Negation, polarity and scale structure: Different inferences of absolute adjectives[1]

Stavroula ALEXANDROPOULOU — *University of Potsdam*
Nicole GOTZNER — *University of Osnabrück*

Abstract. This work investigates the interpretation of absolute gradable adjectives like *clean* in comparison with their stronger scale-mates (*pristine*) and corresponding antonyms (*dirty* and *filthy*) in the scope of negation or in the absence of it. We find that participants distinguish between non-negated absolute terms differing in informational strength (e.g., *clean* vs. *pristine*). However, such distinctions are fewer in the scope of negation. Under negation, weak absolute adjectives entail the antonym of a given pair (e.g., *not clean* ⇒ 'dirty'), while the fine granularity of the underlying measurement scale appears to be responsible for additional interpretations of absolute adjective expressions, such as middling interpretations ('neither clean nor dirty') and inferences to the antonym. Overall, our findings endorse both the standard absolute, contradictory effect of negation on the interpretation of absolute adjectives as well as the less typical attenuating effect of negation, mostly discussed in relation to relative gradable adjectives (Horn, 1989). We conclude that different properties of measurement scales—scale structure and scale granularity—as well as evaluative polarity, play an essential role in the derivation of different (pragmatic) inferences of gradable adjectives (see also Gotzner et al., 2018b).

Keywords: gradable adjectives, scale structure, negation, polarity, granularity, scalar implicature, negative strengthening

1. Introduction

Gradable adjectives like *large* have mostly attracted the attention of semantic theory with respect to their lexical semantic properties (e.g., the scale structure they use), while less consideration has been given to the pragmatic inferences they are associated with. Speakers make use of different alternative gradable expressions to communicate the varying degrees to which a property holds of an individual. This may bring about a number of distinct pragmatic inferences. (1) illustrates different pragmatic inferences gradable adjectives can trigger, depending on whether they appear unembedded (as in B_1) or embedded under negation (B_2), or on whether they have a positive or negative valence (e.g., *large* vs. *small* in B_2 vs. B_3).

(1) A: What size is your apartment?
B_1: It is large.
⇝ 'It is large but not gigantic' (*scalar implicature*)

[1]The research leading to the results reported on in the current paper has received funding from the German Research Foundation (DFG) as part of the Emmy Noether programme: Emmy Noether Grant Nr. GO 3378/1-1 awarded to Nicole Gotzner. We would like to thank our mercator fellows Louise McNally and Ira Noveck, the other SPA lab members—Henrik Discher, Marisha Herb, Radim Lacina, Ella Markham—as well as Stephanie Solt, the LACRE group at Cardiff University, the audiences of SPA lab's kickoff workshop on "Scales, degrees and implicature", of the ESSLLI 2021 workshop on "Approaches to implicature: Rational choice and/or exhaustification", and of the Sinn und Bedeutung 26 conference for useful discussion and valuable feedback.

 In: Daniel Gutzmann and Sophie Repp (eds.): Proceedings of Sinn und Bedeutung 26, pp. 35–54. University of Cologne.

B_2: It is not large.
$\rightsquigarrow$ 'It is (rather) small' (*negative strengthening*)

B_3: It is not small.
$\rightsquigarrow$ 'It is medium-size' (*middling interpretation*)

More specifically, B_1's answer to A triggers the upper-bounded inference, or so-called scalar implicature, that A's apartment is large but not gigantic. This arises via reasoning about alternative expressions B_1 could have uttered, such as *The apartment is gigantic*, but didn't. That is, this reasoning involves the alternative expressions *large* and *gigantic*, which form a Horn lexical scale and stand in an asymmetric entailment relationship such that the stronger expression *gigantic* entails the weak one, *large*, but not vice versa (e.g, Horn, 1989). Another type of pragmatic inference has been observed to occur when gradable adjectives like *large* are embedded under negation, namely, negative strengthening (Horn, 1989). To exemplify this type of inference, B_2's utterance in (1) may be taken to convey that A's apartment is rather small. Arguably, this is a stronger interpretation than the literal interpretation of B_2's utterance, i.e., 'it is less than large', that allows for A's apartment being medium-size. On the other hand, if the negative antonym of *large* appears under negation, i.e., *not small* (see B_3 above), then the resulting inference is usually the aforesaid middling situation where A's apartment is neither large nor small, but just medium-size. In other words, *not small* is not likely to be negatively strengthened into 'large'. This polarity asymmetry in the interpretation of gradable adjectives under negation is well-known in the literature (see, e.g., Brown and Levinson, 1987 and Horn, 1989) and has also been experimentally corroborated. In particular, a number of recent experimental studies (Ruytenbeek et al., 2017; Tessler and Franke, 2019; Gotzner and Mazzarella, 2021; Gotzner and Kiziltan, 2021; Mazzarella and Gotzner, 2021) have conclusively shown that, under negation, positive adjective terms (e.g., *large*) are more likely to receive a negative strengthening interpretation than their corresponding negative antonyms (e.g., *small*).

The above experimental studies have primarily focused on the interpretation of a certain class of gradable adjectives under negation, namely, of relative adjectives, which have a context-dependent standard of comparison and interpretation (see Section 2.1.). Among those studies, only Gotzner and Kiziltan (2021) have investigated the pragmatic inferences of weak relative adjectives in relation to their stronger counterparts (e.g., *large* vs. *gigantic*) and the corresponding negative antonyms with or without negation (however, see Gotzner et al., 2018b, a for an investigation of weak and strong adjective scale-mates from supposed Horn lexical scales). They found that non-negated weak relative adjectives are interpreted distinctly from their stronger counterparts. Under negation, a polarity asymmetry is exhibited by weak terms, such that weak positive adjectives like *large* receive a negative strengthening interpretation ('rather small'), while weak negative adjectives (*small*), as well as the respective stronger adjectives (*gigantic* and *tiny*), receive a middling interpretation ('neither large nor small'/'medium-size'). The authors propose that the scale structure of gradable adjectives should also be taken into consideration in accounting for the pragmatic reasoning they involve. Crucially, there is robust evidence showing that scale structure affects the derivation of pragmatic inferences of weak gradable adjectives (Gotzner et al., 2018b, a; Leffel et al., 2019). That is to say, the two classes of gradable adjectives differing in scale structure, i.e., relative and absolute adjectives, are found to trigger pragmatic inferences differentially: While relative adjectives are less good triggers of

upper-bounded interpretations compared to absolute adjectives, the reverse pattern is observed with respect to negative strengthening. Yet pragmatic inferences of weak and strong absolute adjectives like *clean* and *pristine* have been less systematically explored so far, especially as far as their interpretation under negation is concerned (but see Paradis and Willners, 2006; Gotzner et al., 2018b, a; Leffel et al., 2019). The present work extends on Gotzner and Kiziltan's (2021) systematic study of relative adjectives, zeroing in on absolute adjectives, thereby completing the empirical picture.

Specifically, we aim to probe the interpretation of weak absolute terms (*clean*) in relation to their stronger counterparts (*pristine*) and the corresponding negative antonyms (*dirty*, *filthy*) in the presence and absence of negation. Doing so will allow us to assess the different types of inferences absolute expressions trigger in their non-negated and negated forms, and ultimately determine the role of scale structure in the derivation of pragmatic inferences of gradable adjectives overall. To foreshadow our main results, we find evidence that non-negated weak absolute adjectives are distinguished from their strong counterparts differing in informational strength and that negated weak absolute adjectives entail the antonym of a given pair (e.g., *not dirty* $\Rightarrow$ 'clean') but may also convey an attenuating sense in a context with a fine granularity level (e.g., *not dirty* $\rightsquigarrow$ 'neither (prototypically) clean nor (prototypically) dirty'). Hence, our findings endorse both the typical absolute, contradictory effect of negation on the interpretation of absolute adjectives as well as the less standard attenuating effect of negation.

The paper is structured as follows: Section 2 introduces basic concepts and theories, as well as findings about the meaning and use of absolute adjectives that will be relevant for our study. Section 3 presents our experiment, the obtained results and the interpretation thereof. Section 4 discusses the implications of our findings and Section 5 concludes.

2. Theoretical background

2.1. The relative vs. absolute distinction and the entailments thereof

A prominent semantic approach to gradable predicates models their meaning via so-called measurement scales, which are sets of totally ordered degrees with respect to some dimension (Bartsch and Vennemann, 1973; Cresswell, 1976; Bierwisch, 1989; Kennedy and McNally, 2005; Kennedy, 2007; see Solt, 2015 for an overview). Following Kennedy (2007), a gradable adjective on its positive form relates an object to a degree on a given underlying measurement scale. A gradable adjective like *dirty* is assumed to express that the degree of dirt on the underlying scale of dirt exceeds a standard degree. Gradable adjectives are split into two main classes, absolute and relative adjectives, depending on the type of standard degree they involve: The standard degree is taken to be a fixed and context-invariant value on the underlying measurement scale for absolute adjectives, while, for relative adjectives, it is assumed to be a context-dependent threshold based on the relevant comparison class (Rotstein and Winter, 2004; Kennedy and McNally, 2005; Kennedy, 2007). The class of absolute adjectives further encompasses the classes of minimum-standard adjectives like *dirty* and of maximum-standard adjectives like its antonym *clean*. Minimum standard adjectives require that an individual possess a non-zero/*minimal* degree of the property at stake in order to qualify as such, with the

minimum value on the underlying measurement scale being the relevant standard degree. That is, an individual qualifies as *dirty* as long as it exhibits a non-zero degree of dirt. On the other hand, maximum-standard absolute adjectives require the opposite: That an individual possesses the *maximal* degree of the property at stake, which corresponds to the maximal degree of the relevant measurement scale.[2]

Crucially, the type of standard invoked by gradable adjectives affects entailment relations between antonymic pairs. The negation of an absolute adjective entails the assertion of its antonym (Cruse, 1986; Rotstein and Winter, 2004), as in (2). This is because a minimal positive degree corresponds to a negative maximal degree on the relevant measurement scale (Kennedy, 2001, 2007). This is not the case for relative adjectives, which do not give rise to such an entailment relation, as shown in (3), (Kennedy, 2007). *Large* and *small* make use of the same dimension and degrees, and impose reverse orderings. However, the antonymic relative terms allow for a middle ground between their extensions because their associated, context-dependent, standards may be different (Kennedy, 2007). This is why the entailments in (3) do not go through.

(2) a. The shirt is not clean ⇒ The shirt is dirty
b. The shirt is not dirty ⇒ The shirt is clean

(3) a. The shirt is not large ⇏ The shirt is small
b. The shirt is not small ⇏ The shirt is large

(4) a. The shirt is not very/completely clean ⇏ The shirt is dirty
b. The shirt is not pristine ⇏ The shirt is filthy/dirty

Interestingly, the entailments to the antonym do not hold for modified absolute adjectives or stronger scale-mates of absolute adjectives as shown in (4). Cases like those in (4) are in fact compatible with a number of different possibilities (the shirt being clean, dirty or filthy). Arguably, the availability of the modified forms or stronger scale-mates introduces a fine granularity level, whereby smaller amounts of the property at stake, e.g., of dirt, become relevant (see, e.g., Beltrama, 2022 and references therein, on modified forms). Hence, distinctions between *dirty* and *filthy* or *clean* and *pristine* become relevant, too. Importantly, on such a fine-grained scale, smaller unnoticeable degrees of dirt shift from the interpretation of *clean* to the interpretation of *dirty* (see Sassoon and Zevakhina, 2012), hence, the range of degrees conveyed by *clean* is largely restricted. Rotstein and Winter (2004) further note that, in some contexts, modified absolute terms by *almost*/*slightly* break the complementarity of antonymic pairs of absolute adjectives (e.g., *dirty* vs. *clean*) manifested by their entailment patterns in (2). This is so, as in some contexts modified expressions (like *almost dirty* according to Rotstein and Winter 2004) can be used to describe a middling situation like being 'neither clean nor dirty'. Obviously, one such context is when attention is drawn to the presence of a small amount of dirt, e.g., on a glass (see (5)), thus, a context with a fine-granularity level of dirt.

(5) It is certainly not clean, since it has some small spots on it, but it is not really dirty...
(adapted from Rotstein and Winter, 2004, p. 266)

[2] A third class of absolute adjectives are the so-called totally closed scale adjectives, whose standard degree may be located at the scale's minimum or maximum endpoint (Kennedy, 2007; Sassoon and Toledo, 2015). The pairs *full*/*empty* and *closed*/*open* are examples of this type of absolute adjectives.

Therefore, we may conclude that a context imposing a fine-grained scale can license a middle ground ('neither clean nor dirty') between the extensions of antonymic absolute adjectives, thereby breaking the complementarity of the relevant antonymous adjectives. For antonymic pairs of strong absolute adjectives, such as *pristine/filthy*, that would mean that their semantic interpretation under negation includes all possibilities in the scalar range below the threshold of *pristine/filthy*, including the middling possibility, viz. it amounts to 'less than pristine/filthy', respectively. Moreover, the presence of a middle ground between antonymic terms allows for such constructions to be pragmatically strengthened under negation, as shown in the next section.

2.2. Different pragmatic inferences under negation

One crucial distinction with respect to semantic and pragmatic meanings of negated expressions concerns that between contradictory and contrary antonyms. Contraries, as opposed to contradictories, allow for a middle ground, *the unexcluded middle*, between the extensions of the two antonymic members. Note that contrary antonymic pairs include relative adjectives, e.g., *large* vs. *small*. Horn (1989) observes that weak gradable predicates participating in antonymic pairs that constitute contraries may implicate their antonym under the scope of negation. The resulting pragmatic interpretation is referred to as negative strengthening, a type of implicature that Horn derives on the basis of his R-principle ("Say no more than you must"; R-based strengthening). For example, *not large* may be interpreted as 'rather small' by means of negative strengthening. Stronger relative terms, on the other hand, do not exhibit negative strengthening (e.g., Horn, 1989; Israel, 2004; Gotzner and Kiziltan, 2021).

A further crucial factor in the interpretation of negated adjectives is evaluative polarity, which is the reason why negative strengthening has been argued to be socially motivated (Brown and Levinson, 1987; Horn, 1989): While the negated positive term *not large* tends to implicate 'rather small', this is less likely to be the case for the negative antonym *not small* (see also Bolinger, 1972; Ducrot, 1973; Brown and Levinson, 1987; Levinson, 2000; Solstad and Blutner, 2000).[3] The negated negative antonyms (e.g., *not small*) instead tend to communicate the middle range of degrees (e.g., those qualifying as 'neither large nor small') and are, thus, interpreted as a middling term like *medium-size*. This polarity asymmetry has been corroborated experimentally (Colston, 1999; Fraenkel and Schul, 2008; Ruytenbeek et al., 2017; Tessler and Franke, 2019; Gotzner and Kiziltan, 2021; Gotzner and Mazzarella, 2021). In addition, Gotzner and Kiziltan (2021) find that strong positive (*gigantic*) and negative relative terms (*tiny*), too, receive a middling interpretation under negation.

The situation is different for contradictory terms, which under negation implicate their antonym via an entailment. Evidently, absolute adjectives participate in contradictory antonymic pairs,

[3]Experimental studies have tested the politeness-based explanation of negative strengthening and shown that face-management considerations play a role in negative strengthening in that sociological variables affect the interpretation of negated positive and negative terms (Gotzner and Mazzarella, 2021). Yet the asymmetry in the extent to which positive and negative terms trigger negative strengthening could be based on a more general role of evaluative polarity (independent of the face-threatening potential in context, see Mazzarella and Gotzner, 2021) or a more general interaction of different pragmatic principles based on informativity and brevity or markedness (Horn, 1989; Krifka, 2007b; Ruytenbeek et al., 2017).

e.g., *clean* vs. *dirty*, cf. their entailment patterns in (2) in the previous section. Yet, as we have seen, the availability of modified adjective forms or stronger terms (e.g., *pristine*) may help create a middle ground for absolute adjectives and, thus, may invite pragmatic inferences like those discussed above for contrary antonymic pairs.

Remarkably, there is only very little experimental evidence as to the interpretation of absolute adjectives under negation (see Leffel et al., 2019 on the interpretation of modified absolute adjectives by *very* under negation). Paradis and Willners (2006) tested a limited set of contradictory antonymic pairs in Swedish (i) that use totally closed measurement scales and have absolute interpretations, e.g., *full* vs. *empty*, or (ii) that even contain non-gradable adjectives like *free* (vs. *bound*), and *dead*/*alive*. They subsumed those contradictory antonyms under the so-called category of bounded antonymic pairs and tested their interpretation in their negated and non-negated form. The authors report on different interpretation patterns for contradictory/bounded antonymic terms, some of which indicated that an absolute interpretation can be coerced into a relative interpretation, so negation may then have an attenuating effect like that previously discussed for contrary antonyms and relative adjectives in particular, as opposed to a literal contradictory effect.

As Paradis and Willners's (2006) set of contradictory adjectives is limited and heterogeneous, we cannot firmly tell whether the aforementioned attenuating effect of negation on the interpretation of absolute adjectives like *empty* may be modulated by politeness/evaluative polarity (see footnote 3), or to what extent it generalizes to the class or sub-classes of absolute adjectives. Given the existing evidence, it is not entirely clear to what extent and under which conditions the interpretation of negated absolute adjectives can be coerced into relative interpretations, or, put differently, into the interpretations Gotzner and Kiziltan (2021) attested for negated relative expressions, namely, negative strengthening and middling. This is the starting point of our study. More precisely, in the following, we set out to systematically test for the interpretation of absolute adjectives in the absence or presence of negation.

3. Current study

3.1. Research questions & predictions

In the current study, we aim to investigate the specific ranges weak scalar adjectives like *clean* communicate in comparison with their stronger scale-mates (e.g., *pristine*) and their corresponding negative antonyms (e.g., *dirty* vs. *filthy*), both in their non-negated form and when they appear under negation. This will allow us to identify the variety of meanings the specific ranges reflect and, thus, to determine which inferences are triggered by different absolute expressions. Note that previous work has either investigated the interpretation of antonymic pairs with weak adjective terms only (e.g., Paradis and Willners, 2006; Ruytenbeek et al., 2017; Tessler and Franke, 2019; Gotzner and Mazzarella, 2021), or the relation between different weak and strong adjective terms (e.g., Doran et al., 2012; van Tiel et al., 2016; Gotzner et al., 2018b, a), and only Gotzner and Kiziltan (2021) studied weak and strong antonymic pairs together. While they did so only for relative adjectives, here we investigate absolute adjectives in the same experimental setup. This will allow us to determine the role of scale structure in the derivation of different inferences of gradable adjectives overall.

Different inferences of absolute adjectives

Table 1: Different interpretations of negated positive weak and strong absolute adjective terms, illustrated for the adjective scale of dirt. "??" indicate that the respective inference is not predicted to be available by standard theory.

Adjective type/ Inference	*The door is*	**Absolute**	
		Strong *not pristine*	**Weak** *not clean*
Entailment	'The door is'	'less than pristine'	'dirty'
Indirect scalar implicature	'The door is'	'clean'	??
Middling interpretation	'The door is'	??	??
Negative strengthening	'The door is'	'(rather) dirty'	??

We hypothesize that fewer clear-cut distinctions arise between different absolute terms in negated than in non-negated environments, and that properties of the measurement scales absolute adjectives are associated with are important in understanding their behavior under negation. In what follows, we discuss a number of candidate interpretations predicted for absolute gradable adjectives, stemming from the discussion in Section 2.

In their non-negated form, weak absolute adjectives (*clean/dirty*) are expected to be distinguished from their stronger counterparts (*pristine/filthy*), given that the availability of the stronger terms in the task encourages sensitivity to the informational strength of the different terms. This should be manifested by weak and stronger non-negated absolute adjectives being assigned to distinct ranges of values and, thus, portions of the 5-point response scale by participants.

For negated environments, on the other hand, we shall consider a number of different interpretations of absolute adjectives. Table 1 summarizes these candidate interpretations for the positive antonymic terms of the adjective scale of dirt appearing in the predication sentence *The door is not* ADJ, which is similar to the sentences tested in our experiment.

Let's start from the semantic interpretation (Entailment in Table 1) of the two absolute adjective terms differing in strength. The sentence *The door is not* ADJ, where ADJ is a weak absolute adjective like *clean*, has an entailment to its negative antonym, e.g., *dirty*. When ADJ is a strong adjective like *pristine*, the sentence *The door is not pristine* entails that the door is less than pristine. This interpretation is compatible with a situation where the door is clean, dirty, or even filthy.

The stronger scale-mates of absolute adjectives may also give rise to an indirect scalar implicature under negation (Chierchia, 2004; Cremers and Chemla, 2014; Gotzner and Romoli, 2018; Gotzner et al., 2018a), regardless of their polarity. For instance, *The door is not pristine* may trigger the interpretation 'the door is clean', via standard Quantity-based reasoning about the alternatives of the reversed negated scale <not pristine, not clean>, where *not pristine* is weaker/less informative than *not clean* (note that negation reverses entailment relations).

Moreover, we saw in Section 2.2 that middling interpretations have been discussed in relation to relative gradable adjectives when they occur under negation. That is, by virtue of the middle ground between the extensions of the positive and negative terms of an antonymic pair, a negated relative adjective like *not large* may trigger the strengthened interpretation 'neither large nor small'. Importantly, a middling interpretation like that is not predicted to be available

for absolute adjectives, as their lexical semantics, specifying the scale structure these adjectives use, lacks this middle ground between the positive and negative antonymic terms (see Section 2.1). However, as further discussed in that section, a context imposing a fine granularity, where distinctions between weak and strong terms are made relevant, may license a middle ground between the extensions of antonymic absolute terms (e.g., the middle range qualifying as 'neither clean nor dirty' for the scale of dirt), thereby breaking the complementarity of the relevant antonymous adjectives. Recall once again that Paradis and Willners (2006) found divergent patterns in their experiment on contradictory antonyms, suggesting that the interpretation of such antonymic pairs under negation need not be symmetric. On the contrary, negation may have an attenuation effect on the interpretation of contradictory antonyms resembling that of contrary antonymic pairs of relative adjectives.

Finally, negative strengthening is another pragmatic interpretation that becomes available with gradable adjectives under negation. This interpretation, too, is mainly discussed in connection with weak relative adjectives. To illustrate, *The door is not large* receives the strengthened interpretation 'the door is (rather) small'. Crucially, while weak relative adjectives exhibit a polarity asymmetry in terms of negative strengthening (see Section 2.2), no negative strengthening interpretation is predicted to arise with negated weak absolute adjectives, as those have the entailment to the antonym (e.g., *not clean* ⇒ 'dirty'). Hence, no relevant polarity asymmetry is in turn expected to emerge for weak absolute terms under negation. That is, the standard semantic theory predicts that negated positive weak terms like *not clean* should be rated symmetrically to their negated negative antonyms, e.g., *not dirty*. That being said, we would expect negative strengthening to become available with negated strong terms like *not pristine*.

3.2. Methods

3.2.1. Participants

We recruited 60 participants with US IP addresses on Amazon Mechanical Turk (20 women, 39 men, and 1 who did not give gender information, mean age: 33, age range: 23-69). Participants were screened for native language and only included in the analysis if their self-reported native language was English. On the basis of this criterion, we removed the data of 2 participants who reported a native language other than English (N=58). The experiment lasted about 10 to 15 minutes and participants were paid 1 US Dollar in compensation.

3.2.2. Materials & procedure

We created 8 adjective quadruples, each consisting of a weak and a strong evaluatively positive absolute term, and the corresponding weak and strong negative antonyms. These adjectives were embedded in a simple predication statement, either in their non-negated form or under negation. Hence, there were 8 adjectival expressions in total for a given item and the experiment used a 2 Evaluative Polarity (positive, negative) × 2 Scalar Strength (weak, strong) × 2 Negation (non-negated, negated) design. Table 2 displays all absolute adjectives tested in this experiment.

Table 2: Overview of adjectives in non-negated conditions (negated conditions included the same adjectives preceded by *not*. The top row presents the names of each adjective quadruple.

Item/ Condition	**bolt upright**	**flawless**	**healthy**	**immaculate**	**pristine**	**safe**	**silky soft**	**spotless**
Non-negated negative strong	*twisted*	*imperfect*	*sick*	*broken*	*filthy*	*dangerous*	*cracked*	*filthy*
Non-negated negative weak	*bent*	*impure*	*unwell*	*faulty*	*dirty*	*dodgy*	*rough*	*dirty*
Non-negated positive weak	*straight*	*pure*	*well*	*intact*	*clean*	*riskless*	*smooth*	*clean*
Non-negated positive strong	*bolt upright*	*flawless*	*healthy*	*immaculate*	*pristine*	*safe*	*silky soft*	*spotless*

The predication statements with the resulting 8 adjectival expressions (each corresponding to one condition) were presented concurrently in context. There was a unique context for each adjective quadruple, hence, 8 different contexts/items in total. Each context consisted of a rating scenario that involved an action-based task. That is, participants had to give a rating for each statement. This paradigm was the same as the grading paradigm used by Gotzner and Kiziltan (2021), which was inspired by the best response paradigm of Gotzner and Benz (2018) (see also Tessler and Franke, 2019 for a similar paradigm). Here, too, the idea is that, due to the nature of this action-based rating task, distinctions between different interpretations of adjectival expressions should become relevant. That is, this task provides a rating tool of individuals satisfying the property expressed by each adjectival predication and thus enables the ranking of different adjectival expressions. This ultimately reveals fine distinctions between different interpretations of adjectives.

Also, presenting the statements with all the different alternative adjectival expressions concurrently and coupled with a Likert response scale introduces a fine granularity level for the individual scales the different contexts/items make reference to, rendering distinctions between different expressions relevant. Lastly, as becomes obvious, this type of task capitalizes on the notion of evaluative polarity. Table 3 presents an example context for the item ***pristine***. Participants saw all non-negated and negated forms presented concurrently in one context (see Tessler and Franke, 2019 for a discussion of the difference between concurrent and isolated presentation of stimuli).

In each context, there was a speaker with full knowledge (e.g., an examiner, see Table 3) uttering statements about a set of individuals (people, objects or activities). The participants' task was to indicate which rating each individual (e.g., a hospital in Table 3) would receive in terms of a certain aspect (e.g., hygiene standards) based on the respective statements. The judgments were made on a 5-point Likert scale, where 1 represented the (non-negated) negative strong adjective (e.g., *filthy*) and 5 its positive strong antonym (*pristine*). Thus, participants in this paradigm rated the different individuals (e.g., the Saint Anthony's Hospital, the Saint Joseph Hospital, etc.) they had read an evaluative statement about, taking into account the individual adjectival predications present in the display (e.g., *is not filthy*, *is not dirty*, etc.), thereby drawing finer distinctions when interpreting the different adjectival expressions comparatively as to the given dimension. Consequently, we measured participants' interpretation of different adjectival expressions on the same Likert scale.

Table 3: Example item ***pristine***. Only three of the statements are given here, though in the experiment participants saw all 8 statements.

Context:
The government examines the hospitals of a big city for their hygiene standards. The examiner writes a review.
Please decide which rating each hospital gets for its hygiene standards based on the examiner's statements.
1 = filthy; 5 = pristine
The examiner says:
The Saint Anthony's Hospital is not filthy.
1 2 3 4 5
The Saint Joseph Hospital is not dirty.
1 2 3 4 5
The Saint's Mary's Hospital is pristine.
1 2 3 4 5
. . .

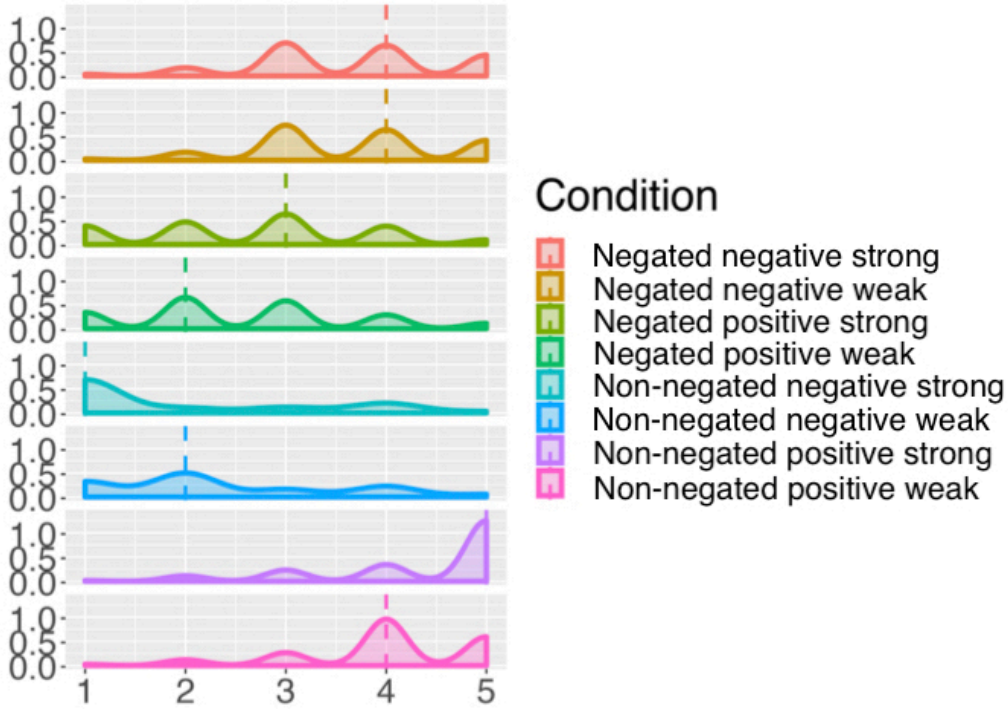
Our three factors, Evaluative Polarity, Scalar Strength and Negation, were all within-subject and within-item. Hence, each participant saw 8 contexts in a randomized order, each with 8 pseudo-randomized statements. The total number of observations was 3712 (64 trials by 58 participants).

The experiment was programmed in HTML and run via Amazon Mechanical Turk's in-built environment. After they responded to the demographic questions, participants read the instructions that also illustrated the experimental task with an example. This example involved an antonymic pair that was not used in the main experiment (i.e., *false* and *true*).

3.3. Results & discussion

We removed ten participants on the basis of their responses to the non-negated strong conditions. Namely, if a participant placed 2 or more (non-negated) strong adjectives at the opposite end of the 1-5 response scale (e.g., if *pristine* was assigned the rating 1, or *filthy* the rating 5, in opposition to what is indicated in the Context, see Table 3), the whole set of data of this participant was excluded from all further analyses. Figure 1 shows density curves for the ratings of the remaining 48 participants across absolute adjective conditions.

Figure 1: Density curves for ratings across absolute adjective conditions (dashed line represents the median).



The resulting data (N = 3072) were analyzed using R (version 4.0.5). Participants' responses were ordered categorical, and, to analyze them, we fitted cumulative link mixed effects models using the `ordinal` package (Christensen, 2019) in R. We included the factors Evaluative Polarity, Scalar strength, and Negation, as well as all their interactions as predictors (fixed effects). We fit a model with sum-coded fixed effects as well as the maximal converging random-effect structure (i.e., random intercepts and slopes for participants, and random slopes for items).

The analysis revealed the following significant effects (summarized in Table 4): (i) a significant main effect of Polarity such that evaluatively positive terms were rated significantly higher than negative terms overall ($\beta = 1.17$, $SD = 0.18$, $z = 6.44$, $p < 0.0001$), (ii) a significant Scalar Strength*Polarity interaction such that overall people distinguish between positive and negative terms differing in scalar/informational strength ($\beta = -0.73$, $SD = 0.14$, $z = -5.31$, $p < 0.0001$), (iii) a 3-way interaction Polarity*Scalar Strength*Negation indicating that the aforesaid distinctions are less pronounced under negation compared to the non-negated environments, ($\beta = -0.71$, $SD = 0.14$, $z = -5.07$, $p < 0.0001$), and (iv) a significant Negation*Polarity interaction effect, such that, averaging over weak and strong terms, the difference in ratings between negative and positive terms is larger in non-negated than in negated environments ($\beta = 2.68$, $SD = 0.38$, $z = 7.14$, $p < 0.0001$).

Crucially, from the main analysis, we conclude that, in the presence of negation, people draw fewer distinctions between adjectival terms than when these terms appear in their non-negated forms, validating our main hypothesis. This becomes obvious if one just glances at the data in Figure 1.

Table 4: Output of cumulative link model with sum coding.
clmm(Score~Negation*ScalarStrength*Polarity + (0+Polarity*Negation:ScalarStrength|Item) + (Polarity*Negation*ScalarStrength|Participant), data = dataset)

	Estimate	SD	z-value	p-value
Negation	0.011850	0.075111	0.158	0.875
ScalarStrength	0.006662	0.084664	0.079	0.937
Polarity	1.169669	0.181666	6.439	0.0001
Negation:ScalarStrength	0.057655	0.065325	0.883	0.377
Negation:Polarity	2.676870	0.375129	7.136	0.0001
ScalarStrength:Polarity	-0.732145	0.137907	-5.309	0.0001
Negation:ScalarStrength:Polarity	-0.709199	0.140001	-5.066	0.0001

In a subsequent analysis, we investigated the effect of Scalar Strength for the individual adjective conditions further. To do so, we ran a model where Scalar Strength was nested under Polarity and Negation. The model revealed that, while the difference in ratings between weak and strong non-negated adjective terms is significant for positive ($\beta = -1.37$, $SD = 0.27$, $z = -5.09$, $p < 0.0001$) and for negative polarity conditions ($\beta = 1.49$, 0.27, $z = 5.43$, $p < 0.0001$), it was not significant for either positive or negative terms under negation (positive: $\beta = -0.08$, $SD = 0.10$, $z = -0.81$, $p = 0.42$; negative: $z = -0.16$, $p = 0.88$). That is to say, participants draw clear-cut distinctions between the interpretation of weak and that of strong non-negated adjective terms (e.g., *clean* vs. *pristine* and *dirty* vs. *filthy*), by using distinct portions of the given scale (cf. in Figure 1, peak on 4 and on 5 for weak and strong positive terms, respectively, and peak on 2 and 1 for weak and strong negative terms, respectively). We take this finding to be evidence of participants clearly distinguishing between terms differing in informational strength. On the other hand, negated weak and strong terms appear to have less clear-cut boundaries (*not clean* ~ *not pristine*, *not dirty* ~ *not filthy*). In what follows, we consider what could be behind the lack of evidence that these terms are treated differently, drawing on our additional hypothesis that properties of the measurement scales used by absolute adjectives may affect their interpretation under negation.

Starting from the two negated negative conditions (e.g., *not dirty/not filthy*), we observe that those receive mostly middle-scale ratings and ratings of 4. We will first be concerned with the weak terms. Ratings of 4 for negated negative weak adjectives like *not dirty* reflect an entailment to the positive antonym, e.g., *not dirty* ⇒ 'clean'. However, negated negative weak terms are not interpreted strictly as their positive antonyms, as statistically confirmed by a post-hoc pairwise comparison (with Tukey-adjusted p: $z = 3.34$, $p < 0.05$; cf. also similar asymmetric findings by Paradis and Willners (2006) on absolute adjectives with totally closed scales like *not empty*). This is primarily because of the presence of the aforesaid middle scores. Together with Gotzner and Kiziltan (2021), we take middle scores to reflect middling interpretations (e.g., 'neither clean nor dirty'). As said in Section 3.1, middling interpretations are not expected to arise with absolute adjectives given that their semantics does not allow for a middle ground. Then what do these interpretations of absolute adjectives amount to? We conjecture that a middling interpretation of absolute adjectives is only acceptable under the assumption that the context imposes a fine granularity (see discussion Section 2.1), where smaller distinctions of the property at stake (e.g., of dirt) are made relevant, and also makes distinctions

between alternative expressions relevant. Thus, a middle range of degrees is licensed, qualifying, e.g., as 'neither clean nor dirty', so as to express that the subject of predication is not prototypically/coarsely-speaking clean nor prototypically/coarsely-speaking dirty but presents at least some (possibly small) amount of dirt/non-cleanness (see similar point in Paradis and Willners, 2006 on the interpretation of *neither dead nor alive*). We will come back to this point in Section 4.1.

Let us now turn to negated negative strong terms (e.g., *not filthy*), which are rated in a similar way to their weak counterparts. While their middle-scale ratings indicate the type of middling interpretations discussed above, we hypothesize that ratings of 4 reflect inferences to the (weak) antonym. Such inferences do not arise via an entailment, as in the case of weak terms (*not dirty* $\Rightarrow$ 'clean'), nor via negative strengthening. They rather arise via an interaction of the minimum standard semantics and granularity, as explained in the following. In our experiment, negative polarity terms (e.g., *dirty, filthy*) are minimum standard adjectives. Based on their minimum standard semantics and a fine-grained scale (e.g., of dirt), one can draw more distinctions with respect to the property at stake (e.g., having a non-zero amount of dirt) than in the case of maximum standard adjectives (e.g., *clean*) that require reaching the maximum of a fine scale (see Sassoon and Zevakhina, 2012, and relevant discussion in Section 2.1). Thus, the scale range communicated by negated minimum standard adjectives given a fine granularity is largely restricted, making it very likely that the range conveyed by the weak term (*not dirty*) overlaps with that conveyed by its strong counterpart (*not filthy*).

Furthermore, although there is seemingly an overlap between the negated positive conditions (e.g., *not clean/not pristine*) too, we argue that the respective underlying reasoning may differ. We take low ratings of weak scalars like *not clean* to mainly reflect an entailment to the antonym (*not clean* $\Rightarrow$ 'dirty'), and similar ratings of strong terms (*not pristine*) to reflect inferences to the antonym possibly via negative strengthening (*not pristine* $\rightsquigarrow$ '(rather) dirty'). As is evident, these two conditions additionally receive ratings from the middle of the scale, suggesting that negated positive weak terms are not interpreted strictly semantically. This has also been found by Paradis and Willners (2006) for absolute terms with totally closed scales like *full* (see above) and is further confirmed by a post-hoc pairwise comparison between the negated weak positive condition and its non-negated antonymic counterpart condition ($z = 3.61$, $p < .01$, Tukey-adjusted).

Let us sum up our findings and compare them against the predicted interpretations of absolute adjectives discussed in Section 3.1 and outlined in Table 1. Our given task revealed that, in non-negated environments, participants make clear distinctions between weak and strong scalemates of both positive and negative adjectives, which they assign to different portions of the 5-point scale. By contrast, fewer distinctions between different terms are found under negation, with negated absolute terms sharing certain interpretations to some extent. Negated weak terms mostly receive semantic interpretations via the entailment to the corresponding antonyms, as predicted by the standard theory (e.g., Rotstein and Winter 2004; Kennedy and McNally 2005). Nevertheless, they are not interpreted strictly symmetrically to their non-negated antonyms, in line with Paradis and Willners's (2006) findings. Middling interpretations seem to be available with negated absolute terms, which we take to be due to the fine granularity and the relevance of the distinctions between alternative expressions that our task imposes. Granularity is further argued to be relevant for the inferences to the antonym triggered by negated negative strong

adjectives like *not filthy*. As to negative strengthening, this seems to be available with negated positive strong terms like *not pristine* (cf. discussion in Section 3.1). Lastly, we did not find any robust indication of negated stronger scale-mates triggering indirect scalar implicatures (e.g., *not pristine* $\nrightarrow$ 'clean' and *not filthy* $\nrightarrow$ 'dirty', respectively) in our task.

4. General discussion

4.1. Non-contradictory effect of negation and middling interpretations

Our findings indicate that, besides the typical contradictory effect of negation on absolute adjectives, whereby the negation of an absolute adjective entails its antonym, negation may bring about a less absolute effect on the interpretation of absolute adjectives. This is the so-called attenuating or mitigating effect of negation, mostly discussed in connection with negated relative adjectives (see Section 2.2), and also hinted at by Paradis and Willners (2006) in relation to the asymmetric interpretative patterns attested for their contradictory antonymic pairs under negation. We further speculated that this effect surfaces in our experiment mostly in the form of a middling interpretation with negated absolute adjectives and less as a negative strengthening interpretation with negated positive strong absolute adjectives like *not pristine*, if at all.

Middling interpretations of absolute adjectives become available by virtue of the fine granularity level our experimental setup imposes, which further makes distinctions between alternative adjective expressions relevant. That is, when speaking in more precise terms, the range of degrees that qualify, for instance, as 'neither clean nor dirty' corresponds to a small non-zero amount of dirt (see also related claim in Rotstein and Winter, 2004 that "in some contexts a moist towel may be deemed neither wet nor dry" and discussion in Section 2.1). Hence, the apparent aforesaid contradiction amounts to 'neither prototypically clean nor prototypically dirty'. Relatedly, Paradis and Willners (2006) claim that a felicitous gradable interpretation of 'neither dead nor alive' *could be 'almost dead' or 'half alive'* (p. 1053). Such an interpretation can be captured in terms of an interplay of Horn's Q/R-principles, also referred to as the "Division of pragmatic labor" (e.g., Horn, 1993) between two truth-conditionally equivalent expressions, like *clean* and *not dirty*:

..the more specialized form — briefer and/or more lexicalized — will tend to become **R**-associated with a particular unmarked, stereotypical meaning, use, or situation, while the use of the periphrastic or less lexicalized expression, typically (but not always) linguistically more complex or prolix, will tend to be **Q**-restricted to those situations outside the stereotype, for which the unmarked expression could not have been used appropriately. (p. 41)

In this light, simpler absolute adjective forms like *clean* are used to (R-)implicate a stereotypical situation like being prototypically or coarsely-speaking clean (note Sassoon and Zevakhina's (2012) relevant claim that approximate/coarse interpretations "are more probable than precise ones" (p. 232), which builds on Krifka's (2007a) work on approximate vs. precise uses of number words). Contrastingly, complex and marked absolute terms, such as the truth-conditionally equivalent *not dirty*, are used to (Q-)convey non-stereotypical meanings like 'neither prototypically dirty nor prototypically clean', which amounts to presenting a non-prototypical non-zero degree of dirt, e.g., when being slightly dirty. For example, imagine a T-shirt that has one

spot on it; it is too clean to put in the washing machine but also too dirty to put it back in the wardrobe (cf. example (5) in Section 2.1, from Rotstein and Winter, 2004, and also Burnett, 2014 for a discussion of whether absolute adjectives may have borderline cases).

Whether the above interpretation of *not dirty* is equivalent to the interpretation of *dirty* when modified by the attenuating degree modifier *slightly*, and, mutatis mutandis, for the interpretation of *not clean* and *almost clean*, is an empirical question that remains to be resolved. Paradis and Willners (2006) partly tackled this question by testing a limited set of heterogeneous data. Future research should investigate more systematically how unmodified/bare absolute adjectives are interpreted in comparison with their modified versions by degree adverbs that linguistically impose a fine granularity scale; put differently, we should consider to what extent negated minimum standard absolute adjectives are interpreted similarly to their modified version by *slightly* and/or negated maximum standard absolute adjectives are interpreted similarly to their modified version by *almost*. A further theoretical possibility to keep in mind is that the potential equivalence between the interpretations of negated bare absolute adjectives and their modified forms may be the result of a standard Quantity-based reasoning on the basis of a fine-grained scale like <…(not very dirty,) not dirty, not slightly dirty> (see also Breheny et al., 2018; Beltrama, 2022).

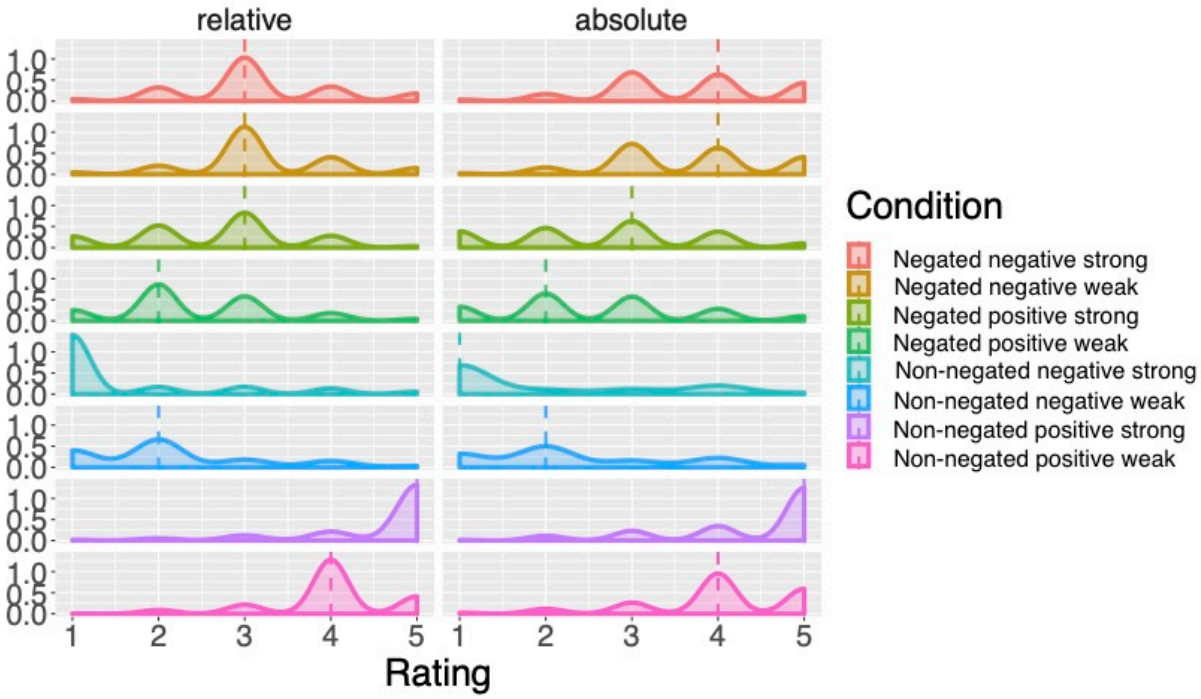
Another issue that remains to be addressed in an experimental setup concerns the nature and generalizability of the reasoning behind middling interpretations of absolute adjectives, as reflected by the middle ratings of the relevant negated conditions. Further research is necessary in order to find out to what extent this type of reasoning pertains to pragmatic principles that regulate communication (e.g., Q/R-principles or Quantity maxim, see above) and whether it hinges on the fine granularity level and the alternative adjective expressions introduced in our specific experimental setup. What would happen to middling interpretations of absolute adjectives when distinctions between alternative expressions are less prominent in the experimental setup?

In the following section, we compare our findings to those by Gotzner and Kiziltan (2021), who used the exact same experimental design and method to test for the interpretation of relative adjective expressions. Doing so will allow us to assess the similarities and differences in the interpretation of the two types of gradable adjectives, viz. of absolute and relative adjectives.

4.2. Relation to Gotzner and Kiziltan's (2021) findings on relative adjectives

Both Gotzner and Kiziltan's (2021) study on relative adjectives and ours on absolute adjectives make use of the same action-based rating task, where participants' responses are given on a Likert scale from 1 to 5. In such a setup (see more details in Section 3.2.2), participants in both experiments made use of distinct portions of the given scale when interpreting statements involving relative or absolute adjectives, especially in non-negated environments (see also Figure 2). That is, the two strong Polarity conditions received ratings from the corresponding endpoints of the scale, while the respective weak conditions were rated significantly differently from them. This is an indication that participants perceived the difference in informational strength between weak and strong expressions. However, under negation, such clear-cut distinctions are significantly fewer, as overlaps between conditions are observed.

Figure 2: Density curves across adjective conditions of the present study (absolute) and of Gotzner and Kiziltan's (2021) study (relative). Dashed lines represent medians.



In particular, Gotzner and Kiziltan (2021) find evidence of a polarity asymmetry in the interpretation of weak relative terms, such that negative strengthening interpretations arise for negated positive weak terms (e.g., *not large* $\rightsquigarrow$ '(rather) small'), but not for their negative counterparts (e.g., *not small*). The latter receive ratings from the middle of the scale, which is taken to indicate that they receive middling interpretations (e.g., *not small* $\rightsquigarrow$ 'neither large nor small'), as predicted by Horn (1989; see also Brown and Levinson, 1987; Levinson, 2000). The overlap with the two negated strong conditions evident in Figure 2 suggests that similar interpretations (e.g., 'neither large nor small') arise for expressions like *not gigantic* and *not tiny*.

In the present study, we, too, find an overlap between negated negative absolute conditions, which we attribute to the availability of a fine granularity level and the presentation mode of the alternative expressions, as concerns both peaks of the bimodal response pattern. Note that relative adjectives receive mostly middle scores, hence, middling interpretations, with the middle ground specified in their semantics being at the very root of such interpretations.

As to the negated positive conditions, in our experiment we did not find a difference between weak and strong terms (e.g., *not clean* / *not pristine*), like that attested for relative terms due to the availability of negative strengthening only for positive weak terms (e.g., *not large*). We speculated that our null finding is in keeping with the availability of seemingly similar interpretations in these two negated positive conditions: i.e., middling interpretations ('neither (prototypically) clean nor (prototypically) dirty') and inferences to the antonym, whereby the latter arise as entailments for weak positive terms (*not clean*), and possibly via negative strengthening for strong positive terms (*not pristine*).

To sum up, middling and negative strengthening interpretations appear to favor relative adjectives as compared to absolute adjectives. Weak absolute terms are mostly interpreted semantically under negation, while overall granularity interacts with the minimum/maximum standard semantics of absolute adjectives triggering additional inferences. Hence, this confirms our hypothesis that properties of the measurement scales underlying the semantics of gradable adjectives, such as endpoint-oriented standards or the lack thereof (scale structure), as well as the availability of a fine granularity level, affect differentially the derivation of inferences of gradable adjectives under negation.

4.3. Inference computation of gradable adjectives: Gotzner's (2021) *measurement mechanism*

The above conclusion comes to add to similar findings as regards different types of inferences exhibited differentially by bare and modified gradable adjectives (Gotzner et al., 2018b, a; Leffel et al., 2019). In this context, Gotzner (2021) proposes the so-called *measurement mechanism*, which aims to capture this interaction of properties of measurement scales of gradable adjectives with the computation of pragmatic inferences. The core idea of this mechanism of computing implicatures of gradable adjectives is that the relevant reasoning is not about lexically stored Horn scale alternatives, as in the case of quantifiers like <some, all>, but it is rather about positions on the underlying measurement scale. To illustrate this for scalar implicatures, rather than negating stronger/more informative lexical alternatives, this mechanism involves excluding all degrees that are ordered higher on the measurement scale than the resolved degrees (see Magri, 2017 for a similar account of Hirschberg scales).

5. Conclusion

While the semantics of gradable adjective classes, i.e., of relative and absolute adjectives, has been extensively studied, very little is known about the pragmatic interpretation of gradable adjectives. With a few very recent exceptions (Gotzner et al., 2018b, a; Leffel et al., 2019; Gotzner and Kiziltan, 2021; van Tiel and Pankratz, 2021), and extending on Gotzner and Kiziltan's (2021) study on relative adjectives, this study started out to investigate the pragmatic interpretation of absolute gradable adjectives. More precisely, we set out to probe the interpretation of weak absolute adjectives in relation to their stronger scale-mates and the corresponding antonyms when in the scope of negation or in the absence of negation.

We hypothesized that fewer distinctions are drawn between absolute terms differing in informational strength in negated than in non-negated environments, and that properties of the measurement scales absolute adjectives are associated with are important in understanding their behavior under negation. Indeed, we found that in the absence of negation participants distinguish between absolute terms differing in informational strength, while such distinctions are fewer in the scope of negation. In particular, under negation weak absolute adjectives are mainly interpreted semantically, while granularity appears to be responsible for additional interpretations of absolute adjectives, such as middling interpretations and inferences to the antonym. Through a comparison of our results to Gotzner and Kiziltan's (2021) findings, which completes the empirical picture of the interpretation of gradable adjectives overall, we firmly conclude that

scale structure, evaluative polarity, and granularity affect the derivation of different inferences of gradable adjectives. Our study comes to provide further support to the idea that properties of the measurement scales of gradable adjectives, such as the scale structure they use and scale granularity, play a central role in the derivation of pragmatic interpretations of gradable adjectives (Gotzner et al., 2018b, a; Leffel et al., 2019). This substantial finding has come to reduce the inscrutable variability as to the likelihood of pragmatic inferences observed across different scalar expressions (e.g., quantificational vs. adjectival; Doran et al., 2012; van Tiel et al., 2016). In general, the need for a model of scalar meaning that integrates multiple semantic and pragmatic factors in a systematic way is glaringly apparent. In search of such a model, we briefly discussed Gotzner's (2021) promising measurement mechanism that aims to account for the pragmatic interpretation of gradable adjectives.

References

Bartsch, R. and T. Vennemann (1973). *Semantic Structures: A Study in the Relation between Syntax and Semantics*. Frankfurt: Athenäum Verlag.

Beltrama, A. (2022). *Just* perfect, *simply* the best: An analysis of emphatic exclusion. *Linguistics and Philosophy 45*, 321–364.

Bierwisch, M. (1989). The semantics of gradation. In M. Bierwisch and E. Lang (Eds.), *Dimensional Adjectives*, pp. 71–261. Berlin: Springer Verlag.

Bolinger, D. (1972). *Degree Words*. The Hague: Mouton.

Breheny, R., N. Klinedinst, J. Romoli, and Y. Sudo (2018). The symmetry problem: Current theories and prospects. *Natural Language Semantics 26*(2), 85–110.

Brown, P. and S. C. Levinson (1987). *Politeness: Some Universals in Language Use*, Volume 4 of *Studies in Interactional Sociolinguistics*. Cambridge: Cambridge University Press.

Burnett, H. (2014). A delineation solution to the puzzles of absolute adjectives. *Linguistics and Philosophy 37*, 1–39.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax / pragmatics interface. In A. Belletti (Ed.), *Structures and Beyond*, pp. 39–103. Oxford: Oxford University Press.

Christensen, R. H. B. (2019). Regression models for ordinal data. R package version: 2019.12-10.

Colston, H. (1999). "Not good" is "bad", but "not bad" is not "good": An analysis of three accounts of negation asymmetry. *Discourse Processes 28*(3), 237–257.

Cremers, A. and E. Chemla (2014). Direct and indirect scalar implicatures share the same processing signature. In S. Reda (Ed.), *Pragmatics, Semantics and the Case of Scalar Implicatures*, pp. 201–227. London: Palgrave Macmillan.

Cresswell, M. J. (1976). The semantics of degree. In B. Partee (Ed.), *Montague Grammar*, pp. 261–292. New York: Academic Press.

Cruse, D. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Doran, R., G. Ward, Y. McNabb, M. Larson, and R. E. Baker (2012). A novel paradigm for distinguishing between what is said and what is implicated. *Language 88*, 124–154.

Ducrot, O. (1973). *Le Rôle de la Négation dans le Langage Ordinaire. La Preuve et le Dire*. Tours: Mame.

Fraenkel, T. and Y. Schul (2008). The meaning of negated adjectives. *Intercultural Pragmatics 5*(4), 517–540.

Gotzner, N. (2021). The use and interpretation of gradable adjectives. Invited talk at the Bochum Language Colloquium, Ruhr University Bochum, 28.1.2021.

Gotzner, N. and A. Benz (2018). The best response paradigm: A new approach to test implicatures of complex sentences. *Frontiers in Communication 2*, 21.

Gotzner, N. and S. Kiziltan (2021). She is brilliant! Distinguishing different readings of relative adjectives. In N. Gotzner and U. Sauerland (Eds.), *Measurement, Numerals and Scales: Essays in Honor of Stephanie Solt*, pp. 117–134. Cham: Palgrave Macmillan.

Gotzner, N. and D. Mazzarella (2021). Face management and negative strengthening: the role of power relations, social distance and gender. *Frontiers in Psychology 12: 602977*.

Gotzner, N. and J. Romoli (2018). The scalar inferences of strong scalar terms under negative quantifiers and constraints on the theory of alternatives. *Journal of Semantics 35*(1), 95–126.

Gotzner, N., S. Solt, and A. Benz (2018a). Adjectival scales and three types of implicature. In *Semantics and Linguistic Theory (SALT)*, Volume 28, pp. 409–432.

Gotzner, N., S. Solt, and A. Benz (2018b). Scalar diversity, negative strengthening, and adjectival semantics. *Frontiers in Psychology 9*, 1659.

Horn, L. (1993). Economy and redundancy in a dualistic model of natural language. In S. Shore and M. Vilkuna (Eds.), *Yearbook of the Linguistic Association of Finland*, pp. 33–72. SKY.

Horn, L. R. (1989). *A Natural History of Negation*. Chicago: University of Chicago Press.

Israel, M. (2004). The pragmatics of polarity. In L. R. Horn and G. Ward (Eds.), *The Handbook of Pragmatics*, pp. 701–723. Oxford: Blackwell.

Kennedy, C. (2001). Polar opposition and the ontology of 'degrees'. *Linguistics and Philosophy 24*, 33–70.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy 30*, 1–45.

Kennedy, C. and L. McNally (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language 81*, 345–381.

Krifka, M. (2007a). Approximate interpretation of number words: A case for strategic communication. In J. Zwarts, G. Bouma, and I. Krämer (Eds.), *Cognitive Foundations of Interpretation*, pp. 111–126. Amsterdam: Mouton de Gruyter.

Krifka, M. (2007b). Negated antonyms: Creating and filling the gap. In U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 163–177. Houndmill: Palgrave Macmillan.

Leffel, T., A. Cremers, N. Gotzner, and J. Romoli (2019). Vagueness in implicature: The case of modified adjectives. *Journal of Semantics 36*(2), 317–348.

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicatures*. Cambridge, MA: The MIT Press.

Magri, G. (2017). Blindness, Short-Sightedness, and Hirschberg's Contextually Ordered Alternatives: A Reply to Schlenker (2012). In S. Pistoia-Reda and F. Domaneschi (Eds.), *Linguistic and Psycholinguistic Approaches on Implicatures and Presuppositions*, Palgrave Studies in Pragmatics, Language and Cognition, pp. 9–54. Cham: Palgrave Macmillan.

Mazzarella, D. and N. Gotzner (2021). The polarity asymmetry of negative strengthening: dissociating adjectival polarity from face-threatening potential. *Glossa: A Journal of General Linguistics 6*(1), 47.

Paradis, C. and C. Willners (2006). Antonymy and negation—The boundedness hypothesis. *Journal of Pragmatics 38*(7), 1051–1080.

Rotstein, C. and Y. Winter (2004). Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics 12*, 259–288.

Ruytenbeek, N., S. Verheyen, and B. Spector (2017). Asymmetric inference towards the antonym: Experiments into the polarity and morphology of negated adjectives. *Glossa: A Journal of General Linguistics 2*(1), 1–27.

Sassoon, G. and A. Toledo (2015). Absolute and relative adjectives and their comparison classes. Unpublished ms., ILLC University of Amsterdam & Utrecht University.

Sassoon, G. and N. Zevakhina (2012). Granularity shifting: Experimental evidence from degree modifiers. In A. Chereches (Ed.), *Proceedings of SALT 22*, pp. 226–246.

Solstad, T. and R. Blutner (2000, 07). Dimensional designation: A case study in lexical pragmatics. In G. Jäger and R. Blutner (Eds.), *Studies in Optimality Theory*, pp. 30–40. Potsdam University Press.

Solt, S. (2015). Measurement scales in natural language. *Language and Linguistics Compass 9*(1), 14–32.

Tessler, M. H. and M. Franke (2019). Not unreasonable: Why two negatives don't make a positive. https://doi.org/10.31234/osf.io/tqjr2.

van Tiel, B. and E. Pankratz (2021). Adjectival polarity and the processing of scalar inferences. *Glossa: A Journal of General Linguistics 6*(1), 32.

van Tiel, B., E. van Miltenburg, N. Zevakhina, and B. Geurts (2016). Scalar diversity. *Journal of Semantics 33*(1), 107–135.