

Scalar diversity and ignorance inferences: An experimental study on *at least* as a modifier of numerals vs. adjectives*

Stavroula Alexandropoulou
University of Potsdam

Abstract This work presents results from an experiment that investigates whether *at least* as a modifier of gradable adjectives (e.g., *at least misleading*) triggers speaker ignorance inferences just as has been established for *at least* as a numeral modifier (e.g., *at least two*). I find that, while *at least* gives rise to ignorance inferences with both types of scalar expressions, this happens in varying degrees, contra existing accounts of *at least* (Geurts & Nouwen 2007; Cohen & Krifka 2014) and in line with experimental evidence on the scalar inferences of unmodified adjectives and numerals (Doran, Baker, McNabb, Larson & Ward 2009), known as scalar diversity. I also find indications that the scale structure of adjectives may affect the availability of ignorance inferences, as in the case of scalar implicature computation for unmodified adjectives (Gotzner, Solt & Benz 2018a), yet in a reverse manner.

Keywords: ignorance inferences, modified numerals, gradable adjectives, scalar diversity, scale structure, vagueness, scalar implicatures.

1 Introduction

It is well established that utterances with superlative numeral modifiers like *at least* convey a robust signal of speaker ignorance (SpI), as first observed by Geurts & Nouwen (2007). Thus, (1) comes across as a weird utterance (note #) exactly because it strongly implies that the speaker does not know how many kids they have.

- (1) # I have at least two kids.
 ↗ ‘Speaker doesn’t know how many kids they have’ (*speaker ignorance*)¹

* The research leading to these results has received funding from the German Research Foundation (DFG) as part of the Emmy Noether programme, Emmy Noether Grant Nr. GO 3378/1-1. I would like to thank Nicole Gotzner, Rick Nouwen, Elli Tourtouri, Myrto Pantazi, Radim Lacina, Marisha Herb, Henrik Discher, and the audiences of CUNY 2021, SALT31, AMGL41, and of the SynSem colloquium in Potsdam for useful discussion and helpful feedback. I would also like to thank all my participants and contributors who forwarded my questionnaire for volunteering their precious time.

¹ Other, secondary, reasons why one could utter (1) can be that the speaker is knowledgeable, but not cooperative or willing to disclose further information to the interlocutor(s) (see, e.g., Fox 2014 on such

There has been an extensive theoretical investigation of *at least* as a numeral modifier (*at least n*) and the SpI inference it triggers (Geurts & Nouwen 2007; Cummins & Katsos 2010; Kennedy 2015; Nouwen 2010; Coppock & Brochhagen 2013a; Cohen & Krifka 2014; Schwarz 2016; Buccola & Haida 2020; Cremers, Coppock, Dotlačil & Roelofsen 2021; a.m.o.). Nouwen (2010) provides cross-linguistic evidence of the robustness with which this inference is triggered by superlative numeral modifiers and their counterparts in a large number of languages (so-called Class B modifiers in Nouwen 2010). The majority of the theoretical proposals derive SpI of *at least n* as a pragmatic inference (e.g., Cummins & Katsos 2010; Coppock & Brochhagen 2013a; Cohen & Krifka 2014; Kennedy 2015; Schwarz 2016), and modeling it as a primary Quantity implicature (à la Sauerland 2004) via a standard Quantity-based reasoning about the relevant stronger alternatives has been a popular take on SpI. There is also a considerable amount of experimental research on SpI inferences of *at least n* confirming their pragmatic status (Geurts, Katsos, Cummins, Moons & Noordman 2010; Cummins & Katsos 2010; Coppock & Brochhagen 2013b; Westera & Brasoveanu 2014; McNabb & Penka 2015; Alexandropoulou 2018; Nouwen, Alexandropoulou & McNabb 2019).

On the other hand, uses of *at least* as an adjective modifier (*at least+adj*), see (2a), have generally been overlooked by the literature on superlative modifiers (but see Geurts & Nouwen 2007; Cohen & Krifka 2014) and it is unclear to what extent *at least* triggers equally robust SpI inferences across different scalar constructions/categories, e.g., *at least+n* vs. *at least+adj*. Interestingly enough, cancellation data suggest that SpI inferences are less robust and less likely to arise with *at least+adj* than with *at least+n*. While the weirdness of (1) (note #) is a sign of the uncontroversial fact that SpI inferences of *at least+n* arise robustly and are hard to be suspended, the *in fact* sentence in (2b) seems to easily cancel the SpI inference about the stronger alternative *complete*, i.e., that the speaker doesn't know whether the issue is treated with complete intellectual honesty, assuming, e.g., the informativity (Horn) scale <rudimentary, complete>.

- (2) a. ... the issue is treated with at least rudimentary intellectual honesty ...²
(NPR Morning, 15/01/2002)
- b. In fact, it is treated with complete intellectual honesty.

The above contrast in ease of SpI cancellation, thus, points to a difference in

uses of disjunction), or that the number being modified is a number/threshold salient in the context and all that matters is whether the speaker has 2 or more, or fewer than 2 children in this particular case. The latter is the so-called *speaker indifference* inference (Alexandropoulou 2018), and is compatible with speaker's knowledgeability. For a difference between uncooperativity/unwillingness to inform and speaker indifference, see the discussion in Alexandropoulou 2018, section 7.1.2.

² Adapted from Corpus of Contemporary American English (COCA, Davies 2008-).

the strength and likelihood with which SpI inferences of *at least+n* vs. *at least+adj* are derived. This is reminiscent of the variability observed among different types of (unmodified) scalar expressions as to the likelihood of triggering secondary Quantity implicatures, aka scalar implicatures (Doran et al. 2009; van Tiel, van Miltenburg, Zevakhina & Geurts 2016; Gotzner et al. 2018a; Gotzner, Solt & Benz 2018b). This phenomenon has been termed *scalar diversity* (van Tiel et al. 2016) and has mainly been discussed in relation to scalar implicatures. In the present work, I aim to investigate to what extent *at least* triggers SpI inferences uniformly across different scalar constructions or, put differently, whether SpI inferences, too, are subject to scalar diversity. To that end, I will probe experimentally SpI inferences of *at least* as a numeral modifier as compared to *at least* as an adjective modifier.

In the following section (Section 2), I briefly describe the theoretical landscape of the superlative modifier *at least*, and also review existing studies on various pragmatic inferences of scalar expressions with implications for the phenomenon of scalar diversity. In Section 3, I present an experiment on SpI inferences of *at least+n* vs. *at least+adj*. Section 4 discusses possible interpretations of the obtained results and Section 5 concludes the paper.

2 Theoretical background

2.1 Accounts of *at least*

There are various views on superlative numeral modifiers like *at least* and the SpI they convey. The dominant pragmatic view includes Quantity-based accounts that derive SpI inferences as primary Quantity implicatures (à la Sauerland, 2004) via a standard Quantity-based reasoning (Büring 2008; Cummins & Katsos 2010; Kennedy 2015; Schwarz 2016). Nouwen's (2015) and Alexandropoulou's (2018) hybrid proposals derive SpI via a combination of presupposition-based alternatives fed into a Manner-based pragmatic mechanism and of Horn scale alternatives used in a standard Quantity-based pragmatic reasoning. Another hybrid account, coached in the framework of speech acts, comes from Cohen & Krifka (2014), whereby SpI inferences come about via a Quantity-based reasoning and are obligatorily derived in order to obtain the truth-conditions of a sentence with *at least n*.

There's also a view featuring a more conventional encoding of the SpI inference of *at least n*, or even a semantic/lexical encoding deriving SpI as an entailment, thereby capturing its persistence and robustness. Geurts & Nouwen (2007) capture SpI by positing a lexical semantics for *at least* that contains an epistemic modal component. Nouwen (2010) models SpI via the application of a silent LF existential modal operator required to license superlative numeral modifiers.³ The import of

³ The import of this operator into the LF of a sentence with a superlative numeral modifier results in

this operator is pragmatically motivated but the resulting speaker ignorance is an obligatory inference derived via compositional semantics.⁴ Coppock & Brochhagen (2013a), Ciardelli, Coppock & Roelofsen (2018), and Spsychalska (2015) derive SpI via observing a conversational maxim tied to the semantics of superlative numeral modifiers, which resembles the Gricean maxim of Quality.⁵ Spector (2015), on the other hand, derives SpI effects of *at least* in grammar as Quantity implicatures and captures their robustness via the application of an obligatory exhaustification operation above the speaker belief operator that propositions are prefixed with (as in Meyer 2013). Buccola & Haida (2020) offer a similar account, where ignorance effects are derived in grammar as entailments across the board and are additionally derived via a pragmatic route when there is a *how many* Question under Discussion (QuD; see also Cremers et al. 2021).

What most of the above accounts have in common is that they derive the following SpI implications for a sentence with *at least n* (see, e.g., (1)): the speaker doesn't know whether *exactly n* is the case and she doesn't know whether *more than n* is the case (but see Cohen & Krifka 2014 and Coppock & Brochhagen 2013a). Applying the standard Quantity-based recipe, these follow from the primary Quantity inferences together with the possibility implications that the speaker considers it possible that *exactly n* is the case and she considers it possible that *more than n* is the case. The latter implications, called *specific ignorance* by Alexandropoulou (2018), are entailed by the Quality assumption that the speaker believes that *at least n* is true along with the primary inferences. There is also experimental evidence provided by Mendia (2016) that the SpI inferences of *at least n* must include the exhaustive interpretation of the prejacent (i.e., 'exactly *n*') as an epistemic possibility for the speaker, namely the first half part of specific ignorance (see also Alexandropoulou, Dotlačil, McNabb & Nouwen 2015 and Alexandropoulou 2018 for a similar claim as to variation inferences of embedded *at least n*). Further support of *at least n* triggering a specific SpI implicature comes from Alexandropoulou's (2018) eye-tracking findings.

Of the aforesaid accounts, Geurts & Nouwen (2007) and Cohen & Krifka (2014) extend their account of superlative modifiers to also capture uses with non-numerical scales, such as adjectival scales. Although they do not go into great detail considering, for instance, how the representation of adjectival scales or properties of the measurement scales underlying the meaning of (gradable) adjectives may affect SpI interpretations, as compared to uses with numerical scales, one can extract the following prediction from these two analyses: *At least* triggers SpI interpretations

creating an epistemic range of values with the numeral modified being an endpoint on that range.

⁴ In a sense, this account could also be classified as a hybrid account.

⁵ The Quality maxim is argued to be inescapable and hard to violate compared to other maxims, such as Quantity or Manner (Grice 1989).

uniformly across numerical and non-numerical scales.

Note, finally, that a salient interpretation of adjectival uses of *at least* (often with adjectives of a positive valence as in (3a)) is the so-called concessive interpretation or ‘settling for less’ interpretation according to Nakanishi & Rullmann (2009). This is the following interpretation for (3a), when prominent phonological focus falls on the verb: ‘Getting good grades is considered to be a satisfactory outcome, though not the optimal one’. This is the same when *at least* takes a propositional argument as in (3b) (see relevant discussion in Geurts & Nouwen 2007).

- (3) a. Her grades were at least good.
- b. At least, her grades were good.

A concessive interpretation of a sentence with *at least+adj* like (3a) may obscure the availability of a SpI interpretation (e.g., ‘the speaker doesn’t know whether her grades were just good and doesn’t know whether they were more than good’).

In the following, I review experimental findings on a variety of pragmatic interpretations triggered differentially by different types of scalar expressions, including numerical and adjectival scales.

2.2 Scalar diversity: Different scalar expressions trigger pragmatic inferences differentially

For the most part, scalar diversity has been discussed in connection with scalar implicatures and refers to the observation that different scalar expressions, such as numerals, quantifiers, adjectives, verbs, modals, connectives, trigger scalar implicatures to varying degrees (Doran et al. 2009; Doran, Ward, Larson, McNabb & Baker 2012; McNabb 2015; van Tiel et al. 2016), with adjectives giving rise to particularly low rates of scalar implicatures (but see Gotzner et al. 2018a,b below). Scalar diversity is found across, but also within, categories of scalar expressions.

Focussing on adjectival scales here, those are argued to trigger scalar implicatures to a small extent, compared, for instance, to standard quantificational scales, because the scale-mates of the trigger may not be salient enough (Doran et al.’s (2009) salience explanation of scalar diversity) or distinguishable from each other (van Tiel et al.’s (2016) distinctness explanation of scalar diversity). In a similar vein, Sassoon & Zevakhina (2012) hint at a causal relation between bare adjectives being less likely to trigger upper-bounded interpretations and them being *interpreted along coarse-grained scales, with high approximation levels* (p. 237), and *overlapping intervals* (p. 244), unlike numerical scales that typically have salient and clearer reference points. Zeroing in even further on adjectival scales, Gotzner et al. (2018a,b) show that adjectives do not receive low rates across the board, as they find that absolute and relative gradable adjectives, differing in the scale structure they use (Rotstein &

Winter 2004; Kennedy 2007, a.o.), give rise to different rates of scalar implicature, with absolute adjectives being better triggers than relative ones. Gotzner et al. conclude that properties of the measurement scale of gradable adjectives, such as scale structure, account for a large amount of the variability within the category of adjectives (see also Leffel, Cremers, Gotzner & Romoli 2019 on implicatures of gradable adjectives in the *not very* ADJ construction). In this context, it would not be surprising to find differences between numerals and adjectives modified by *at least*, by virtue of differences pertaining to the respective scales they are associated with, semantically or lexically (cf. measurement vs. Horn scales; see Solt 2015 and Gotzner & Kiziltan forthcoming for a discussion of different notions of scale).

Having said that, the debate on scalar diversity has drawn way less attention to other inferences triggered by diverse scalar expressions. An exception is the study by Dieuleveut, Chemla & Spector (2019), who investigated the likelihood of different scalar expressions triggering primary Quantity implicatures (i.e., ‘it is not the case that the speaker believes that [A STRONGER ALTERNATIVE] holds’). They found no evidence of such interpretations for bare numerals or for plural morphology on nouns, as opposed to the scalar items *some* and *almost*. Mendia (2016), too, tested ignorance inferences triggered by superlative quantifiers when modifying totally ordered scales (numerals) vs. partially ordered scales (conjunctions). He found that in the former case, the exhaustive interpretation of the prejacent (i.e., ‘exactly *n*’) is an epistemic possibility for the speaker (see also previous section), which need not be the case for the latter scale type. In addition, in previous work (Alexandropoulou 2018), I studied different types of pragmatic inference triggered within the category of modified numerals (see also Alexandropoulou et al. 2015; McNabb, Alexandropoulou, Blok, Bimpikou & Nouwen 2016; Alexandropoulou, Dotlačil & Nouwen 2017; Nouwen et al. 2019). I found that class A modified numerals (e.g., *more than n*) trigger variation/free-choice-like effects and speaker ignorance inferences to a lesser extent compared to truth-conditionally equivalent class B modified numerals (e.g., *at least n+1*; see Nouwen 2010 for the class A/B distinction). I attributed this finding to a difference in the source of (part of) the alternatives associated with the two classes, which are further fed into a single implicature mechanism shared by the two inferences in question.

It is apparent that scalar diversity is relevant for and exhibited in other domains of pragmatics as well. The present study seeks to investigate SpI effects induced by the superlative quantifier *at least* across categories, looking into its use as a numeral modifier vs. as an adjective modifier. Doing so will allow us to contribute novel data and possible new factors to the unresolved debate on scalar diversity.

3 Current study

3.1 Research questions & design

In this study, I set out to answer the following research question: To what extent speaker ignorance inferences arise uniformly for different scalar expressions modified by *at least*. Specifically, I aim to systematically test for uses of *at least* as a numeral modifier and of *at least* as an adjective modifier.

This research was carried out in Greek in order to avoid concessive interpretations of English *at least* (see Section 2.1) that may interfere with the inferences I am after. Greek has two counterparts of *at least*, *tulahiston* and *to lighotero*, and only the former can have a concessive function (Barouni 2018). Barouni further argues that *to lighotero* gives rise to stronger SpI effects compared to *tulahiston*.

In particular, I compared the lower-bound quantifier *to lighotero* as a numeral modifier (***at least+num*** condition) with uses of *to lighotero* as a modifier of a gradable adjective (***at least+adj*** condition). This was one manipulation (Scalar Expression). In order to test for the availability of ignorance/the likelihood of SpI to arise, I compared *at least* to a condition with an exclusive adverb modifying the scalar expressions, which was the control condition (***just*** condition). This was the second manipulation (Scalar Modifier), which also included a condition where numerals and gradable adjectives appeared unmodified/bare (so-called **null** condition). So we had a 2 Scalar Expression \times 3 Scalar Modifier design.

3.2 Methods

The study used an inference judgement task where participants are asked to provide their judgements on a 7-point Likert scale. The questionnaire was administered online and was created on Google forms.

3.2.1 Participants

54 native speakers of Greek (36 women, 17 men, 1 diverse, mean age: 38.6, age range: 20-72) were recruited. Data of 2 participants were discarded because they were not completely naive as to the purpose of the study. The experiment lasted for about 8 to 10 minutes. Participation was voluntary and without compensation.

3.2.2 Materials & procedure

The stimuli consisted of pairs of sentences: An utterance by the main character, Maria, followed by a conclusion. Maria's utterance contained both manipulations (Scalar Modifier and Scalar Expression), while the conclusion was always an ignorance statement that Maria doesn't know whether a stronger relevant alternative

from the scale at stake holds. Participants had to judge to what extent the conclusion is valid given Maria's utterance. This setup was inspired by Westera & Brasoveanu's (2014) study on the context/QuD sensitivity of speaker ignorance inferences of the truth-conditionally equivalent quantifiers *at most n* and *less than n+1* (assuming discrete numbers for *n*). (4) and (5) present experimental items in all 6 conditions, which are approximately translated from Greek (see corresponding original items in a repository on OSF).⁶

Numeral condition

(4) Maria says:

There were $\left\{ \begin{array}{c} \textit{at least} \\ \textit{just} \\ \emptyset \end{array} \right\}$ *thirteen* actors on stage during the performance we saw yesterday.

Conclusion: Maria doesn't know **the exact number** of actors that were on stage during the performance she saw yesterday.

Adjective condition

(5) Maria says:

Kostas' overall performance at school is $\left\{ \begin{array}{c} \textit{at least} \\ \textit{just} \\ \emptyset \end{array} \right\}$ *good*.

Conclusion: Maria doesn't know whether Kostas' overall performance at school is **excellent**.

How valid is the conclusion given Maria's utterance?
(not at all valid) 1 2 3 4 5 6 7 (absolutely valid)

As already said, the target sentence in Maria's original utterance in Greek contained the quantifier *to lighotero n* ('at least *n*', lit. the less *n*) in the *at least* condition, whereas the *just* control items used different exclusive modifiers across Scalar Modifier conditions: In the **numeral (num)** condition the adverb used was *akrivos* 'exactly', while in the **adjective (adj)** condition the adverb *aplos* 'simply' was used. This was a necessary compromise, as the Greek equivalents of *only* or

⁶ <https://mfr.osf.io/render?url=https://osf.io/749gn/?direct%26mode=render%26action=download%26mode=render>)

just, which can modify both numerals and adjectives in English, turned out not to work equally well with numerals and adjectives, given my informants' feedback.⁷ That being said, the two exclusive elements used have a similar function: They exhaust the relevant set of alternatives evoked by the expression they modify, which amounts to the negation of all (stronger/more informative) relevant alternatives and an upper-bounded interpretation of the modified expression (see Beltrama 2021 and references therein on *simply*; see Gotzner & Benz 2021 on *only n* and *exactly n* and Sauerland's (2013) analysis of *exactly* as a focus sensitive element similar to *only*).

The scalar expression in the target sentence was always a non-round number in the **numeral** condition, as round numbers may be taken to already signal speaker ignorance on their own, given that they are frequent in imprecise/approximate contexts or have a coarse/approximate interpretation (see Krifka 2009 and references therein; but also Alexandropoulou, Dotlačil & Nouwen 2016; Alexandropoulou 2018). This was done in order to keep the differences between **numeral** and **adjective** conditions, being compared in terms of speaker ignorance interpretations, to a minimum.

In the **adjective** condition, two types of gradable adjectives were included as scalar expressions in the target sentence: Relative and absolute adjectives (see, e.g., Kennedy & McNally 2005; Kennedy 2007). The two types of gradable adjectives differ in the scale structure they use: Relative adjectives use a totally open scale, whereas the scale used by absolute adjectives has an endpoint (Rotstein & Winter 2004; Kennedy & McNally 2005). Gradable adjectives that use scales without an endpoint, i.e., relative adjectives, are characterized by vagueness (Kennedy 2007), which is an important difference of the two types of gradable predicates. Namely, there is uncertainty about whether a relative predicate applies to an individual, e.g., what heights count as tall or what the threshold/standard of tallness is, while this is not the case for absolute adjectives, which have fixed and context-invariant standards: For instance, an individual entity qualifies as dirty as long as it possesses a non-zero/minimal degree of dirt (minimal degree on scale of dirt). Absolute adjectives are further split into minimum standard and maximum standard adjectives. *Dirty* exemplifies the former subclass, while the latter requires that an individual possesses the maximal degree of the property at stake, which corresponds to the maximal degree of the relevant scale (e.g., *clean*, *dry*). The adjective scales used in the experiment were borrowed from Gotzner et al. 2018a and were translated accordingly into Greek (they can be accessed [here](#)). Note that for absolute adjectives we used both minimum standard ($n = 4$) and maximum standard ($n = 2$) adjectives as scalar expressions in the target sentence in the **adjective** condition. As in Gotzner et al. 2018a, the scales of maximum standard adjectives had a weaker maximum

⁷ The Greek counterparts we considered were *mono*, *monaha*, and *molis*.

standard adjective, while the respective stronger scale-mates denoted *that standard interpreted at a higher level of precision* (p. 9; e.g., <clean, spotless>).

As specified in Section 2.2, according to the prevalent analysis of SpI effects of *at least n* and experimental evidence by Mendia (2016) and Alexandropoulou (2018), the ignorance inference of the target sentence in the **numeral** condition is that Maria does not know whether there were just/exactly 13 actors and she doesn't know whether there were more than 13 actors on stage in (4). So the rationale is the following: This composite inference follows from a reasoning (above a belief operator) assuming, e.g., a two-scale analysis of *at least+n*. Namely, with substitutions of *at least* from the scale <at least, just/only/exactly> and of *n* from the lexical number scale (see, e.g., Schwarz 2016, a.o.). Likewise for (5), with the difference of substituting for the adjective from the Horn scale <good, excellent>. Hence, the composite ignorance inference of the target sentence in the **adjective** condition is that Maria does not know whether Kostas' overall performance was just good and she doesn't know whether it was more than good. In both Scalar Expression conditions, the conclusion sentence is compatible with the aforementioned composite ignorance inference of Maria's utterance when it contains *at least*. Thus, if participants calculate the composite SpI implicature in Maria's utterance with *at least* (**at least** items), they will judge the conclusion as valid, giving it higher ratings on the 7-point Likert scale. Lower ratings on that scale would then indicate that the conclusion stating Maria's ignorance about relevant stronger alternatives for a given scale is not justified given Maria's utterance, hence, that Maria's utterance does not convey any such information. In our control condition with *just* (**just** items), Maria's utterance is evidently incompatible with the conclusion of Maria's ignorance about whether a relevant stronger alternative holds, as Maria's utterance conveys the falsity of all relevant alternatives. So participants will pick ratings from the bottom of the scale in **just** items.

I constructed 12 items where the scalar expression was a numeral and 12 where it was a gradable adjective, half of which were relative adjectives (**RelAdj**) and the other half were absolute adjectives (**AbsAdj**). I also included 24 filler items, 12 designed to receive low ratings on the validity scale (bad fillers) and 12 intended to receive high ratings on the scale (good fillers). In the bad fillers, the conclusion was unambiguously invalid or contradictory given Maria's utterance, see (6) in Appendix, while in the good ones it was a valid conclusion given Maria's utterance, as in (7) (Appendix). As is obvious, the conclusion of the filler items, too, was stating Maria's ignorance ("Maria doesn't know..."), but did not target information similar to that targeted in the test items.

Finally, the experiment included two practice trials, which exemplified the task and also served to familiarize participants with the task. Practice trials were similar to the filler items: One involved a conclusion that was (stated to be) invalid given

Maria's utterance and instructed participants to pick a rating from the left end of the scale, while the other one involved a conclusion that was valid given Maria's utterance and encouraged participants to pick a score from the right end of the scale.

12 items appeared in six conditions (3 Scalar Modifier \times 2 Scalar Expression) and were rotated through 6 pseudo-randomized lists in a Latin square design. Participants received 38 trials in total, including practice items, and were randomly assigned to a list. All materials were created on the basis of or adapted from naturally occurring sentences in the Hellenic National Corpus (HNC)⁸ or Greek Google, and were further judged by three native speakers of Greek.

After giving their consent to participate in the study, participants had to answer some demographic questions. Next, they read the instructions and completed the practice phase. Then they were let to proceed to the questions of the actual experiment.

3.3 Predictions

The *at least+num* condition is predicted to present high validity rates of concluding SpI on the part of Maria, given the firm cross-linguistic evidence of the robustness of the SpI signal of *at least* and other class B expressions used as numeral modifiers (see Nouwen 2010). Also, Barouni (2018) claims that *to lighotero behaves like a typical class B element giving rise to strong ignorance effects* (p. 200).

Geurts & Nouwen (2007) and Cohen & Krifka (2014), which derive SpI inferences in a similar manner for *at least+n* and *at least+adj* and to the same extent, predict that the respective conditions will exhibit a similar pattern of responses.

As to the **null** condition, if the Competence Assumption (van Rooij & Schulz 2004; Geurts 2011; or the epistemic step à la Sauerland 2004) is met that the speaker is knowledgeable about the stronger alternatives, upper-bounded interpretations should arise⁹, reflected by low validity ratings, in a similar fashion to the *just* control condition. However, they should arise differentially for bare numerals and bare adjectives, given Doran et al.'s (2009) findings that bare numerals are more

⁸ Institute for Language and Speech Processing - Athena Research Center (2015)

⁹ According to the Grammatical approach to scalar implicatures (e.g., Chierchia, Fox & Spector 2012), the derivation of a scalar implicature associated with a bare scalar expression (secondary Quantity implicature) does not rely on the Competence Assumption and is independent of the derivation of primary Quantity implicatures (ignorance implicature). Given that, one may assume that participants in our experiment may think that Maria is not fully informed/competent about the truth of the stronger alternative in the conclusion, but still derive a scalar implicature in Maria's utterance in the **null** conditions. Hence, the conclusion about Maria's ignorance may be justified even if participants have derived a scalar interpretation in Maria's utterance. That is, higher ratings in the **null** condition are not necessarily an indication of the derivation of an ignorance interpretation in Maria's utterance, and can be given by participants even if they have derived a scalar implicature in Maria's utterance.

likely to trigger upper-bounded interpretations compared to bare adjectives. That is, **null+num** items should receive lower rates of concluding SpI than **null+adj** items, as compared to the respective control *just* conditions. Furthermore, considering [Götzner et al.'s \(2018a\)](#) findings that scale structure of gradable adjectives affects implicature derivation and that bare relative adjectives are worse at triggering upper-bounded interpretations compared to bare absolute adjectives, bare absolute adjectives are expected to receive lower validity rates than bare relative adjectives, as compared to the respective *just* control conditions.

3.4 Results & discussion

The data of 6 participants were removed on the basis of their mean responses to the bad and good filler items. More precisely, if a participant gave ratings greater than 3 to the bad fillers or smaller than 5 to the good fillers on average, the whole set of data of this participant was excluded from further analyses. Figure 1 shows the data of the remaining 46 participants, used in the statistical analyses reported on here.

Overall, we see that our bad and good fillers were rated as expected, with the former (**bad-fil**) receiving scores from the bottom of the scale (mean=1.27) and the latter (**good-fil**) mostly receiving high scores (mean=6.02). *At least+num* items, too, were rated rather high (mean=5.73), confirming the robustness of SpI inferences of *at least* as a numeral modifier in Greek as well. Remarkably, Figure 1 reveals diverse patterns of responses for *at least* as a numeral modifier vs. *at least* as an adjective modifier. Likewise, the **null** conditions, where numerals and adjectives appear unmodified, present a clear contrast, when compared to the respective *just* control conditions (albeit quite different from each other).

The data were analyzed using R (version 4.0.5). Participants' responses were ordered categorical, and to analyze them cumulative link mixed effects models were fit using the ordinal package ([Christensen 2019](#)) in R. I included the factors Scalar Modifier and Scalar Expression, along with their interaction as predictors. I fit a model with treatment-coded fixed effects (baselines: *just* for Scalar Modifier and **adj** for Scalar Expression) as well as the maximal converging random-effect structure. Table 1 summarizes the output of this analysis.

At least+adj items were rated significantly higher than *just+adj* ($p < .05$), showing that *at least+adj* does trigger SpI effects, though to a lesser extent than *at least+num* does, as demonstrated by the significant interaction in the *just* vs. *at least* comparison with the **num** vs. **adj** items ($p < .0001$). This is consistent with our aforementioned observation that *at least* shows different patterns as a numeral modifier vs. as an adjective modifier in terms of SpI effects. The analysis also revealed a simple effect of **null** such that **null+adj** items received significantly higher rates than *just+adj* ($p < .05$). This difference suggests that the target sentences

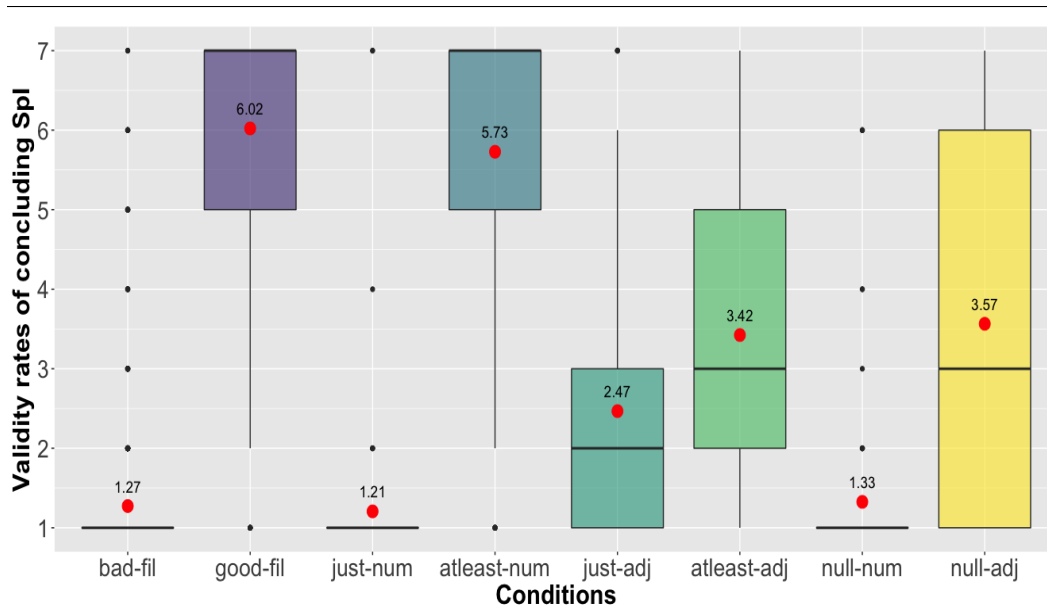


Figure 1 Box plots of validity ratings per condition. Red dots represent mean scores.

with bare adjectives in Maria's utterance (**null+adj** items) were not wholly assigned an upper-bounded interpretation (i.e., 'just ADJ'; but see footnote 9). Lastly, the significant interaction in the *just* vs. **null** comparison with the **num** vs. **adj** items ($p < .05$) manifests that **null+adj** items are less likely to receive an upper-bounded interpretation compared to **null+num** items, confirming our relevant observation.

Going back to the predictions (Section 3.3), the high rates of SpI interpretations detected for the Greek numeral modifier *to lighotero* ('at least') are in line with Barouni's (2018) relevant claim and add to the cross-linguistic profile of *at least* as a trigger of robust SpI effects. We further found far less robust SpI effects triggered by *to lighotero* ('at least') as an adjective modifier, which is unexpected given the proposals by Geurts & Nouwen (2007) and Cohen & Krifka (2014) in their existing form and brings forward an additional domain of pragmatics where scalar diversity seems to be operative. Finally, the interaction of the *just* vs. **null** contrast with the Scalar Expression factor is taken to be in line with Doran et al.'s (2009) findings that bare numerals are better at triggering upper-bounded interpretations than bare adjectives (but see footnote 9).

In what follows, I zoom in on the two gradable adjective classes included in the **adjective** items, namely, relative and absolute adjectives. Below I report on the results of a post-hoc analysis (baselines: *just* for Scalar Modifier and **RelAdj** for

	Estimate	SE	z-value	p-value
ScalarExpressionNum	-8.9537	2.1238	-4.216	2.49e-05 ***
ScalarModifierAtleast	0.5891	0.2878	2.047	0.0407 *
ScalarModifierNull	0.7007	0.3290	2.130	0.0332 *
ScalarExpressionNum:ScalarModifierAtleast	11.0838	2.2914	4.837	1.32e-06 ***
ScalarExpressionNum:ScalarModifierNull	5.7371	2.3628	2.428	0.0152 *

Table 1 Output of cumulative link model with treatment coding.

Scalar Expression) similar to the main analysis, setting the two gradable adjective classes as levels in place of **adjective** for the Scalar Expression factor. As this analysis is run on fewer data, possibly with lower power, the relevant findings are to be interpreted with caution. Figure 2 shows the validity ratings of concluding SpI per adjective class.

Interestingly, it is suggested by this analysis that the significant simple effect of *at least* in the main analysis (see Table 1) seems to be mostly driven by such an effect for relative adjectives ($\beta = .79$, $SE = .30$, $z = 2.58$, $p < .01$), as opposed to absolute adjectives (Tukey-adjusted $p = .34$). This indicates that a SpI interpretation becomes available with *at least* as a relative adjective modifier, but not with *at least* as an absolute adjective modifier. Hence, SpI inferences seem to target a specific class of gradable adjectives. A potential source of this preference is the scale structure gradable adjectives use, namely, whether it has an endpoint (absolute adjectives) or not (relative adjectives; see also Gotzner et al. 2018a,b on scalar implicatures and Leffel et al. 2019 on negative strengthening). Recall that adjectives that use scales without any endpoint (open scales) and, thus, the relevant threshold of application may be anywhere on the respective scale, lead to vague interpretations. This sensitivity of SpI inferences to scale structure, and vagueness in particular, appears to carry over to the **null** condition: **null+RelAdj** items were rated significantly higher than **just+RelAdj** items ($z = 3.73$, $p < .001$), while this effect was not significant for absolute adjectives (Tukey-adjusted $p = .76$). This is in line with Gotzner et al.'s (2018a) finding that scale structure affects implicature calculation as far as unmodified gradable adjectives are concerned, however the predicted interaction in **RelAdj** vs. **AbsAdj** with *just* vs. **null** (see Section 3.3) was not found reliable by the present analysis ($p = .18$).

To sum up the findings, we found evidence that SpI inferences do arise with the Greek counterpart of *at least* (*to lighotero*) when used as an adjective modifier, though to a lesser extent than when used as a numeral modifier. We also find indications that scale structure and/or vagueness may affect the availability of SpI inferences in the former construction, as in the case of scalar implicature computation with bare adjectives (Gotzner et al. 2018a,b). In the next section, I elaborate on these

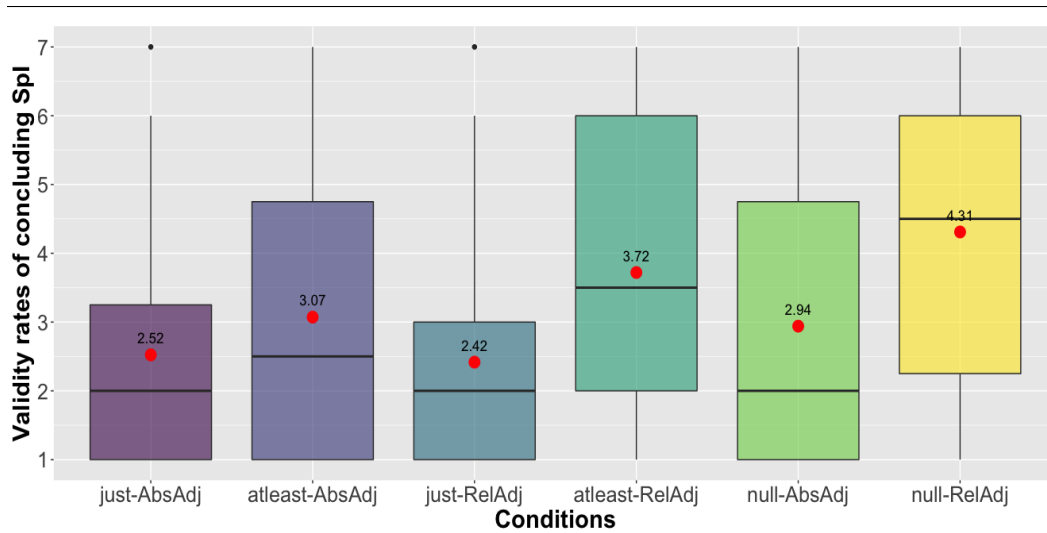


Figure 2 Box plots of validity ratings per gradable adjective condition. Red dots represent mean scores.

findings as well as on certain contested issues of the present study.

4 General discussion

4.1 Why is *at least*+numeral a better trigger of SpI than *at least*+adjective?

The fact that SpI inferences do not arise uniformly across uses of *at least* as a numeral modifier and as an adjective modifier suggests that the phenomenon of scalar diversity is exhibited by SpI implicatures associated with *at least*, too. In order to account for this, it is argued that theories of superlative modifiers like *at least* that derive SpI across different scalar expressions in a similar manner (e.g., [Geurts & Nouwen 2007](#); [Cohen & Krifka 2014](#)) should incorporate certain characteristics that distinguish the different types of scalar expressions. Here, I focus on the types of scales tested in the present study.

I conjecture that the number of (pragmatic) scales involved in the implicature mechanism is a relevant difference. On many accounts of *at least*, specific SpI is derived on the basis of the interaction of two scales: I.e., the scale $\langle \text{at least, exactly} \rangle$ and the Horn number scale (e.g., [Nouwen 2015](#); [Schwarz 2016](#)). For adjectives, it may be the case that SpI is derived only on the basis of one scale, namely, the scale associated with *at least*. Note that a recent proposal by [Gotzner \(2021\)](#) posits that implicatures of adjectives are computed via a reasoning about

positions on the relevant underlying measurement scale rather than about possible lexically-determined alternatives (so-called *measurement mechanism* account). The difference in the number (and/or type) of scales used in the implicature mechanism could be behind the diverse rates of SpI inferences attested for the constructions *at least+n* and *at least+adj*.

Even if we typically assume lexical alternatives fed into the implicature mechanism of gradable adjectives, the representation of adjectival scales as compared to numerical scales constitutes another crucial difference. As has been advocated in the literature of scalar diversity (Doran et al. 2009; van Tiel et al. 2016; Gotzner et al. 2018a,b, a.o.), alternatives of adjectival scales are less salient, less readily available, and/or less distinguishable from each other compared to alternatives of quantificational or numerical scales (see also Section 2.2). This makes it harder to access adjectival alternatives, unlike numerical alternatives, hence, less likely to feed them into the implicature mechanism responsible for SpI calculation. This would again explain away the observed diverse rates of SpI interpretations for the constructions *at least+n* vs. *at least+adj*.

Another explanation, suggested to me by Rick Nouwen (p.c.), concerns an assumed intrinsic difference between the possibility of ignorance with judging quantities, as in the case of numerals, and with judging qualities, as in the case of adjectives. Although it is conceivable to have ignorance about the number of people one saw yesterday (see (4)), it is less conceivable to have ignorance about whether one's school performance is excellent or not (see (5)), while being at the same time in a position of judging their performance as good. What lack of information could then keep one from judging it as excellent? This contrast hints that judging qualities is more likely to involve speaker knowledgeability, though we could imagine a situation where the speaker does not have knowledge on one's grades in *all* school subjects.

Nevertheless, a possible alternative reason that could restrain a competent/knowledgeable speaker from committing to the truth of the stronger alternative in the adjective case (*at least+adj*) is politeness.¹⁰ That is, the less informative *at least+adj* expression that conveys a less negative evaluation is used by the speaker as a type of hedging in a bid to preserve the hearer's face to some extent, cf. *this is at least misleading, if not wrong* (from Cohen & Krifka 2014, p. 64). Gradable adjectives, unlike numbers, are characterized in terms of evaluative polarity, which in turn has been argued to affect their interpretation under negation (so-called polarity asymmetry of negative strengthening interpretations, see Bolinger 1972; Brown & Levinson 1987; Levinson 2000; Gotzner & Kiziltan forthcoming; Gotzner & Mazzearella 2021; but see also Mazzearella & Gotzner 2021). Further research should

¹⁰ Thanks to Muffy Siegel for bringing up this point.

look into the possibility of an effect of evaluative polarity on the interpretation of adjectives modified by *at least* and the hypothesis that it is more likely to use *at least* with an evaluatively negative (weak) adjective thereby avoiding a stronger, direct face-threatening negative evaluation (e.g., *wrong* above), rather than using it with an evaluatively positive adjective.

4.2 Inherent uncertainty vs. speaker uncertainty

The secondary, tentative analysis I conducted, including different adjective classes, points to an effect of scale structure, or vagueness in particular, on the availability of speaker ignorance interpretations with *at least+adj* constructions, but also with unmodified uses of adjectives (**null+adj** items). I hypothesize that this seeming preference of SpI inferences for the class of relative adjectives in the *at least* and **null** conditions may indicate the following: Our participants were not judging the speaker's uncertainty about the truth of the stronger adjective in the conclusion on the basis of whether the speaker has enough evidence (SpI interpretation), but they were rather judging the inherent uncertainty of the given relative, vague adjective about what the corresponding threshold is. To illustrate with the scale <attractive, stunning>, participants' relevant judgements that Maria does not know whether the stronger alternative (*stunning*) holds of an individual might reflect that they consider that Maria cannot make up her mind about whether that individual (that is attractive) also counts as stunning due to the vagueness of the predicate (i.e., what counts as *stunning*), rather than due to not having seeing that individual well enough to be able to judge their appearance (incomplete evidence). That would necessitate that in the *just* condition participants do not make the same consideration about the alternative mentioned in the conclusion. Rather all they care about in this condition is the (applicability of the) term (in Maria's utterance) modified by *just*, which gives an early signal that any other alternative expression is to be ruled out.

More importantly, the analysis under discussion strongly suggests that future research should take into account and systematically control for more properties of the measurement scales associated with the different adjective scales tested here, e.g., evaluative polarity (see also above), semantic distance of scale-mates, boundedness, type of standard (van Tiel et al. 2016; Gotzner et al. 2018a,b). Hence, the interaction of scale structure and (ignorance) implicature needs to be looked into further (see also Gotzner et al. 2018a,b; Leffel et al. 2019).

4.3 On the choice of the *just* control condition

As revealed by Figure 1 and briefly noted in Section 3.4, the two *just* conditions received remarkably different validity ratings, with *just+adj* getting higher rates

overall (mean=2.47) than **just+num** (mean=1.21). This is most probably due to the fact that these conditions used different exclusive adverbs: *exactly* with numerals, *simply* with adjectives. Higher rates in the latter case could be attributed to participants interpreting *aplos* ‘simply’ with an emphatic (cf. *simply the best*; Beltrama 2021) rather than an exclusive function. Such a use could be compatible with speaker’s ignorance about whether a stronger alternative holds. If anything, possible emphatic interpretations of *simply*+ADJ should mask the interaction of **just** vs. **null** with **adj** vs. **num**, or the likelihood of bare adjectives triggering upper-bounded interpretations (simple effect of **null**), which were both already found to be significant (see Table 1).

Another candidate explanation could be that participants do not take the adjective modified in Maria’s utterance and that mentioned in the conclusion to sit on the same scale or to be in a competition (as to scalar strength). If that were the case, we would not expect to see much of a difference in validity ratings across the three **adjective** conditions, contrary to fact (cf. simple effects of **at least** and of **null** in Table 1).

Finding an alternative control condition that is common for both modified numerals and adjectives (though see Section 3.2.2 and footnote 7) or using a different manipulation of speaker’s knowledgeability (about the stronger alternative) could eliminate the possible confounding emphatic interpretations of the control condition of adjectives.

5 Conclusion

In this paper, I set out to explore to what extent speaker ignorance inferences arise uniformly for different scalar expressions modified by *at least* in Greek (*to lighotero*), namely numerals and adjectives. I present experimental evidence that speaker ignorance inferences are less likely to arise when the Greek counterpart of *at least* is used as an adjective modifier than as a numeral modifier. I also found indications that scale structure, or vagueness more precisely, may affect the availability of speaker ignorance inferences in the case of adjectives, in line with recent findings on the pragmatic interpretation of gradable adjectives (see, e.g., Gotzner et al. 2018a,b; Leffel et al. 2019). These findings reveal that scalar diversity is exhibited by speaker ignorance inferences of superlative modifiers, too. I discussed possible factors to account for the observed variability and argued that the theory of superlative modifiers needs to be more fine-grained to accommodate them. Finally, further investigation into the interplay of scale structure of gradable adjectives and implicature computation is deemed necessary.

References

- Alexandropoulou, S., Jakub Dotlačil & Rick Nouwen. 2016. *At least* ignorance inferences come at a processing cost: Support from eye movements. In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 26, 795–813.
- Alexandropoulou, Stavroula. 2018. *On the pragmatics of numeral modifiers: The availability and time-course of variation, ignorance and indifference inferences*: Utrecht University PhD dissertation.
- Alexandropoulou, Stavroula, Jakub Dotlačil, Yaron McNabb & Rick Nouwen. 2015. Pragmatic inferences with numeral modifiers: Novel experimental data. In Sarah D’Antonio, Mary Moroney & Carol Rose Little (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 25, 533–549.
- Alexandropoulou, Stavroula, Jakub Dotlačil & Rick Nouwen. 2017. Pragmatic effects of *more than* and *at least* in incremental interpretation. In Dan Burgdorf & Jacob Collard (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 27, 680–699.
- Barouni, Maria. 2018. *Topics in the Syntax-Semantics of Greek particles*: University of Crete PhD dissertation.
- Beltrama, Andrea. 2021. *Just perfect, simply the best*: An analysis of emphatic exclusion. *Linguistics and Philosophy* doi:10.1007/s10988-021-09326-x.
- Bolinger, Dwight. 1972. *Degree Words*. Den Haag: Mouton.
- Brown, Penelope & Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Use*, vol. 4 Studies in Interactional Sociolinguistics. Cambridge: Cambridge University Press.
- Buccola, Brian & Andreas Haida. 2020. How obligatory irrelevance, symmetric alternatives, and dense scales conspire: The case of modified numerals and ignorance. In Joseph Rhyne, Kaelyn Lamp, Nicole Dreier & Chloe Kwon (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 30, 464–484.
- Büring, Daniel. 2008. The least *at least* can do. In Charles B. Chang & Hannah J. Haynie (eds.), *West Coast Conference on Formal Linguistics (WCCFL)*, vol. 26, 114–120.
- Chierchia, Genaro, Danny Fox & Benjamin Spector. 2012. Scalar implicature as a grammatical phenomenon. In Claudia Maienborn, Klaus von Stechow & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, 2297–2331. Berlin: De Gruyter Mouton.
- Christensen, Rune Haubo Bojesen. 2019. Ordinal—Regression models for ordinal data. [R package version 2019.12-10]. <https://CRAN.R-project.org/package=ordinal>.
- Ciardelli, Ivano, Liz Coppock & Floris Roelofsen. 2018. Implicatures of modified

- numerals: Quality or Quantity? In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Sinn und Bedeutung*, vol. 21 1, 283–300.
- Cohen, Ariel & Manfred Krifka. 2014. Superlative quantifiers and meta-speech acts. *Linguistics and Philosophy* 37(1). 41–90. doi:[10.1007/s10988-014-9144-x](https://doi.org/10.1007/s10988-014-9144-x).
- Coppock, Elizabeth & Thomas Brochhagen. 2013a. Raising and resolving issues with scalar modifiers. *Semantics & Pragmatics* 6(3). 1–57. doi:[10.3765/sp.6.3](https://doi.org/10.3765/sp.6.3).
- Coppock, Elizabeth & Thomas Brochhagen. 2013b. Diagnosing truth, interactive sincerity, and depictive sincerity. In Todd Snider, Brendan Jackson & Tanya Matthews (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 23, 358–375.
- Cremers, Alexandre, Liz Coppock, Jakub Dotlačil & Floris Roelofsen. 2021. Ignorance implicatures of modified numerals. *Linguistics and Philosophy* doi:[10.1007/s10988-021-09336-9](https://doi.org/10.1007/s10988-021-09336-9).
- Cummins, Chris & Napoleon Katsos. 2010. Comparative and superlative quantifiers: pragmatic effects of comparison type. *Journal of Semantics* 27(3). 271–305. doi:[10.1093/jos/ffq006](https://doi.org/10.1093/jos/ffq006).
- Davies, Mark. 2008-. The Corpus of Contemporary American English (COCA). Available online at <http://www.american corpus.org>.
- Dieuleveut, Anouk, Emmanuel Chemla & Benjamin Spector. 2019. Distinctions between primary and secondary scalar implicatures. *Journal of Memory and Language* 106. 150–171. doi:[10.1016/j.jml.2019.02.008](https://doi.org/10.1016/j.jml.2019.02.008).
- Doran, Ryan, Rachel E. Baker, Yaron McNabb, Meredith Larson & Gregory Ward. 2009. On the non-unified nature of scalar implicature: An empirical investigation. *International Review of Pragmatics* 1(2). 211–248. doi:[10.1163/187730909X12538045489854](https://doi.org/10.1163/187730909X12538045489854).
- Doran, Ryan, Gregory Ward, Meredith Larson, Yaron McNabb & Rachel E. Baker. 2012. A novel experimental paradigm for distinguishing between ‘what is said’ and ‘what is implicated’. *Language* 88(1). 124–154. doi:[10.1353/lan.2012.0008](https://doi.org/10.1353/lan.2012.0008).
- Fox, Danny. 2014. Cancelling the maxim of quantity: Another challenge for a Gricean theory of scalar implicatures. *Semantics & Pragmatics* 7(5). 1–20. doi:[10.3765/sp.7.5](https://doi.org/10.3765/sp.7.5).
- Geurts, Bart. 2011. *Quantity Implicatures*. Cambridge: Cambridge University Press.
- Geurts, Bart, Napoleon Katsos, Chris Cummins, Jonas Moons & Leo Noordman. 2010. Scalar quantifiers: Logic, acquisition and processing. *Language and Cognitive Processes* 25(1). 130–48. doi:[10.1080/01690960902955010](https://doi.org/10.1080/01690960902955010).
- Geurts, Bart & Rick Nouwen. 2007. At least et al.: The semantics of scalar modifiers. *Language* 83(3). 533–559. doi:[10.1353/lan.2007.0115](https://doi.org/10.1353/lan.2007.0115).
- Gotzner, Nicole. 2021. The use and interpretation of gradable adjectives. Slides for an invited talk at Bochum Language Colloquium, 28.1.2021, Ruhr University Bochum.

- Gotzner, Nicole & Anton Benz. 2021. Implicatures in (non-) monotonic environments. Slides for paper presented at Sinn und Bedeutung 26.
- Gotzner, Nicole & Sybille Kiziltan. forthcoming. She is brilliant! Distinguishing different readings of relative adjectives. In Nicole Gotzner & Uli Sauerland (eds.), *Measurement, Numerals and Scales. Essays in Honour of Stephanie Solt* Palgrave Studies in Pragmatics, Language and Cognition, Palgrave Macmillan.
- Gotzner, Nicole & Diana Mazzarella. 2021. Face management and negative strengthening: The role of power relations, social distance and gender. *Frontiers in Psychology* doi:10.3389/fpsyg.2021.602977.
- Gotzner, Nicole, Stephanie Solt & Anton Benz. 2018a. Scalar diversity, negative strengthening, and adjectival semantics. *Frontiers in Psychology* 9. 1–13. doi:10.3389/fpsyg.2018.01659.
- Gotzner, Nicole, Stephanie Solt & Anton Benz. 2018b. Adjectival scales and three types of implicature. In Sireemas Maspong, Brynhildur Stefánsdóttir, Katherine Blake & Forrest Davis (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 28, 409–432.
- Grice, H. Paul. 1989. *Studies in the Way of Words*. Harvard: Harvard University Press.
- Institute for Language and Speech Processing - Athena Research Center. 2015. Hellenic national corpus.
- Kennedy, Chris. 2007. Vagueness and grammar: The semantics of relative and absolute gradable predicates. *Linguistics and Philosophy* 30(1). 1–45. doi:10.1007/s10988-006-9008-0.
- Kennedy, Christopher. 2015. A "de-Fregean" semantics (and neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics* 8(10). 1–44. doi:10.3765/sp.8.10.
- Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification and the semantics of gradable predicates. *Language* 81(2). 345–381. doi:10.1353/LAN.2005.0071.
- Krifka, Manfred. 2009. Approximate interpretations of number words: A case for strategic communication. In Erhard Hinrichs & John Nerbonne (eds.), *Theory and Evidence in Semantics*, 109–132. Stanford: CSLI Publications.
- Leffel, Timothy, Alexandre Cremers, Nicole Gotzner & Jacopo Romoli. 2019. Vagueness in implicature: The case of modified adjectives. *Journal of Semantics* 36(2). 317–348. doi:10.1093/jos/ffy020.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Mazzarella, Diana & Nicole Gotzner. 2021. The polarity asymmetry of negative strengthening: Dissociating adjectival polarity from face-threatening potential. *Glossa: A Journal of General Linguistics* 6(1). 47. doi:10.5334/gjgl.1342.

- McNabb, Yaron. 2015. Differences and similarities between scalar inferences and scalar modifiers: The case of quantifiers. In Sarah D'Antonio, Mary Moroney & Carol Rose Little (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 25, 267–287.
- McNabb, Yaron, Stavroula Alexandropoulou, Dominique Blok, Sofia Bimpikou & Rick Nouwen. 2016. The likelihood of upper-bound construals among different modified numerals. In Nadine Bade, Polina Berezovskaya & Anthea Schöller (eds.), *Sinn und Bedeutung*, vol. 20, 497–514.
- McNabb, Yaron & Doris Penka. 2015. An experimental investigation of ignorance inferences and authoritative interpretations of superlative modifiers. Ms.
- Mendia, Jon Ander. 2016. Reasoning with partial orders: Restrictions on ignorance inferences of superlative modifiers. In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 26, 489–505.
- Meyer, Marie-Christine. 2013. *Ignorance and Grammar*: MIT, Massachusetts PhD dissertation.
- Nakanishi, Kimiko & Hotze Rullmann. 2009. Epistemic and concessive interpretations of 'at least'. Slides for paper presented at CLA, Carleton University, May 24.
- Nouwen, Rick. 2010. Two kinds of modified numerals. *Semantics and Pragmatics* 3(3). 1–41. doi:[10.3765/sp.3.3](https://doi.org/10.3765/sp.3.3).
- Nouwen, Rick. 2015. Modified numerals: The epistemic effect. In Luis Alonso-Ovalle & Paula Benito-Ménendez (eds.), *Epistemic indefinites: Exploring modality beyond the verbal domain*, chap. 11, 244–266. Oxford: Oxford University Press.
- Nouwen, Rick, Stavroula Alexandropoulou & Yaron McNabb. 2019. Experimental work on the semantics and pragmatics of modified numerals. In Chris Cummins & Napoleon Katsos (eds.), *Handbook of Experimental Semantics and Pragmatics*, 178–194. Oxford: Oxford University Press.
- van Rooij, Robert & Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Logic, Language and Information* 13. 491–519. doi:[10.1007/s10849-004-2118-6](https://doi.org/10.1007/s10849-004-2118-6).
- Rotstein, Carmen & Yoad Winter. 2004. Total adjectives versus partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics* 12. 259–288. doi:[10.1023/B:NALS.0000034517.56898.9a](https://doi.org/10.1023/B:NALS.0000034517.56898.9a).
- Sassoon, Galit. & Natalia Zevakhina. 2012. Granularity shifting: Experimental evidence from degree modifiers. In Anca Chereches (ed.), *Semantics and Linguistic Theory (SALT)*, vol. 22, 226–246.
- Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27. 367–391. doi:[10.1023/B:LING.0000023378.71748.db](https://doi.org/10.1023/B:LING.0000023378.71748.db).

- Sauerland, Uli. 2013. Presuppositions and the alternative tier. In Todd Snider, Brendan Jackson & Tanya Matthews (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 23, 156–173.
- Schwarz, Bernhard. 2016. Consistency preservation in quantity implicature: The case of *at least*. *Semantics and Pragmatics* 9(1). 1–47. doi:[10.3765/sp.9.1](https://doi.org/10.3765/sp.9.1).
- Solt, Stephanie. 2015. Measurement scales in natural language. *Language and Linguistics Compass* 9(1). 14–32. doi:[10.1111/lnc3.12101](https://doi.org/10.1111/lnc3.12101).
- Spector, Benjamin. 2015. Why are class B modifiers global PPIs? Hurford disjunctions as a model for class b modifiers. Handout for a paper presented at the Workshop on Negation and Polarity in Jerusalem.
- Spychalska, Maria. 2015. At least not false, at most possible: Between truth and assertibility of superlative quantifiers. *Synthese* 195(2). 1–32. doi:[10.1007/s11229-014-0615-y](https://doi.org/10.1007/s11229-014-0615-y).
- van Tiel, Bob, Emiel van Miltenburg, Natalia Zevakhina & Bart Geurts. 2016. Scalar diversity. *Journal of Semantics* 33(1). 137–175. doi:[10.1093/jos/ffu017](https://doi.org/10.1093/jos/ffu017).
- Westera, Matthijs & Adrian Brasoveanu. 2014. Ignorance in context: The interaction of modified numerals and QUDs. In Todd Snider, Sarah D’Antonio & Mia Weigand (eds.), *Semantics and Linguistic Theory (SALT)*, vol. 24, 414–431.

Stavroula Alexandropoulou
 Campus Golm
 Cognitive Sciences
 Linguistics Department
 Karl-Liebknecht-Straße 24-25
 14476 Potsdam
stavroulaale@gmail.com

6 Appendix: English translations of examples of bad and good filler items

Bad filler

- (6) Maria says:

The administration of the tech companies continue to protest.

Conclusion: Maria doesn’t know whether the administration of the tech companies stopped protesting.

Scalar diversity and ignorance inferences

Good filler

(7) Maria says:

I could not distinguish from so far whether that silhouette belonged to Artemi.

Conclusion: Maria doesn't know whether the silhouette she saw from afar belonged to Artemi.