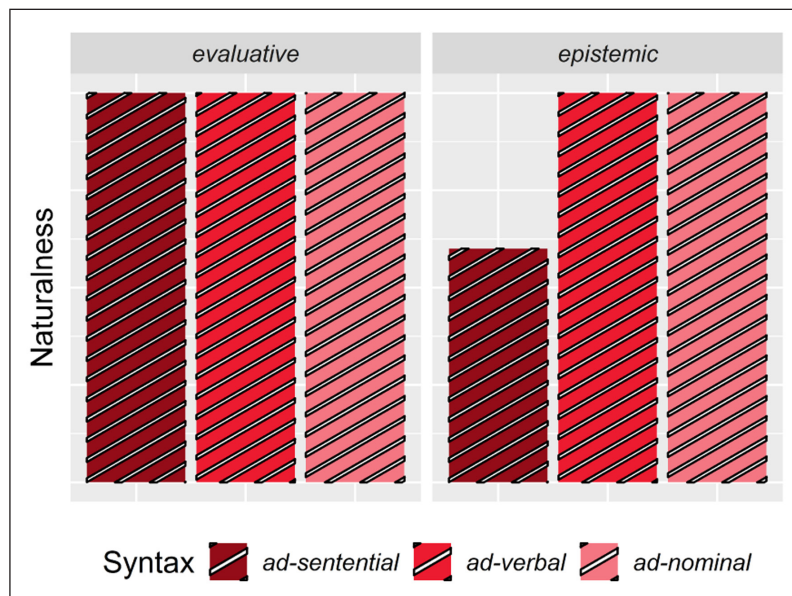
24 participants were recruited from Prolific.ac and paid US\$1.60 each. Two participants were excluded based on rating the 'good' practice item as good or worse than the 'bad' practice item, leaving 22 participants for data analysis.

### 3.1.4 Predictions & Coding

As laid out in Section 2.1, the symmetric view and the asymmetric view only differ in their assumptions about the syntactically mediated availability of evaluative *at least*: while evaluative *at least* is ruled out in ad-nominal position by the symmetric view, the asymmetric view argues that it is in fact available. Applied to the experiment, the symmetric view hence predicts lower ratings for ad-nominal *at least* relative to other positions in evaluative contexts, while the asymmetric view predicts no such difference. For epistemic *at least*, on the other hand, both views take this interpretation to be ruled out in ad-sentential position, therefore predicting a decrease relative to other positions in epistemic contexts. Crucially, these predictions rely on the assumption that an epistemic interpretation requires higher alternatives to be open and that an evaluative interpretation entails its prejacent. An idealized pattern of the results predicted by each view – presupposing an even baseline for the two contexts – is shown in **Figures 1** and **2**.



**Figure 1:** Idealized pattern predicted by symmetric view, Experiment 1.



**Figure 2:** Idealized pattern predicted by asymmetric view, Experiment 1.

In terms of corresponding statistical effects, both views predict a significant difference between ad-sentential and the other two positions within epistemic contexts, with the ad-sentential position receiving lower ratings. In contrast, only the symmetric view predicts significantly lower ratings for ad-nominal position within evaluative contexts relative to the other two

positions, whereas the asymmetric view predicts no such effect. Since this latter point is where the two views diverge, the factors will be coded such that *evaluative* is the dummy reference level of CONTEXT and *ad-nominal* the Helmert reference level of SYNTAX. The comparison of *ad-sentential + ad-verbal.VS.ad-nominal* would thus be the crucial effect. The relevant effect for the shared prediction in this model would be the interaction of *evaluative.VS.epistemic* and *ad-sentential.VS.ad-verbal*.

### 3.1.5 Results

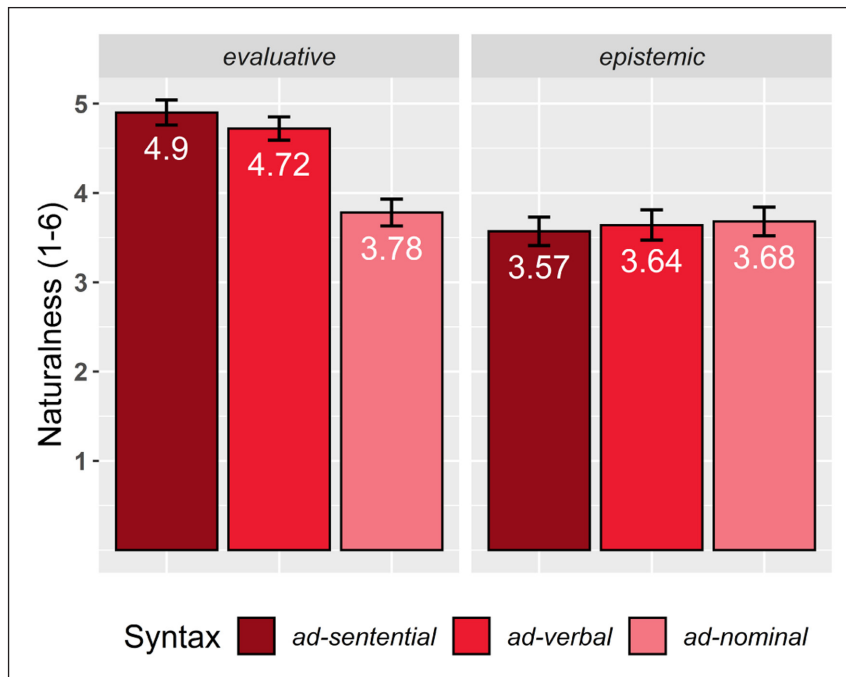
Data were analyzed with an ordinal mixed effects model with random intercepts for items and participants, which was the most maximal model that converged (see Barr et al. 2013), using the `clmm` function from the `ordinal` package (Christensen 2019) in R (R Core Team 2018), the output of which is shown in **Table 3**. The average ratings per condition are shown in **Figure 3**. Looking first at the pattern of results for evaluative contexts, we can see a decrease for ad-nominal position relative to the ad-sentential and ad-verbal ones. This difference is reflected in a significant effect of *ad-sentential + ad-verbal.VS.ad-nominal* ( $z = 6.675$ ,  $p < .001^{***}$ ). The numerical decrease for ad-verbal position relative to ad-sentential position was not significant ( $z = -1.269$ ,  $p = 0.204$ ).

Turning to the results for epistemic contexts, we first note that ratings are overall lower than in evaluative contexts, supported by the model (*evaluative.VS.epistemic*:  $z = -7.652$ ,  $p < .001^{***}$ ). Secondly, the position of *at least* does not seem to affect the ratings noticeably, hence differing from evaluative contexts: the interaction between *ad-sentential.VS.ad-verbal* and *evaluative.VS.epistemic*, which was relevant to the predictions of the symmetric and asymmetric views, did not reach significance ( $z = 1.204$ ,  $p = 0.229$ ). Moreover, there was a significant effect of the interaction between *ad-sentential + ad-verbal.VS.ad-nominal* and *evaluative.VS.epistemic* ( $z = -5.170$ ,  $p < .001^{***}$ ).

	Est.	Std. Err.	z-value	p-value
EVAL VS epi	-0.725	0.095	-7.652	<.001***
AD-NOM VS (ad-sen + ad-ver)	0.939	0.141	6.675	<.001***
ad-sen VS ad-ver	-0.211	0.166	-1.269	.204
EVAL VS epi * AD-NOM VS (ad-sen + ad-ver)	-1.013	0.196	-5.170	<.001***
EVAL VS epi * ad-sen VS ad-ver	0.276	0.229	1.204	.229

**Table 3:** Experiment 1 model output: `clmm(Rating ~ Context*Syntax + (1|Subject) + (1|Item))`. CONTEXT dummy coded with *evaluative* as reference level [eval = 0, epi = 1], SYNTAX Helmert coded with *ad-nominal* as reference level [ad-sen =  $\frac{1}{3}$ ,  $-\frac{1}{2}$ ), ad-ver =  $(\frac{1}{3}, \frac{1}{2})$ , ad-nom =  $(-\frac{2}{3}, 0)$ ]. Small capitals indicate reference level of the relevant factor, if applicable.





**Figure 3:** Mean ratings (with standard errors) by condition, Experiment 1.

### 3.1.6 Discussion

The experiment yielded mixed results regarding the predictions laid out in Section 3.1.4. In evaluative contexts, when higher alternatives were false, ratings showed a contrast between ad-sentential and ad-verbal position on the one hand and ad-nominal position on the other, with the latter showing a decrease. This pattern is in line with the symmetrical view on the availability of the evaluative interpretation of *at least*: *at least* in ad-nominal position only allows the epistemic interpretation, leading to a decrease in acceptability due to a violation of the requirement that higher alternatives be left open. The results are hence at odds with the predictions of the asymmetrical view, which holds that evaluative *at least* is available even in ad-nominal position.<sup>10</sup>

In contrast, for epistemic contexts, we found no effect of *at least*'s syntactic position: ratings for ad-sentential, ad-verbal and ad-nominal position were not significantly different from each other. This lack of effect is surprising in light of the predictions, based on which we would have expected a decrease for ad-sentential *at least* relative to other positions. The ratings for epistemic contexts also were overall lower than for evaluative ones, numerically approximating the lowest

<sup>10</sup> Note, however, that naturalness ratings serve as a proxy for the availability of an interpretation rather than being a direct reflection of it. That is, in the absence of a negative baseline to compare unambiguous instances of each interpretation of *at least* to, a decrease in naturalness may also be an indicator of additional effort to reach a certain interpretation instead of that interpretation being not available at all.

average rating in ad-nominal position. A possible explanation for this difference in baseline may be that the dialogues in evaluative contexts are in some sense complete, whereas the reply in epistemic contexts does not fully resolve A's question. While this pattern is unexpected, it does not invalidate the interpretation of the results with respect to our main research question insofar as it allows us to assess differences between the positions of *at least* reliably. Given that the average still hovers around 4, it would have been easy for participants to give lower ratings, such that we can rule out a floor effect.<sup>11</sup>

An alternative explanation for the lack of an effect could be that the manipulation was simply too weak to register. Specifically, the manipulation relied on a conflict between speaker B conveying uncertainty regarding A's question by uttering *not sure* or a similar expression, and the potential entailment of the prejacent being at odds with this uncertainty. There are two ways in which this manipulation could then be contrasted with the one for evaluative contexts. First, the manipulation in evaluative contexts boils down to rendering epistemic *at least* redundant due to no higher alternatives being left open, which Chen (2018) treats as a case of semantic vacuity. The epistemic context manipulation, in contrast, depends on a conflict within speaker B's belief states, which may be considered more pragmatic in nature. Second, while the evaluative context manipulation simply requires the participant to interpret *at least* relative to the context, for the intended violation in the epistemic context to be noticeable it would be necessary to represent multiple belief states, one of which relies on an inference from *at least's* prejacent. Epistemic contexts may then be subject to instances of shallow processing where participants make a judgment without going through all the necessary steps to notice the violation.

If we take the results at face value however, the question then is which aspect of previous characterizations of *at least* is inaccurate: either evaluative *at least* does not entail its prejacent, or the epistemic interpretation is available in ad-sentential position. The next experiment pursues this second option by examining another property associated with evaluative *at least*, namely the desirability ranking.

## 3.2 Experiment 2: Desirability Ranking

### 3.2.1 Design & Materials

This experiment again used short dialogues like Experiment 1 and a similar design, but used the context to manipulate a different property that has been argued to distinguish the two interpretations: evaluative *at least* ranks its alternatives on a desirability scale, where higher alternatives are marked as more desirable and lower alternatives as less desirable, while epistemic *at least* does not convey this type of evaluation. Through testing in what syntactic

---

<sup>11</sup> Another possible explanation for the lower ratings for epistemic contexts would be that in fact all positions show a decrease due to entailing the prejacent. However, this explanation would amount to arguing that there is no epistemic interpretation of *at least*, which we should see at least in ad-nominal position given the decrease in evaluative contexts.



position a desirability ranking occurs, the present experiment allows us to test to what extent an epistemic interpretation is available even in ad-sentential position, as suggested by the results of Experiment 1. The dialogues were modeled after the epistemic condition in Experiment 1:

(18) Sample Item Experiment 2

a. *Desirable* CONTEXT:

A: You were worried that the teacher might not pass the students.

How many of the students do you think she passed?

B: I'm not sure. (**At least**) she (**at least**) passed (**at least**) some of the students.

b. *Undesirable* CONTEXT:

A: You were worried that the teacher might fail the students.

How many of the students do you think she failed?

B: I'm not sure. (**At least**) she (**at least**) failed (**at least**) some of the students.

The crucial manipulation was whether a desirability ranking would result in a positive or a negative statement by changing the main verb of the target sentence and prior utterances leading up to it: in the *desirable* CONTEXT in (18a), B's reply with *pass* with a desirability ranking would convey that B considers some students passing less desirable than all students passing, which is in line with A's first sentence about B's attitude toward the possible outcomes. In contrast, interpreting the reply in (18b) with a desirability ranking would convey that B finds a lower number of students failing less desirable than a higher number, which violates basic social norms. While this violation may be sufficient to render the sentence less acceptable, A's first statement was added to make the internal contradiction explicit. Additionally manipulating the syntactic position of *at least* as in Experiment 1, we get a  $2 \times 3$  design.

There were three further differences between the dialogues used here and the epistemic contexts in Experiment 1. First, A's question is a *how many*-question rather than a polar question targeting the highest alternative to minimize repetition from the preceding context sentence. Second, B's reply was split into two sentences rather than conjoined with *but* to render the syntax of the target clause and the position of *at least* in it even more clear. Third, the word assumed to receive Focus in the target sentence was always *some*, which is mutually compatible with the higher alternative *all* rather than being mutually exclusive, which was necessary to create a sufficient number of items that allowed antonym-like minimal pairs as in (18).

There were 24 item sets like (18), Latin-squared as before, in addition to 24 filler items from an unrelated experiment on *even*. Both item sets are again accessible through the OSF repository linked in Supplementary Files.

### 3.2.2 Procedure

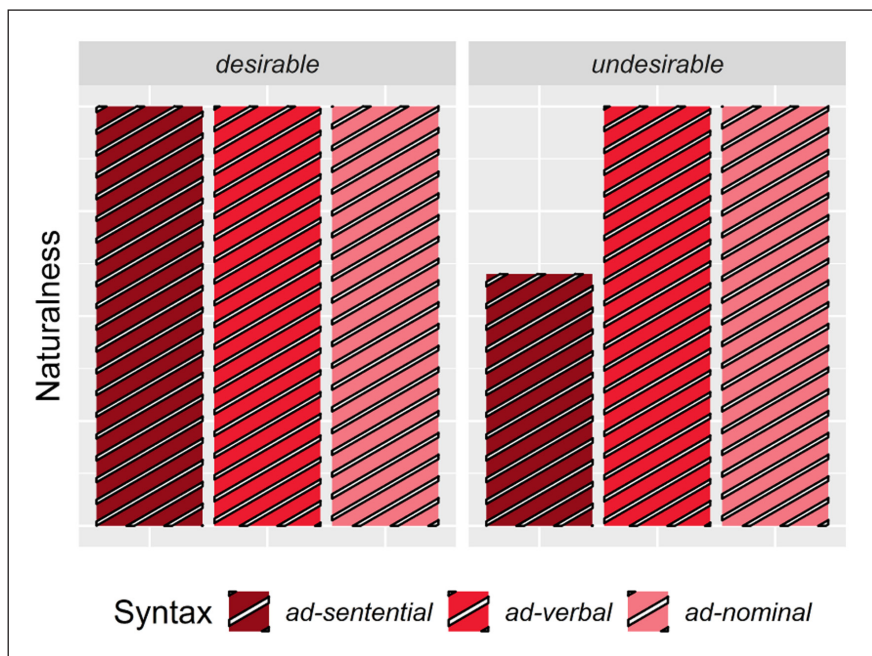
The procedure was identical to Experiment 1. A demo link to the experiment can be found [here](#).

### 3.2.3 Participants

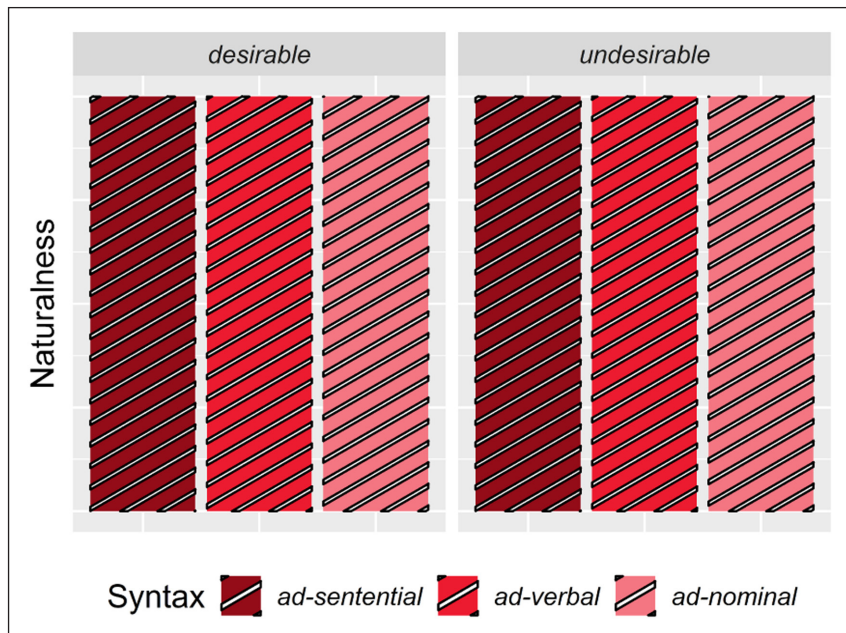
26 participants were recruited from Prolific.ac and were paid US\$1.60 each. Two participants were excluded due to at least 25% of trials having a response time below 2s or above 40s, with 24 participants remaining for data analysis.

### 3.2.4 Predictions & Coding

The symmetric and the asymmetric view differ on the availability of evaluative *at least*, not epistemic *at least*, with the latter being assumed to be ruled out in ad-sentential position by both accounts. The evaluative reading is assumed to be associated with a desirability ranking, resulting in infelicity if the asserted alternative is not in fact preferable to relevant alternatives, as is the case in undesirable contexts. If an epistemic interpretation of *at least* is impossible in ad-sentential position, we should see a deprecation in ratings in such cases, while the other positions, which uncontroversially allows for an epistemic reinterpretation, should not show this effect. On the other hand, if the epistemic reading is possible in ad-sentential position, as the results of Experiment 1 seem to suggest, then we should not see a relative decrease of ad-sentential position in undesirable contexts. For desirable contexts, where either interpretation should be felicitous, no differences between positions is predicted, presupposing again an even baseline. An idealized pattern of results for the two possibilities is again shown in **Figures 4 and 5**.



**Figure 4:** Idealized pattern predicted if ad-sentential position disallows epistemic interpretation, Experiment 2.



**Figure 5:** Idealized pattern predicted if ad-sentential position allows epistemic interpretation, Experiment 2.

As a complicating factor, however, Biezma (2013) and Chen (2018) additionally assume that an evaluative interpretation only arises when higher alternatives are known to be false. Since the stimuli use mutually compatible alternatives that stand in an entailment relationship, this requirement would not be met. In principle, *some* in the target utterance could be interpreted exhaustively as *some but not all*, but this would in turn conflict with the requirement of epistemic *at least* that higher alternatives are left open. These two accounts thus predict that there should not be any decrease for the ad-sentential position in undesirable contexts, since the discourse conditions for an evaluative interpretation are not met, even if the syntax would make it available.<sup>12</sup> Their predictions would thus conform to the pattern in **Figure 5** as well.

Since the differences in predictions concern the status of ad-sentential *at least* in undesirable contexts, it might seem natural to use both as – dummy coded – reference levels of their respective factors. However, we saw ad-verbal *at least* trend toward an intermediate status for evaluative contexts in Experiment 1 due to its ambiguity, which might also be the case here, while the ad-nominal position should be restricted to an epistemic interpretation. A comparison of the ad-verbal position to the ad-nominal position might thus be more informative than a comparison

<sup>12</sup> Notably, they still assume that an epistemic interpretation should be unavailable in ad-sentential position, resulting in the odd prediction that ad-sentential *at least* in a context where higher alternatives are left open should not be interpretable either epistemically or evaluatively. Such a state of affairs could also be reflected in decreased ratings but also raises further questions.

of ad-verbal with ad-sentential. We will thus code ad-nominal as the dummy reference level. The most relevant statistical effects are then *ad-sentential.VS.ad-nominal*, as well as its interaction with *desirable.VS.undesirable*. The status of ad-verbal position relative to ad-nominal is more explorative in this respect.

### 3.2.5 Results

Data analysis was the same as for Experiment 1, the output given in **Table 4**. The average rating per condition is shown in **Figure 6**. First looking at ratings for the undesirable context, we see a stairlike pattern, with ad-sentential *at least* receiving the lowest and ad-nominal the highest ratings, and ad-verbal position in between. Both decreases relative to ad-nominal position are statistically significant ( $z = -6.10, p < .001^{***}$ ;  $z = -2.64, p < .01^{**}$ ). In desirable contexts, we see a similar pattern, but much more subtle, supported by a significant interaction of *ad-sentential.VS.ad-nominal* and *desirable.VS.undesirable* ( $z = 3.00, p < .01^{**}$ ).

	Est.	Std. Err.	z-value	p-value
des VS UNDES	0.234	0.156	1.502	.133
ad-sen VS AD-NOM	-0.952	0.156	-6.101	$1.05e^{-09***}$
ad-ver VS AD-NOM	-0.410	0.155	-2.638	.008**
des VS UNDES * ad-sen VS AD-NOM	0.656	0.219	2.997	.003**
des VS UNDES * ad-ver VS AD-NOM	0.243	0.221	1.104	.270

**Table 4:** Experiment 2 model output: `clmm(Rating ~ Context*Syntax + (1|Subject) + (1|Item))`. CONTEXT dummy coded with *undesirable* as reference level [des = 1, undes = 0], SYNTAX dummy coded with ad-nominal as reference level [ad-sen = (1, 0), ad-ver = (0, 1), ad-nom = (0, 0)]. Small capitals indicate reference level of the relevant factor.

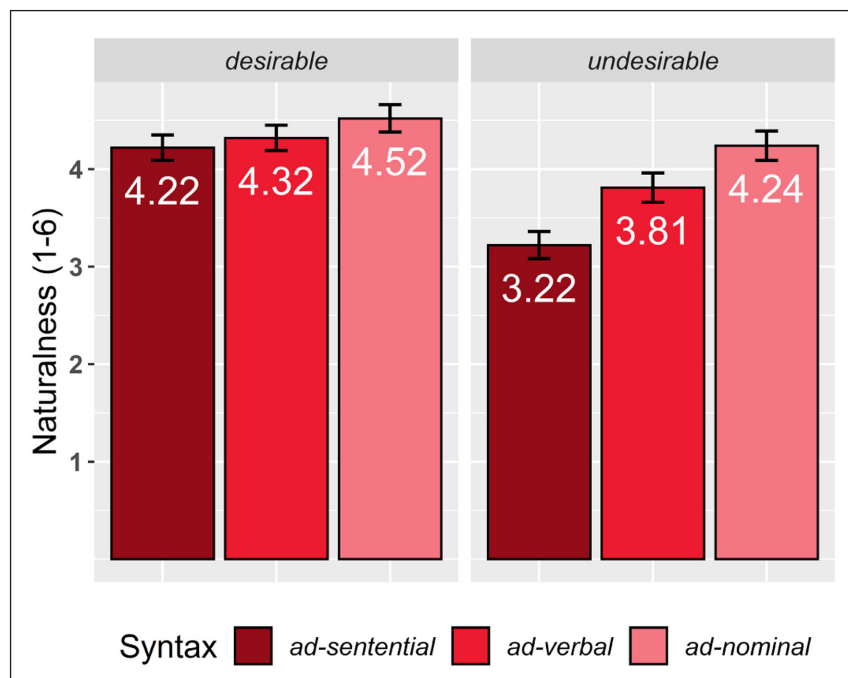
### 3.2.6 Discussion

The results provide evidence for ad-sentential *at least* inducing a desirability ranking: in undesirable contexts, where ranking alternatives according to their desirability was inconsistent with the discourse and social norms, ad-sentential *at least* received lower ratings than in ad-nominal position, where only the epistemic interpretation is available.<sup>13</sup> This difference was smaller in

---

<sup>13</sup> The intermediate status of ad-verbal *at least* can be accounted for by appealing to either a cost due to the participants having to determine which interpretation to choose given that both the epistemic and the evaluative one are available, or to participants sometimes choosing the “wrong” interpretation and not being able to reanalyze. These two explanations could be distinguished based on the distribution of individual ratings: if participants never reanalyze but sometimes choose incorrectly, the distribution should be bimodal, whereas a general cost of reanalysis should yield a unimodal distribution of slightly reduced ratings.

desirable contexts that are compatible with a desirability ranking. This outcome is thus in line with the view that ad-sentential *at least* is restricted to an evaluative interpretation and that this interpretation has such a desirability ranking: if an epistemic interpretation was available in ad-sentential position and came without a desirability ranking, no decrease should be expected. Similarly, if higher alternatives were required to be false for the evaluative interpretation and its desirability ranking to be available, the explanation of the decrease of ad-sentential *at least* in undesirable contexts adopted above would no longer work. Therefore, the data suggest that evaluative *at least* does not require higher alternatives to be false to become available, contrary to Biezma (2013) and Chen (2018).



**Figure 6:** Mean ratings (with standard errors) by condition, Experiment 2.

These findings furthermore bear on the interpretation of the results of Experiment 1. As a reminder, Experiment 1 failed to provide evidence for the ad-sentential position being restricted to an evaluative interpretation when using the assumed prejacent entailment component of this interpretation, while finding support for an interpretive restriction of the ad-nominal position to epistemic *at least* with regard to higher alternatives being left open. This contrast was compatible with either a (differently specified) asymmetry in the syntax-semantics mapping – with ad-sentential *at least* being compatible with either interpretation but ad-nominal *at least* being restricted to an epistemic one – or there being no entailment component associated with evaluative *at least*. Since the present experiment did find support

for an interpretive restriction of *at least* in ad-sentential position, we thus have support for the symmetrical view for both epistemic and evaluative *at least*, based on one property specific to the respective interpretation. As a consequence, the explanation that is left for the asymmetry observed in Experiment 1 is that evaluative *at least* in fact does not entail the truth of its prejacent, contrary to what has been assumed in the previous literature. The next experiment follows up on this issue.

### 3.3 Experiment 3: Truth of Prejacent & Desirability Ranking

#### 3.3.1 Design & Materials

The main goal of this experiment was to reassess the finding from Experiment 1 that there was no penalty for ad-sentential *at least* when the ad-sentential position was compatible with uncertainty about the prejacent. This pattern is unexpected if the ad-sentential position disambiguates toward the evaluative reading, which all current analyses assume entails the prejacent. Nakanishi & Rullmann (2009) offer a potential explanation for this result: They claim that the ad-sentential position only leads to a preference for an evaluative reading, but take the epistemic reading to be available as well. So just based on the null effect in Experiment 1 we might conclude that an epistemic interpretation is available in ad-sentential position if the context makes it sufficiently clear that the speaker is uncertain about the prejacent. Experiment 2, however, contradicts this interpretation: The evaluative effect in ad-sentential position provides evidence for a disambiguation toward the evaluative reading after all. The question then is whether evaluative *at least* always entails its prejacent. There are two aspects of the previous experiments we modified to answer this question. First, in light of the concerns described in Section 3.1.6 regarding Experiment 1's sensitivity to finding an effect of entailment, we included a control condition without *at least* since the entailment status of a "bare" proposition should be uncontroversial. To keep the design compact, the ad-verbal position was dropped since it is of lesser interest given the research question here. Additionally, the target sentence was adjusted to make a possible conflict more explicit. Secondly, we tested the entailment component in conjunction with the desirability component from Experiment 2 as a way to test whether a positional effect is in principle present.

The basic design was the same  $2 \times 3$  as Experiment 2, crossing CONTEXT (*desirable* vs *undesirable*) and SYNTAX, but with a bare prejacent as *bare control* replacing ad-verbal *at least*. As an additional between-item factor that was already employed in Experiment 1 but proved inconsequential (see footnote 8), we manipulated the ALTERNATIVE TYPE: alternatives were either mutually *compatible* as in Experiment 2, shown in (19), or mutually *exclusive*, as in (20).



(19) Sample Item Experiment 3, Mutually Compatible Alternativesa. *Desirable* CONTEXT:

A: You were really hoping that Josephine might pass the students.  
Do you know how many students she passed?

B: I'm not sure. (**At least**) She passed (**at least**) some of the students, but I'm hoping she in fact passed all of them.

b. *Undesirable* CONTEXT:

A: You were really hoping that Josephine might not pass the students.  
Do you know how many students she passed?

B: I'm not sure. (**At least**) She passed (**at least**) some of the students, but I'm worried she in fact passed all of them.

(20) Sample Item Experiment 3, Mutually Exclusive Alternativesa. *Desirable* CONTEXT:

A: You were really hoping that Yvette might win gold at the race.  
Do you know what place she ended up getting?

B: I'm not sure. (**At least**) She won (**at least**) silver, but I'm hoping she in fact won gold.

b. *Undesirable* CONTEXT:

A: You were really hoping that Yvette might not win gold at the race.  
Do you know what place she ended up getting?

B: I'm not sure. (**At least**) She won (**at least**) silver, but I'm worried she in fact won gold.

This additional manipulation was necessary since it is only with mutually exclusive alternatives that the two interpretations are taken to differ in the status of the prejacent: since *some* is compatible with *all*, the truth of the two alternatives is not at odds with each other, whereas *win silver* and *win gold* cannot be true at the same time (see Section 2.2). The experiment thus essentially combines aspects of the two previous experiments by taking the design of Experiment 2 and adding exclusive alternatives to it. Using exclusive alternatives, however, necessitated adjusting the way the desirability component was manipulated, since there were not enough antonym-like pairs available. Instead, A's statement varied in ascribing B either an interest in a positive or a negative outcome, which was also indicated by the contrast between *hoping* and *worried* in the target sentence. The present experiment thus fully relied on varying the context rather than the target clause.

There were 24 item sets, evenly split by ALTERNATIVE TYPE, in addition to 24 fillers. Of the fillers, 12 were a variant of the *even*-fillers from Experiment 2, 4 were bad catch-fillers with *even* rated low in Experiment 2, and 8 were items with epistemic and evaluative *at least*, respectively,

intended to be maximally acceptable. All items are again available in the OSF repository, see Supplementary Files.

### 3.3.2 Procedure

The procedure was identical to Experiments 1 and 2. A demo link is provided [here](#).

### 3.3.3 Participants

38 participants were recruited from Prolific.ac and paid US\$2.00 each. Two participants were excluded due to having more than 25% of trials with response times below 2 seconds, leaving 36 participants for data analysis. The number of participants was increased relative to Experiments 1 and 2 given the additional between-item factor.

### 3.3.4 Predictions & Coding

The first base prediction concerned validation that the experimental design is sensitive to the manipulation meant to get at the entailment status of the target clause: for mutually exclusive alternatives, taking the target clause (= *she won silver*) to be true should be inconsistent with the speaker first conveying uncertainty (“*I’m not sure*”) and the continuation that expresses hope about an alternative – no longer possible – outcome. In contrast, when alternatives are mutually compatible there should be no conflict between the target clause (= *she passed some of the students*) and speaker uncertainty about or speaker desire for the truth of higher alternatives (= *she passed all of the students*). Since the truth of the target clause should be uncontroversially entailed in the absence of *at least* and equally uncontroversially left open with ad-nominal *at least* (see also results from Experiment 1), we predict the difference in ratings between the bare control and ad-nominal *at least* to be larger for mutually exclusive alternatives when there is the potential for conflict than for mutually compatible alternatives when there should be none, resulting in an interaction of the comparison between the bare control and ad-nominal *at least* and the comparison between mutually compatible and mutually exclusive alternatives. We will refer to this prediction as the Baseline Prediction. An idealized illustration of the prediction being borne out or not, focusing only on the most relevant conditions, is given in **Figures 7 and 8**.

The more crucial second prediction concerns ad-sentential *at least*: if ad-sentential *at least* entails the truth of its prejacent – in this case the target clause – it should pattern with the bare control and contrast with ad-nominal *at least*, reflected in an interaction of the comparison between ad-sentential and ad-nominal *at least* and ALTERNATIVE TYPE; if, however, ad-sentential *at least* does not entail the truth of its prejacent – as the results of Experiment 1 suggest – then it should pattern with ad-nominal *at least* instead. We will call this prediction the *At least* Entailment Prediction. Corresponding illustrations are given in **Figures 9 and 10**.



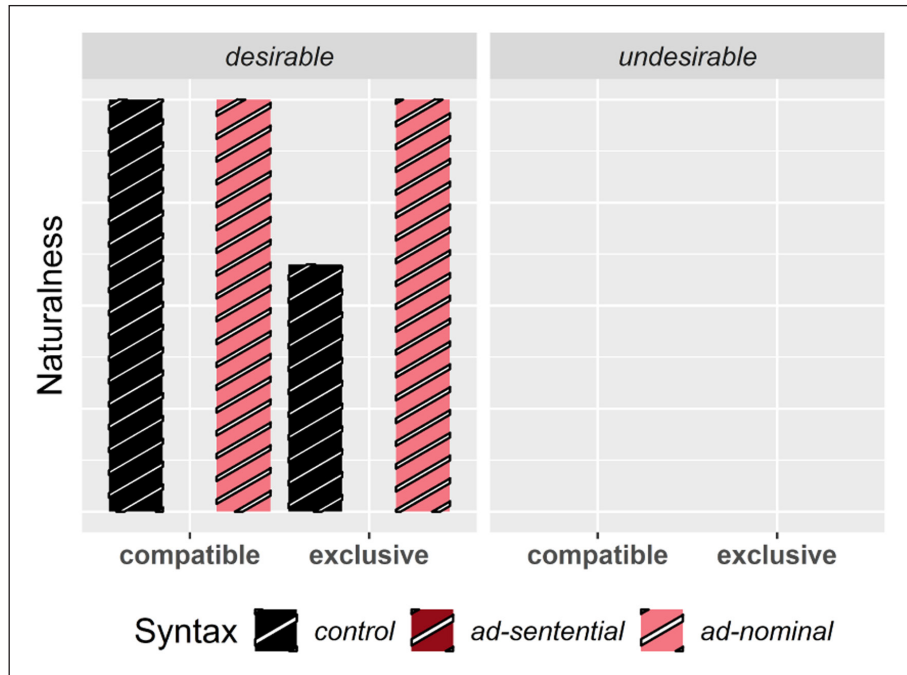


Figure 7: Idealized pattern for Baseline Prediction, Experiment 3.

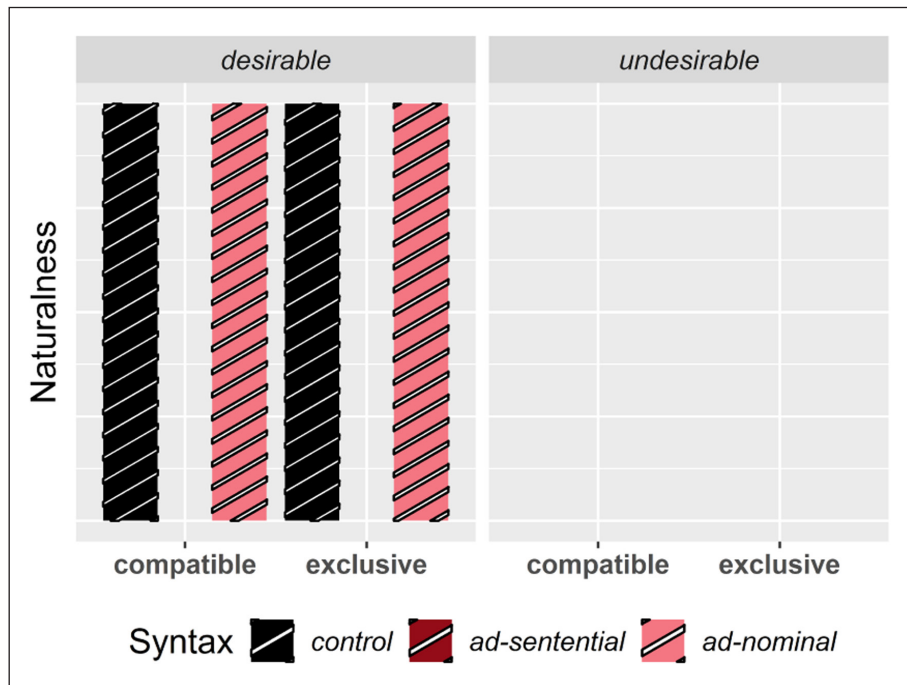


Figure 8: Idealized pattern inconsistent with Baseline Prediction, Experiment 3.

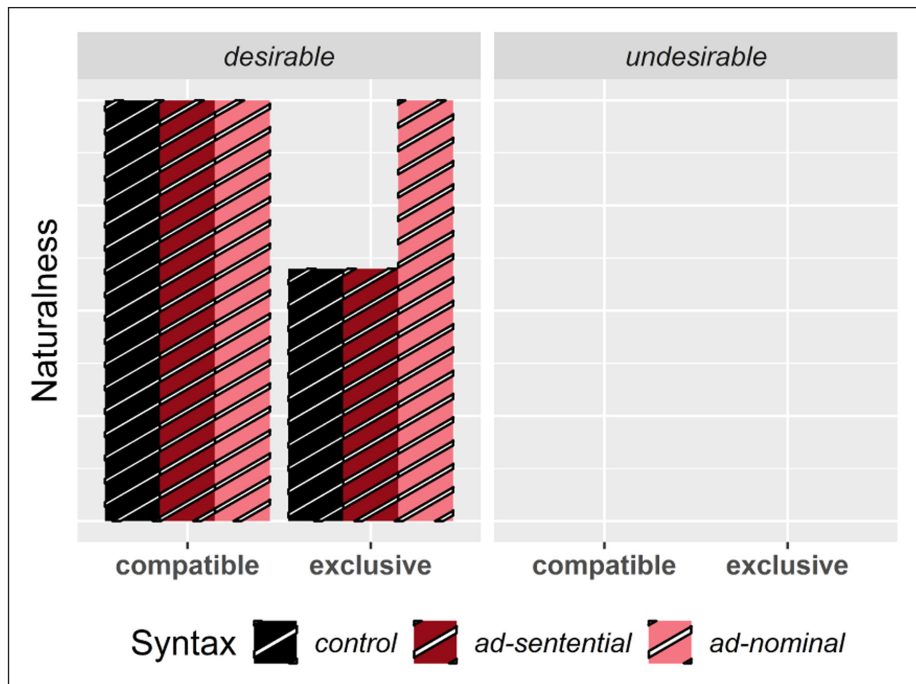


Figure 9: Idealized pattern for Entailment Prediction, Experiment 3.

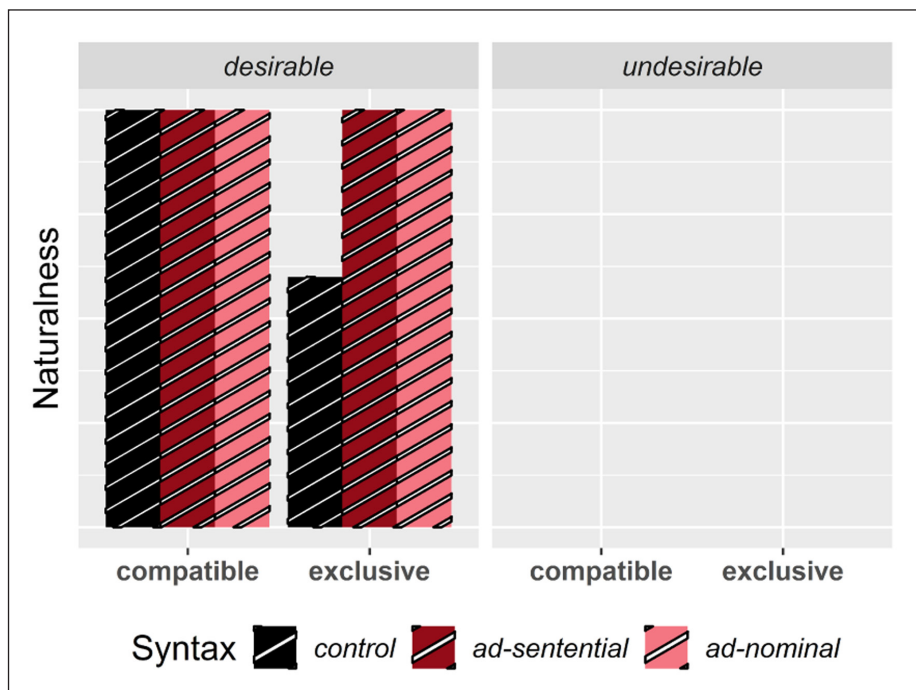
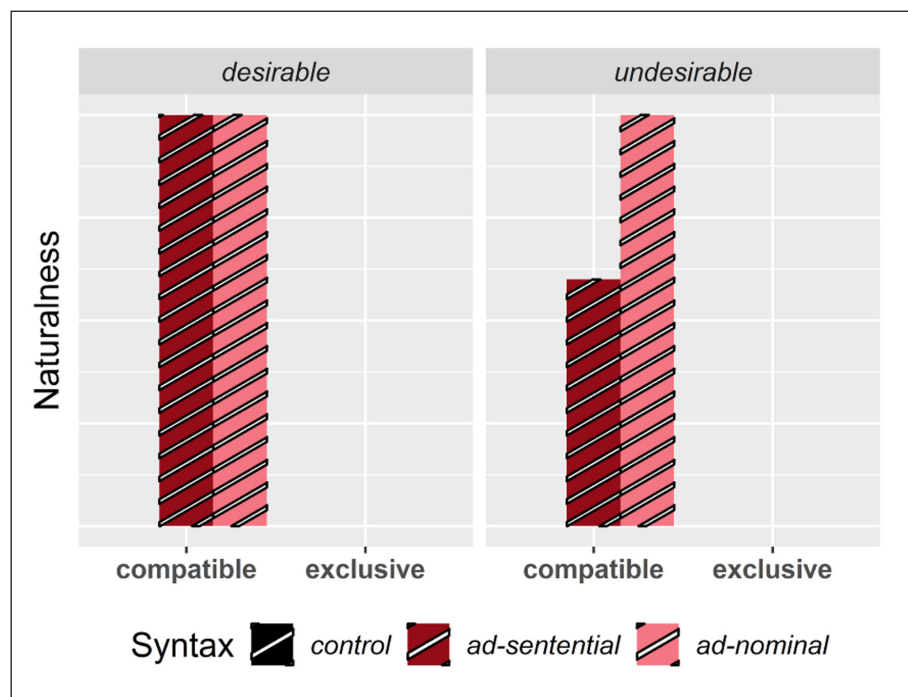


Figure 10: Idealized pattern inconsistent with Entailment Prediction, Experiment 3.

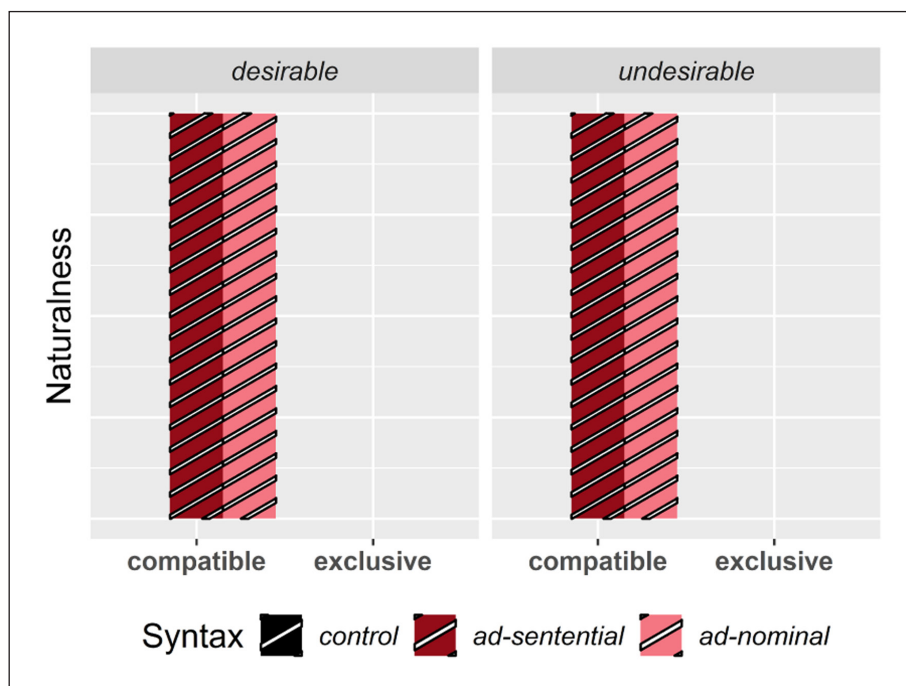
In addition to these main predictions, there are two further relevant aspects of the design. First, the present design subsumes the design of Experiment 2 and therefore allows us to see whether the observed pattern of results regarding the desirability ranking replicates. Thus, we expect there to be a larger difference between ad-sentential and ad-nominal *at least* in undesirable contexts than in desirable contexts due to ad-sentential position necessarily evoking a desirability ranking that violates the discourse. The relevant term would thus be an interaction of the ad-sentential and ad-nominal *at least* difference and CONTEXT. Let's call this the Replication Prediction, with illustrations given in Figures 11 and 12.



**Figure 11:** Idealized pattern for Replication Prediction, Experiment 3.

Additionally, by virtue of the added ALTERNATIVE TYPE manipulation, the experiment allows us to reassess the question of whether higher alternatives need to be false for an evaluative interpretation to be available. Experiment 2 only used mutually compatible alternatives and left higher alternatives implicitly open in the context such that the finding is at odds with Biezma's (2013) and Chen's (2018) assumption that evaluative *at least* – and its desirability ranking – require higher alternatives to be false and hence should not have been available. Viewing the present experiment with respect to this issue, these accounts would thus predict that the desirability violation – the decrease of ad-sentential *at least* relative to ad-nominal *at least* in undesirable contexts compared to desirable contexts – should only occur with mutually exclusive

alternatives but not mutually compatible alternatives. The reason for that is that mutually compatible alternatives should leave higher alternatives open, failing to fulfill the requirement for evaluative *at least* to become available; in contrast, mutually exclusive alternatives could in principle rule out higher alternatives and make evaluative *at least* available – if the pre-jacent is taken to be true. This prediction – which we will dub the False Alternative Prediction – thus depends on the *At least* Entailment Prediction to be borne out, again illustrated in **Figures 13** and **14**. The relevant effect would be a three-way interaction of the comparison between ad-sentential and ad-nominal *at least*, CONTEXT, and ALTERNATIVE TYPE.



**Figure 12:** Idealized pattern inconsistent with Replication Prediction, Experiment 3.

Since all predictions feature either a comparison between the bare control and ad-nominal *at least* (Baseline Prediction) or ad-sentential and ad-nominal *at least* (*At least* Entailment Prediction, Replication Prediction, False Alternative Prediction), SYNTAX will be dummy coded with ad-nominal as reference level. Secondly, CONTEXT will also be dummy coded, with desirable as reference level, to provide the most neutral grounds when assessing the Baseline Prediction. Finally, ALTERNATIVE TYPE will be dummy coded as well, with mutually compatible as reference level, in order to make it most comparable to Experiment 2, which used this type of alternative. As a consequence, the simple effects of SYNTAX will compare differences between the bare control and ad-nominal *at least*, and ad-sentential and ad-nominal *at least* for mutually compatible alternatives in desirable contexts.

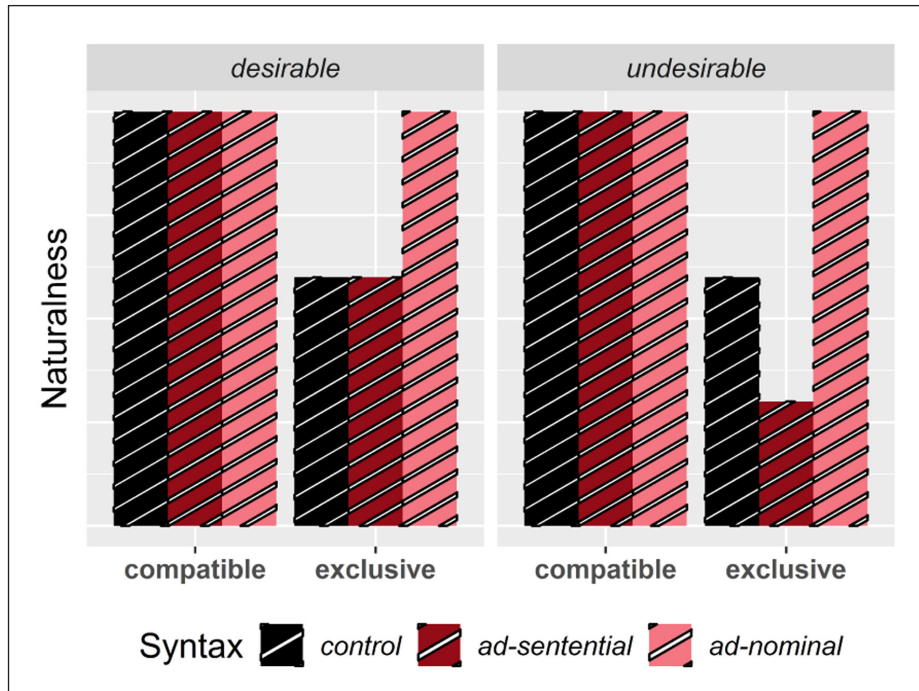


Figure 13: Idealized pattern for False Alternative Prediction, Experiment 3.

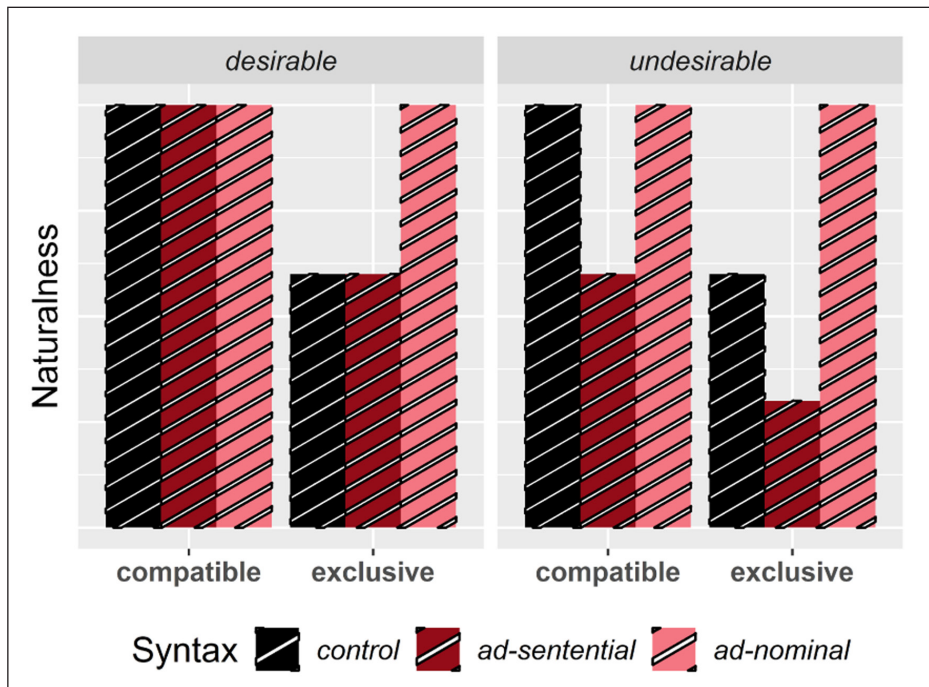


Figure 14: Idealized pattern inconsistent with False Alternative Prediction, Experiment 3.

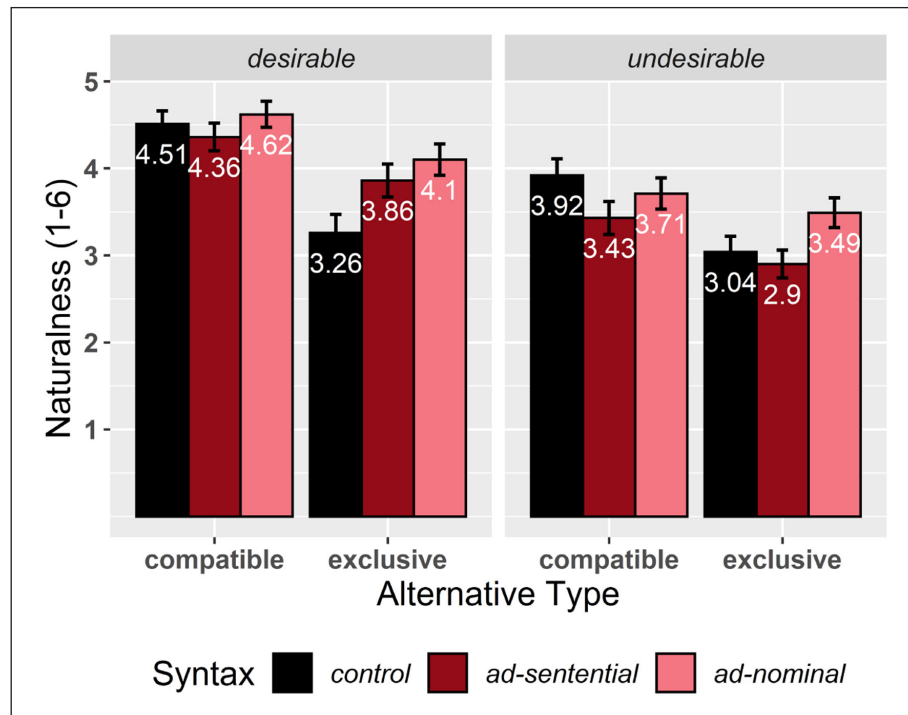
### 3.3.5 Results

Data were again analyzed with an ordinal mixed effects model, as in Experiments 1 and 2, with the full output shown in **Table 5**. The average ratings by condition are shown in **Figure 15**. Focusing on the effects of interest one by one, we can first look at desirable contexts only and the difference between the bare control and ad-nominal position for the Baseline Prediction. While bare control and ad-nominal position are numerically close with mutually compatible alternatives, the bare control decreases in ratings relative to ad-nominal position with mutually exclusive alternatives, supported by a significant interaction of *barecontrol.VS.ad-nominal* and *compatible.VS.exclusive* ( $z = -2.407$ ,  $p < .05^*$ ). Next looking at how ad-sentential position patterns relative to this comparison to assess the Entailment Prediction, the difference between ad-sentential and ad-nominal position seems to remain stable across alternative type in desirable contexts, as the interaction between *ad-sentential.VS.ad-nominal* and *compatible.VS.exclusive* does not reach significance ( $z = 0.386$ ,  $p = .700$ ).

	Est.	Std. Err.	z-value	p-value
DES VS undes	-0.799	0.180	-4.441	<.001***
bare VS AD-NOM	-0.084	0.181	-0.461	.645
ad-sen VS AD-NOM	-0.277	0.180	-1.537	.124
COMP VS excl	-0.454	0.205	-2.218	.027*
DES VS undes * bare VS AD-NOM	0.277	0.253	1.097	.273
DES VS undes * ad-sen VS AD-NOM	0.078	0.252	0.310	.756
DES VS undes * COMP VS excl	0.270	0.251	1.075	.283
bare VS AD-NOM * COMP vs excl	-0.611	0.254	-2.407	.016*
ad-sen VS AD-NOM * COMP vs excl	0.098	0.253	0.386	.700
DES VS undes * bare VS AD-NOM * COMP vs excl	0.031	0.355	0.088	.930
DES VS undes * ad-sen VS AD-NOM * COMP vs excl	-0.353	0.355	-0.994	.320

**Table 5:** Experiment 3 model output: `clmm(Rating ~ Context * Syntax * Alternative Type + (1|Subject) + (1|Item))`. CONTEXT coded as [des = 0, undes = 1], SYNTAX coded as [bare = (1, 0), ad-sen = (0, 1), ad-nom = (0, 0)], ALTERNATIVE TYPE coded as [comp = 0, excl = 1]. Small capitals indicate reference level of the relevant factor, if applicable.





**Figure 15:** Mean ratings (with standard errors) by condition, Experiment 3.

Next, we can look at differences between contexts to assess the remaining two predictions. Comparing the results for mutually compatible alternatives across desirable and undesirable contexts, the patterns seem largely the same, and there was indeed no significant interaction of *ad-sentential.VS.ad-nominal* and *desirable.VS.undesirable* ( $z = 0.310$ ,  $p = .756$ ). The latter interaction would thus speak against the Replication Prediction. In contrast, when looking at differences between contexts for mutually exclusive alternatives, there does seem to be the expected pattern that *ad-sentential* *at least* decreases more in undesirable contexts than *ad-nominal* *at least*. However, since the relevant three-way interaction between *ad-sentential.VS.ad-nominal*, *desirable.VS.undesirable*, and *compatible.VS.exclusive* did not reach significance ( $z = -0.994$ ,  $p = .320$ ), we conducted a pairwise comparison for levels of SYNTAX within ALTERNATIVE TYPE and CONTEXT for our model to further investigate this pattern, using the `lsmeans` package (Lenth 2016), see **Table 6** for details. There, the only significant comparison between *ad-sentential* and *ad-nominal* position is the one for undesirable contexts and mutually exclusive alternatives ( $z = -2.585$ ,  $p < .05^*$ ), in line with the visual impression. Although this pattern hence suggests weak evidence in favor of the False Alternative Prediction, which predicted an effect of desirability for mutually exclusive alternatives only, it just does so at first glance since the *At least* Entailment Prediction was not borne out, which will be discussed in detail in the next section.

	Est.	Std. Err.	z-value	p-value
<i>desirable + compatible</i>				
bare VS ad-sen	0.193	0.180	1.076	.529
bare VS ad-nom	-0.084	0.181	-0.461	.890
ad-sen VS ad-nom	-0.277	0.180	-1.537	.274
<i>desirable + exclusive</i>				
bare VS ad-sen	-0.516	0.178	-2.893	.011*
bare VS ad-nom	-0.695	0.178	-3.905	<.001***
ad-sen VS ad-nom	-0.179	0.178	-1.010	.571
<i>undesirable + compatible</i>				
bare VS ad-sen	0.392	0.177	2.217	.068*
bare VS ad-nom	0.194	0.176	1.101	.514
ad-sen VS ad-nom	-0.199	0.176	-1.128	.497
<i>undesirable + exclusive</i>				
bare VS ad-sen	0.068	0.176	0.385	.921
bare VS ad-nom	-0.386	0.176	-2.200	.071*
ad-sen VS ad-nom	-0.455	0.176	-2.585	.026*

**Table 6:** Experiment 3 model output for pairwise comparisons for SYNTAX within ALTERNATIVE TYPE and CONTEXT.

### 3.3.6 Discussion

We will assess each of the four predictions in turn. First, the Baseline Prediction regarding the control condition without *at least* was borne out: while the ratings for the bare prejaent and ad-nominal *at least* in the desirable context with mutually compatible alternatives were not significantly different from each other, ratings for the bare prejaent decreased more with mutually exclusive alternatives. This effect follows from the prejaent being asserted as true in the absence of *at least*, resulting in a conflict with the speaker conveying uncertainty regarding other alternatives in a situation when these alternatives are mutually exclusive and consequently ruled out by a true prejaent. This conflict does not occur with ad-nominal *at least* where alternatives are conveyed to be left open independently of their logical relationship. The data thus provide evidence for the sensitivity of the materials to these types of semantic-pragmatic conflicts and allow us to interpret the behavior of ad-sentential *at least* in this respect more conclusively in the following.



Regarding the crucial *At least* Entailment Prediction, we failed to find evidence for ad-sentential *at least* necessarily entailing its prejacent. Looking at the same difference between mutually compatible alternatives and mutually exclusive alternatives in desirable contexts, there was no evidence for a change in the difference between ad-sentential *at least* and ad-nominal *at least* for ALTERNATIVE TYPE, contrary to what we would have expected if ad-sentential *at least* always conveyed the truth of its prejacent. Note that this cannot be attributed to a lack of statistical power, since the analogous effect was present for the bare control. This finding thus suggests that the failure to find an entailment conflict in Experiment 1 was not due to the manipulation being insufficiently sensitive but simply that there was no effect to be found.

Turning to the data bearing on the type of ranking associated with *at least*, neither the Replication Prediction – which predicted a decrease for ad-sentential *at least* compared to ad-nominal *at least* to be larger in undesirable contexts than desirable contexts – nor the False Alternative Prediction – which predicted the effect expected on the Replication Prediction to be restricted to mutually exclusive alternatives – was borne out directly. Looking only at mutually compatible alternatives, Experiment 2 found a decrease in ratings for ad-sentential and ad-verbal *at least* compared to *at least* in ad-nominal position when a desirability ranking would lead to a conflict in the discourse, which was not present when such a ranking was consistent with the discourse. The analogous comparison between ad-sentential and ad-nominal *at least* in the present experiment, however, did not show such a mediation by context – contrary to the Replication Prediction. On the other hand, post-hoc pairwise comparisons showed that the two syntactic positions differed in undesirable contexts for mutually exclusive alternatives, suggesting that the expected effect of context was not completely absent.

In fact, this pattern would be in line with the False Alternative Prediction, although the crucial three-way interaction was not significant (this may be due to lack of statistical power). However, the False Alternative Prediction relied on the entailment of ad-sentential *at least*, which we found no evidence for: if ad-sentential *at least* does not by default entail its prejacent, then higher alternatives are left open both for mutually exclusive and mutually compatible alternatives, and an evaluative interpretation is not supposed to arise in the first place. We are thus left with the question why the evidence for ad-sentential *at least* inducing a desirability ranking is weak compared to Experiment 2. We want to suggest two possible explanations.

First, items varied in whether the context manipulation was done by adding *not* to A's statement, as in (19)/(20), or by changing *hoping* to *concerned*, which was evenly split by items. While this alteration should not have led to the apparent asymmetry between alternatives where the pattern found in Experiment 2 seemed to be only present for mutually exclusive alternatives but not mutually compatible ones, it may have resulted in misparses: participants might have paid more attention to the initial attitude verb and occasionally overlooked the negation, increasing overall variation and weakening the effect of context. A second possibility is that participants

focused on the entailment manipulation such that the sensitivity to the desirability manipulation was weakened or overridden. That is, participants may have adopted a criterion for providing a rating that focused on logical inconsistency rather than contextual inconsistency. To conclude, while the results on their own thus provide only weak evidence in support of a desirability ranking associated with ad-sentential *at least*, we will nonetheless assume that ad-sentential *at least* does in fact necessarily involve such a desirability ranking in light of the results from Experiment 2, and leave an investigation into the possible explanations for the variation for future research.

## 4 General Discussion

This paper presented results from three naturalness rating experiments to assess what properties are associated with *at least* in different positions. Experiment 1 provided evidence that the ad-nominal position requires higher alternatives to be left open, but failed to provide evidence that there are syntactic restrictions on when *at least* entails its prejacent. Experiment 2 showed that the ad-sentential position is associated with a ranking that orders alternatives according to their desirability. Experiment 3 followed up on the lack of entailment effect from Experiment 1 and showed that participants are sensitive to contradictions caused by entailed alternatives being incompatible with each other while failing to find such an effect with ad-sentential *at least*. Both Experiment 1 and Experiment 3 thus failed to provide any evidence that the evaluative reading necessarily entails the prejacent. An overview of the results is shown in **Table 7**, with the three relevant properties and the interpretation they have been linked to related to the syntax of *at least*.

	<i>ad-sentential</i>	<i>ad-verbal</i>	<i>ad-nominal</i>
(i) higher ALT required to be open (EPI)	✗	✗	✓
(ii) prejacent necessarily entailed (EVAL)	✗	✗	✗
(iii) prejacent necessarily desirable (EVAL)	✓	✗	✗

**Table 7:** Overview of experimental results regarding the syntax-semantics mapping of *at least*.

Regarding our main question about the relation between *at least*'s syntactic position and its available interpretations, these results thus provide evidence for the symmetrical view by Nakanishi & Rullmann (2009) and Chen (2018) – according to which the epistemic interpretation is unavailable ad-sententially and the evaluative interpretation is unavailable ad-nominally – and against the asymmetrical view by Biezma (2013) and others: higher alternatives have to be left open in ad-nominal position because *at least* is restricted to an epistemic interpretation there, and alternatives have to be ranked according to desirability in ad-sentential position because only an

evaluative interpretation is available. On a view that allows *at least* to be interpreted more freely in either of these two positions the experimental results would not be explained without further assumptions.

However, the results also revealed a tension with prior claims about the semantic-pragmatic properties of *at least*: contrary to the generally adopted assumption that the evaluative interpretation entails the truth of its prejacent, ad-sentential *at least* patterned with ad-nominal *at least* rather than the bare prejacent in a context where a true prejacent should result in infelicity.<sup>14</sup>

The natural question this raises is how this finding fares with the formal accounts of *at least* discussed in Section 2.2. Since all three relevant accounts took the prejacent entailment of evaluative *at least* for granted, we can ask how easily they can be adjusted to capture the new data.

On Nakanishi & Rullmann's (2009) account, the prejacent entailment is implemented as part of the assertion of evaluative *at least*, which could easily be dropped without additional consequences. In contrast, Biezma (2013) and Chen (2018) derive the entailment through an interaction of three components: non-entailed alternatives below the prejacent being ruled out (either by directly being presupposed as false (Biezma 2013), or indirectly via domain restriction (Chen 2018)), the requirement that higher alternatives are false for an evaluative interpretation to arise, and some alternative in the alternative set being true. While this more elaborate derivation may make a minimal adjustment to capture the data seem less straightforward, the experimental results crucially bear on one of those three components: Experiment 2 yielded evidence that the evaluative interpretation is available even when higher alternatives are not known to be false, given that a conflict with the desirability ranking associated with this interpretation was reflected in decreased ratings. Once this requirement is retracted – in line with the empirical evidence – the prejacent entailment disappears as well from Biezma and Chen's accounts. The results can thus be captured by all three accounts with minimal adjustments.<sup>15</sup>

We may nonetheless wonder how these assumptions about the evaluative interpretation concerning its entailment status and the falsity of higher alternatives came to be. What we want to suggest here is that these claims pick up on how evaluative *at least* may be used most frequently: by virtue of contributing a positive counterpoint analogous to “it could have been worse”, evaluative *at least* will most naturally be used in response to a negative outcome, which corresponds to higher, or in this case better, alternatives being false. The falsity of a higher

---

<sup>14</sup> Since the presented experiments relied on naturalness ratings as measures, a follow-up study may directly assess the inferences participants draw to further examine this issue. Thanks to an anonymous reviewer for this suggestion.

<sup>15</sup> Note that this view does not mean that evaluative *at least* is thus equivalent to epistemic *at least*, with the exception of the desirability ranking. Rather, evaluative *at least* is merely compatible with what might be called an epistemic interpretation by leaving the truth of higher alternatives open, while epistemic *at least* requires higher alternatives to be open, as Experiment 1 showed.

alternative and the resulting prejacent entailment may thus be viewed as inferences about the type of contexts evaluative *at least* is used in rather than a property inherent to the meaning of evaluative *at least* itself.

The next section concludes the paper with some further remarks for future research.

## 5 Concluding Remarks

This paper started with the question of how the syntactic position of *at least* determines its possible interpretations, motivated by previous conflicting claims in the literature. In three experiments, we found evidence in favor of a symmetrical view where ad-sentential *at least* is restricted to an evaluative interpretation and ad-nominal *at least* to an epistemic interpretation. We also found evidence against the claim that evaluative *at least* entails the truth of its prejacent. We argued that these new data can be straightforwardly implemented into existing formal accounts once the assumption that an evaluative interpretation requires higher alternatives to be false is given up, and suggested that this misconception is due to an extrapolation about what contexts evaluative *at least* is used in most commonly.

The above articulated conjecture about where this misconception comes from reveals a disparity of formal accounts to evaluative *at least*, however, in particular Chen (2018). Spelling out the contribution of the evaluative interpretation according to Chen’s entry, repeated in (21), *at least* conveys that there is a true proposition in the set of alternatives as good as or better than the prejacent, which could be paraphrased as “it could have been better”. This paraphrase directly contrasts with the intuition that evaluative *at least* communicates a sense of optimism by comparing the state of affairs with a worse outcome. An empirical argument for this view can be found in relation to the contrastive connective *but*. Of the two paraphrases, the one “looking up” in (22a) is infelicitous, while the one “looking down” in (22b) is not. Crucially, the evaluative interpretation in (22c) patterns with the latter, not the former.

$$(21) \quad \llbracket \text{at least}(C) \rrbracket^{\text{w,c}} = \lambda \alpha_{\langle s,t \rangle}. \exists \gamma [\gamma \in C \wedge \gamma_w \wedge \forall \beta [\beta \in C \wedge \beta \neq \alpha \rightarrow \mu_c(\alpha) < \mu_c(\beta)]]$$

- (22) a. A: It’s too bad Emma only won silver.  
 B: #True, but it could’ve been better. She could’ve won gold.
- b. A: It’s too bad Emma only won silver.  
 B: True, but it could’ve been worse. She could’ve won bronze.
- c. A: It’s too bad Emma only won silver.  
 B: True, but **at least** she won silver.

Taking evaluative *at least* to be oriented “downward” in this way by marking lower alternatives as worse may also provide a different perspective on unifying its meaning with epistemic *at least*,

namely by taking the epistemic interpretation to convey that all lower alternatives are false.<sup>16</sup> The two interpretations would thus only differ on the dimension with respect to which alternatives are assessed, either in terms of truth or in terms of desirability, respectively, similarly to what is at the heart of the contextual sensitivity of the ranking endorsed by Biezma (2013) and Chen (2018).

An independent motivation for this perspective comes from other Focus-particles exhibiting a similar ambiguity such as *only*. While *only* is most commonly featured with respect to its exhaustive inference, it has also been discussed as contributing a scalar (or evaluative) reading (Beaver & Clark 2008; Alxatib 2017). Interestingly, the difference between these two interpretations can be described analogously to *at least* as varying minimally in the relevant ranking – truth versus desirability – but with an upward orientation: Exhaustive *only* conveys that all higher alternatives are false whereas scalar *only* marks all higher alternatives to be more desirable. As an additional similarity, scalar *only* has been noted to arise when the exhaustive inference would be vacuous due to the prejacent already ruling out higher mutually exclusive alternatives (Klinedinst 2004), similar to how evaluative *at least* has been taken to arise when higher alternatives are false and epistemic *at least* would be redundant.<sup>17</sup>

A final relevant property that the two interpretations differ on that has not been much discussed and relates to the comparison with *only* as well is their accommodation behavior or Strong Contextual Felicity condition (Tonhauser et al. 2013). Biezma (2013) notes that evaluative *at least* is infelicitous out of the blue, which is in line with experimental results by Göbel (2022). Moreover, Göbel also shows that *only* leads to a penalty in ratings in contexts where the scalar component lacks contextual grounding. In contrast, the exhaustive inference of *only* is usually taken to be an at-issue component, suggesting that the contrast between truth and desirability correlates with how the relevant dimension is encoded, either as at-issue or a presupposition. While it goes beyond the scope of this paper to integrate these remarks into a full fledged account of *at least*, we believe that they point toward an interesting possibility that may help in understanding the syntax-semantics mapping investigated here.

---

<sup>16</sup> This view of epistemic *at least* would be more in line with an account like Cohen & Krifka (2014) that rules out speech acts endorsing a lower alternative, rather than accounts that focus on how to model what alternatives are left open (e.g. Büring 2008; Mendia 2022), although the latter could be rephrased accordingly given the logical equivalence. However, these accounts from the literature on numeral modifiers also rely on complex interactions with pragmatics to derive all relevant inferences associated with epistemic *at least*, which we do not have the space to go into here insofar as our results primarily require reconsidering prior views of evaluative *at least*.

<sup>17</sup> A gap in the potential connection between *at least* and *only* would be the ambiguity of *only* does not seem to be syntactically mediated, but see Winterstein & Davis (2022) for an analysis of *only* as a sentential connective.

## Supplementary files

Experimental files, results files and analysis files can be accessed through the OSF repository at <https://osf.io/rw9jn>.

## Ethics and consent

The experimental studies reported here were approved by the McGill Research Ethics Board under protocols #401-0409/#342-0118 to Michael Wagner.

## Funding information

This research was funded by a Feodor-Lynen Fellowship of the Humboldt-Foundation to the first author and an SERC Discovery Grant to the second author.

## Acknowledgements

We want to thank Luis Alonso-Ovalle, Chris Davis, Lisa Matthewson, Bernhard Schwarz, Duane Watson, Gregoire Winterstein, three anonymous reviewers, our editor Lyn Tieu, as well as the audience at AMLaP 2021 for feedback and comments on the project. All errors are our own.

## Competing interests

The authors have no competing interests to declare.

---

## References

- Alexandropoulou, Stavroula. 2018. *On the pragmatics of numeral modifiers: The availability and time course of variation, ignorance and indifference inferences*. Utrecht dissertation.
- Alxatib, Sam. 2017. The scalar presupposition of 'only' and 'only if'. In Cremers, Alexandre & van Gessel, Thom & Roelofsen, Floris (eds.), *Proceedings of Amsterdam Colloquium 21*. 96–105. Amsterdam: ILLC.
- Barr, Dale J. Barr & Levy, Roger & Scheepers, Christoph & Tily, Harry J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278. DOI: <https://doi.org/10.1016/j.jml.2012.11.001>
- Beaver, David & Clark, Brady. 2008. *Sense and sensitivity: How focus determines meaning*. Oxford: Blackwell. DOI: <https://doi.org/10.1002/9781444304176>
- Biezma, María. 2013. Only one 'at least': Refining the role of discourse in building alternatives. In Shwayder, Kobey (ed.), *Proceedings of Penn Linguistics Colloquium 36*. 11–19.
- Büring, Daniel. 2008. The least 'at least' can do. In Chang, C. B. & Haynie, H. J. (eds.), *Proceedings of West Coast Conference in Formal Linguistics 26*. 114–120. Cascadilla Proceedings Projects.

- Chen, Yi-Hsun. 2018. *Superlative modifiers: Ignorance and concession*. Rutgers University dissertation.
- Christensen, Rune Haubo B. 2019. Cumulative link models for ordinal regression with the R package ordinal. [https://rdrr.io/cran/ordinal/f/inst/doc/clm\\_article.pdf](https://rdrr.io/cran/ordinal/f/inst/doc/clm_article.pdf).
- Cohen, Ariel & Krifka, Manfred. 2014. Superlative quantifiers and meta-speech acts. *Linguistics and Philosophy* 37. 41–90. DOI: <https://doi.org/10.1007/s10988-014-9144-x>
- Erlewine, Michael Yoshitaka. 2014. *Movement out of focus*. Massachusetts Institute of Technology dissertation.
- Geurts, Bart & Nouwen, Rick. 2007. ‘At least’ et al.: the semantics of scalar modifiers. *Language* 83. 533–559. DOI: <https://doi.org/10.1353/lan.2007.0115>
- Göbel, Alexander. 2022. On the role of focus-sensitivity for a typology of presupposition triggers. *Journal of Semantics* 39. 617–656. DOI: <https://doi.org/10.1093/jos/ffac011>
- Grosz, Patrick Georg. 2011. A uniform analysis for concessive ‘at least’ and optative ‘at least’. In Ashton, Neil & Chereches, Anca & Lutz, David (eds.), *Proceedings of Semantics and Linguistic Theory* 21. 572–591. DOI: <https://doi.org/10.3765/salt.v21i0.2627>
- Kay, Paul. 1992. At least. In Lehrer, A. & Kittay, E. F. (eds.), *Frames, fields, and contrasts: new essays in semantic and lexical organization*, 309–332. Hillsdale, NJ: L. Erlbaum Associates.
- Klinedinst, Nathan. 2004. Only scalar *only*. Talk presented at Presupposition & Implicature Workshop. [http://www.linguist.univ-paris-diderot.fr/~amsili/Rech/jpi04/nklinedinst\\_only\\_paris.pdf](http://www.linguist.univ-paris-diderot.fr/~amsili/Rech/jpi04/nklinedinst_only_paris.pdf).
- Krifka, Manfred. 1999. At least some determiners aren’t determiners. In Turner, K. (ed.), *The semantics/pragmatics interface from different points of view*, vol. 1. 257–291. Oxford: Elsevier.
- Lenth, Russell V. 2016. Least-squares means: The R package lsmeans. *Journal of Statistical Software* 69(1). 1–33. DOI: <https://doi.org/10.18637/jss.v069.i01>
- Mendia, Jon Ander. 2022. Structural effects on implicature calculation. *Journal of Semantics*, 1–34. DOI: <https://doi.org/10.1093/jos/ffac004>
- Nakanishi, Kimiko & Rullmann, Hotze. 2009. Epistemic and concessive interpretation of *at least*. Talk presented at Canadian Linguistics Association. [https://linguistics.sites.olt.ubc.ca/files/2018/03/2009.Nakanishi\\_Rullmann.CLA\\_-1.pdf](https://linguistics.sites.olt.ubc.ca/files/2018/03/2009.Nakanishi_Rullmann.CLA_-1.pdf).
- Nouwen, R. 2015. Modified numerals: the epistemic effect. In Alonso-Ovalle, L. & Menéndez-Benito, P. (eds.), *Epistemic indefinites*, 244–266. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199665297.003.0011>
- R Core Team. 2018. R: A language and environment for statistical computing. <https://www.R-project.org/>.
- Roberts, Craige. 2012. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5. 1–69. Earlier version appeared in OSU Working Papers in Linguistics 49 in 1996. DOI: <https://doi.org/10.3765/sp.5.6>
- Schwarz, Bernhard. 2016. Consistency preservation in quantity implicature: The case of ‘at least’. *Semantics and Pragmatics* 9. 1–47. DOI: <https://doi.org/10.3765/sp.9.1>



Tonhauser, Judith & Beaver, David & Roberts, Craige & Simons, Mandy. 2013. Toward a taxonomy of projective content. *Language* 89. 66–109. DOI: <https://doi.org/10.1353/lan.2013.0001>

Winterstein, Grégoire & Davis, Christopher. 2022. From exclusive particles to adversative connectives. In Pratley, Breanna & Bakay, Özge & Neu, Eva & Deal, Peyton (eds.), *Proceedings of North East Linguistic Society 52*. Amherst, MA: GLSA. <https://semanticsarchive.net/Archive/zc5ZDM5Z/>.

