Not all gradable adjectives are vague – Experimental evidence from adults and children[1]

Merle WEICKER — *Goethe University Frankfurt*
Petra SCHULZ — *Goethe University Frankfurt*

Abstract. This paper investigates whether absolute gradable adjectives such as *clean* show characteristics of vagueness, as do relative gradable adjectives such as *big*. We addressed this question by examining adults' interpretation of the antonyms *clean* and *dirty* in a picture-selection task. Children between the ages of 3 and 5 were tested with the same task to see how early the interpretation pattern shown by the adults emerges in the course of acquisition. Our findings indicate that absolute gradable adjectives do not show characteristics of vagueness in either the adults or the children. Absolute gradable adjectives are interpreted as having a minimum or maximum standard of comparison, which is not context-sensitive, and they do not give rise to borderline cases.

Keywords: gradable adjective, vagueness, context-sensitivity, standard of comparison, borderline case.

1. Introduction

According to Kennedy (2007), a sentence such as (1) is vague because it has the following three characteristics: (i) contextually variable truth conditions, (ii) the existence of 'borderline cases', and (iii) trigger of the Sorites Paradox.

(1)   The cake is big.

It is widely agreed that the gradable adjective *big* is responsible for the uncertain interpretation of the sentence in (1). This interpretive uncertainty results from the observation that the interpretation of *big* is sensitive to the context (i). The sentence in (1) may be true in the context of cakes that a specific person, let's say, Merle, usually bakes, but false in the context of wedding cakes offered in a catalogue. The interpretive uncertainty of (1) also results from the existence of borderline cases (ii). These are entities, for which it is difficult to judge whether the sentence in (1) is true or false. Imagine a cake buffet at a birthday party. A cupcake would clearly not be judged big, a Black Forest cake would clearly be called big, but for a middle-sized cheesecake it would be more difficult to judge it as big or not big. It constitutes a so-called 'borderline case' for which it is unclear whether the predicate *is big* is true or false. Finally, vague predicates have been shown to give rise to the Sorites Paradox (iii), which is illustrated in (2). Although the two premises P1 and P2 seem true, the conclusion C is false.

(2)   P1. A 1-meter cake is big.

P2. Any cake that is 1 millimeter smaller than a big one is big.
C. Any 4-centimeter cake is big.

Kennedy (2007) argues that not all gradable adjectives show characteristics of vagueness: so-called absolute gradable adjectives (e.g., *full, dirty*) pattern differently from so-called relative gradable adjectives as *big* in (1). Other approaches, however, postulate a less sharp distinction between relative and absolute gradable adjectives (Burnett, 2012; Toledo and Sassoon, 2011).

The aim of the present study is therefore to examine whether adults treat absolute gradable adjectives as vague predicates, parallel to what has been found for relative gradable adjectives. By also testing children between the ages of 3 and 5, we examine how early the interpretation pattern shown by the adults emerges in the course of acquisition. The paper is organized as follows: the three characteristics of vagueness are explained in detail in Section 2 in the context of relative and absolute gradable adjectives. Previous empirical studies regarding adults' and children's interpretation of gradable adjectives are summarized in Section 3. Our study on the interpretation of absolute gradable adjectives in German is presented in Section 4. In Section 5, we discuss our findings.

2. Gradable adjectives and vagueness

2.1. Context-sensitivity

Gradable adjectives must be interpreted relative to a standard of comparison in order to evaluate whether the property introduced by the adjective holds of an individual. How this standard of comparison is determined differs between relative gradable adjectives (e.g., *big, tall, long*) and absolute gradable adjectives (e.g., *full, dirty, straight*): for relative gradable adjectives, the standard is contextually defined, whereas the standard for absolute gradable adjectives is less context-sensitive.

This analysis of gradable adjectives can be implemented in a degree-based framework. It is widely assumed that gradable adjectives denote relations between entities and degrees (Cresswell, 1976; von Stechow, 1984; Kennedy, 2007, among others). Degrees are understood as representations of measurement. A set of degrees that is totally ordered with respect to some dimension (e.g., height) is called a scale. Kennedy and McNally (2005) argue that gradable adjectives can differ regarding their scale structure. Adjectives can be associated with (i) totally closed scales, which have a minimal and a maximal endpoint (e.g., *full, empty*), with (ii) partially closed scales, which have either a minimal or a maximal endpoint (e.g., *clean, dirty; dangerous, safe*) or with (iii) totally open scales, which lack a minimal and a maximal endpoint (e.g., *big, small*). These differences in scale structure affect the interpretation of gradable adjectives. For gradable adjectives with open scales (referred to as relative gradable adjectives), the standard is context-dependent. In contrast, for gradable adjectives with (partially) closed scales (referred to as absolute gradable adjectives), the standard is less context-dependent: as a default, the standard corresponds to the minimal or maximal endpoint of the scale. Consequently, the adjective's argument must show either a maximal degree of the property (e.g., *clean, straight*) or a non-zero degree of the property (e.g., *dirty, bent*). Therefore, the

examples in (3) are contradictory (marked as #) with absolute gradable adjectives ((3a) and (3b)) but not with relative gradable adjectives ((3c) and (3d)) (Kennedy, 2007, p. 26).

(3)
     a.   # The rod is not bent, though there is a small bend in the middle.
     b.   # The line is STRAIGHT, but you can make it straighter.[2]
     c.   Sam is not tall, but his height is normal for his age.
     d.   That film is interesting, but it could be more interesting.

The difference between relative and absolute gradable adjectives regarding context-sensitivity is also mirrored by the relation of antonyms exemplified in (4). For relative gradable antonyms, the negation of one form does not entail the assertion of the other, that is, they are 'non-complementary' antonymous adjectives (Cruse, 1986).

(4)   The cake is not small. $\not\Rightarrow$ The cake is big.

Because the standard for relative gradable adjectives is context-dependent, the standard for *big* and *small* need not be the same degree (Kennedy, 2007; Kennedy and McNally, 2005). Therefore, an entity may exceed the standard for smallness without meeting the standard for bigness. Unlike relative gradable adjectives, absolute gradable adjectives license the inference 'not ADJECTIVE $\Rightarrow$ antonym of ADJECTIVE', as illustrated in (5); hence they are 'complementary' antonymous adjectives.

(5)   The plate is not dirty. $\Rightarrow$ The plate is clean.

According to Kennedy (2007), this follows because *dirty* imposes a minimum standard and *clean* imposes a maximum standard. Consequently, an entity without a minimal degree of dirtiness has a maximal degree of cleanliness.

Note that there are exceptions to the generalization that adjectives with closed scales have absolute standards that are the scales' minimal or maximal endpoint (Kennedy, 2007; Kennedy and McNally, 2005). These have led to the proposal of alternative approaches. Toledo and Sassoon (2011) argue that absolute gradable adjectives are interpreted relative to the context, just like relative gradable adjectives: the standard for *clean*, for instance, is lower for kitchen knives than for surgical instruments and the standard for *dirty* is lower for a child's shirt than for a tuxedo. Using the maximum absolute gradable adjective *full*, McNally (2011) further shows that its standard need not to be the maximal degree, for instance if wine glasses are considered. Example (6) by Rotstein and Winter (2004) shows that absolute gradable antonyms need not always be complementary. There seems to be a degree of dirtiness for which both *clean* and *dirty* would be judged false.

(6)   This glass is almost dirty. It is certainly not clean, since it has some small spots on it, but it is not really dirty, and I am willing to drink from it if you insist.

---

[2] The capitals signal focal stress, which forces a precise interpretation of the maximum gradable adjective. According to Kennedy (2007), this precise interpretation is necessary for this kind of entailment.

2.2. Borderline cases

Besides differences regarding context-sensitivity, the distinction between absolute and relative gradable adjectives is reflected in the existence/absence of borderline cases. Example (7) (taken from Kennedy, 2007, p. 24) is unnatural because the rod in question constitutes a borderline case: for this rod it is unclear whether the predicate *is long* applies.

(7)   ??We need a long rod for the antenna, but since long means 'greater than 10 meters' and this one is 1 millimeters short of 10 meters, unfortunately, it won't work.

In contrast, borderline cases do not exist for absolute gradable adjectives as the felicity of (8) illustrates (taken from Kennedy, 2007, p. 25).

(8)   The rod for the antenna needs to be straight, but this one has a 1 mm bend in the middle, so unfortunately it won't work.

Burnett (2012) calls the non-existence of borderline cases for absolute gradable adjectives into question, however. According to Burnett, absolute gradable adjectives such as *bald* have borderline cases: in addition to clear cases for bald and non-bald people, there may be less clear borderline cases as for instance people with a quarter head of hair.

2.3. Sorites Paradox

As shown in (2), relative gradable adjectives give rise to the Sorites Paradox: although the two premises are judged as true, the conclusion is false. Absolute gradable adjectives, in contrast, do not give rise to a paradoxical conclusion because the second premise is judged as false. This is illustrated in (9): P1 is true and P2 is false, hence the conclusion should be false. This is indeed the case.

(9)   P1. A plate without any spot of dirt is clean.
        P2. A plate with one spot of dirt is clean.
        C. A plate with very many spots of dirt is clean.

Again, it has been noted that this generalization may not hold for all absolute gradable adjectives. Once more, the adjective *bald* is a case in point, as exemplified in (10). It is questionable whether adding one hair to a bald person's head turns the person from bald to not bald; hence it is unclear whether the second premise is judged to be false. Therefore, the conclusion is paradoxical, similar to the example with the relative gradable adjective *big* (see example (2)).

(10)   P1. A person without any hair is bald.
        P2. A person with one hair is bald.
        C. A person with a head full of hair is bald.

2.4. Summary

Theoretical accounts of the semantics of gradable adjectives largely agree that relative gradable and absolute gradable adjectives exhibit a number of differences. Relative gradable adjectives show three central characteristics of vagueness: the interpretation of relative gradable adjectives is always context-sensitive, there is no context such that borderline cases disappear, and relative gradable adjectives always give rise to the Sorites Paradox. Absolute gradable adjectives show no characteristics of vagueness. However, as we saw in the preceding sections in some cases characteristics of vagueness are present in absolute gradable adjectives as well. Different proposals have been suggested to account for the differences and similarities between relative and absolute gradable adjectives. The approach by Kennedy (2007, see also Kennedy and McNally, 2005), for instance, assumes that the apparent similarities between relative and absolute gradable adjectives regarding the characteristics of vagueness in fact result from different phenomena. According to this analysis, vagueness (i.e., context-sensitivity, borderline cases, and the presence of the Sorites Paradox) is part of the semantics of relative gradable adjectives but not of absolute gradable adjectives. Absolute gradable adjectives can have imprecise uses, which is a pragmatic phenomenon, but they are not vague. That is, speakers can tolerate some deviation from the minimum or maximum standard in some contexts. An alternative approach by Burnett (2012) suggests that two characteristics of vagueness— borderline cases and the Sorites Paradox—arise through pragmatic processes that are the same for all gradable adjectives. The differences regarding context-sensitivity are assumed to be part of the semantics of relative and absolute gradable adjectives. In summary, whereas relative gradable adjectives are commonly characterized as vague predicates, the nature of absolute gradable adjectives is still debated. In particular, it is open whether—in line with relative gradable adjectives—they should be analyzed as (potentially) vague predicates. If vagueness is part of the semantics of relative but not of absolute gradable adjectives, participants' default interpretation in experimental settings should reflect the characteristics of vagueness for relative gradable but not for absolute gradable adjectives. Alternatively, if vagueness arises through pragmatic processes that are the same for relative and absolute gradable adjectives, participants' readiness for a vague interpretation should be similar for relative and absolute gradable adjectives given the appropriate context.

Acquisition data may help to shed light on the similarities and differences between different types of gradable adjectives. In our view, there are two basic alternative theoretical approaches: one could assume two different underlying meanings for relative and absolute gradable adjectives, which may result from different scale structures. In this case, children need to discover that absolute gradable adjectives, which have a precise meaning, can also be used imprecisely. Alternatively, relative and absolute gradable adjectives could be assumed to share the same—potentially vague—meaning. In this case, children need to discover that absolute but not relative gradable adjectives can also be used precisely. By looking at children's interpretation of gradable adjectives across different ages we can learn which interpretation learners choose initially. This in turn can help us to address the question of which of the two theoretical assumptions may be on the right track. If learners initially permit absolute gradable adjectives to have an imprecise reading, this finding would support the approach that all gradable adjectives are potentially vague. However, if learners initially do not permit absolute gradable adjectives to have an imprecise reading, this finding would support the view that absolute and relative gradable adjectives have different underlying meanings.

In what follows, we summarize the empirical findings of previous adult and child studies on the interpretation of gradable adjectives regarding context-sensitivity and borderline cases.

3. Empirical findings on the interpretation of gradable adjectives by adults and children

Let us first turn to the findings regarding the interpretation of relative gradable adjectives. Previous studies on adults' and children's interpretation of relative gradable adjectives support a semantic analysis as vague predicates. Findings by Syrett, Bradley, Kennedy and Lidz (2006) show that, just like adults, English-speaking children as young as age 3 are able to shift the standard of comparison for *big* and *long* depending on the context (see also Ebeling and Gelman, 1994; Sera and Smith, 1987; Syrett, Kennedy and Lidz, 2010). Moreover, by age 3, children locate the standard around the center of the corresponding scale (Syrett et al., 2006, for English; Foppolo and Panzeri, 2013, for Italian). Findings by Barner and Snedeker (2008) indicate that 4-year-old children interpret the antonyms *tall* and *short* as non-complementary. These findings all suggest a context-sensitive interpretation of relative gradable adjectives. Concerning the existence of borderline cases, it has been shown for adults that for some objects the predicate *is big* is judged as neither clearly true nor clearly false (Alxatib and Pelletier, 2011; Égré and Zehr, 2018; Solt and Gotzner, 2010). Findings by Weicker and Schulz (2020) indicate that this sensitivity to borderline cases emerges already as early as age 3.

Let us now turn to studies on the interpretation of absolute gradable adjectives, where theoretical approaches disagree on whether they should be analyzed as vague predicates (see Section 2). Syrett et al. (2006) and Syrett et al. (2010) used a 'Presupposition Assessment Task' to investigate whether participants shift the standard of comparison for *spotted, full, straight* and *bumpy* depending on the context. In this task, participants saw two objects and were for instance asked *Please give me the full one*. Because *full* has a maximum standard, participants should accept this request in a felicitous context with a maximally and a partially full container, but reject it in an infelicitous context with two partially full containers filled to different degrees. The adults performed as expected and rejected the request in the infelicitous context for all four absolute gradable adjectives. The 3- to 5-year-old children, in contrast, selected the fuller of the two partially full containers instead of rejecting the request. This non-target pattern was also found in the 3-year-olds' responses to the maximum absolute gradable adjective *straight*. Notably, the children rejected infelicitous requests containing the minimum absolute gradable adjectives *spotted* and *bumpy*. These findings for maximum absolute gradable adjectives suggest that up to age 5 children do not yet restrict context-sensitive interpretations to relative gradable adjectives.

To assess whether participants have an endpoint-standard for absolute gradable adjectives, Foppolo and Panzeri (2013, for Italian) and Syrett et al. (2006, for English) used a 'Scalar Judgement Task'. In this task, participants saw a series of objects that displayed the same property but to different extents. For each object a participant was asked 'Is this *ADJ*?', e.g., *Is this full?* That is, participants had to judge for each of the objects in the series whether the property introduced by an adjective is true. For absolute gradable adjectives with a minimum standard (e.g., *spotted*), a yes-response was expected for all objects except the object with a zero degree of the property, i.e., all objects with one spot or more. In contrast, for maximum absolute gradable adjectives (e.g., *full*) a yes-response was expected only for objects with a maximal degree of the property. For minimum absolute gradable adjectives adults and children

responded as expected. Unexpected results were found for *full*: 3- to 5-year-old English- and Italian speaking children—as well as some adults—judged containers that were not maximally full as *full*. This result points to the possibility of a non-endpoint standard for maximum absolute gradable adjectives.

The finding for *full* was replicated and extended to the maximum absolute gradable adjective *empty* in Foppolo and Panzeri's (2013) study. This study allows first insights into the question of whether maximum and minimum absolute gradable adjectives are interpreted as complementary because antonym pairs were included in their Scalar Judgement Task. The results for antonym pairs such as *clean* and *dirty* indicate a complementary interpretation: adults and children only judged objects without any dirt as *clean*. Conversely, all objects with some amount of dirt were judged as *dirty*. Note that participants seem to have received test sentences either with the positive or with the negative adjective. Accordingly, this study leaves open whether the interpretation of antonyms as complementary is maintained when participants are presented with both adjectives using the same series of objects. First results from 5-year-old German-speaking children and adults (Weicker and Schulz, 2018) address this question: for the same series of objects, participants selected only objects without any dirt as *clean* and conversely, all objects with some amount of dirt as *dirty*. In other words, there were no objects that satisfied neither *clean* nor *dirty*.

In short, previous empirical findings for adults and children support the theoretical analysis that relative gradable adjectives are vague predicates. Regarding the question of whether absolute gradable adjectives differ from relative gradable adjectives in terms of vagueness, the empirical evidence to date is mixed. The present study contributes to the empirical research on absolute gradable adjectives by using a novel experimental design for German.

4. Our study on absolute gradable adjectives

The present study examines whether German-speaking children and adults interpret absolute gradable adjectives as not vague. More specifically, we asked four research questions:

(Q1) Do adults and children have an endpoint-standard for absolute gradable adjectives?
(Q2) Do adults and children interpret absolute gradable antonyms as complementary?
(Q3) Do borderline cases exist for absolute gradable predicates?
(Q4) Do children's interpretation patterns change with age?

Children between ages 3 and 5 were tested. This is because children at age 3 are known to use adjectives from different semantic classes productively (Blackwell, 2005; Tribushinina et al., 2014; Weicker, 2019) and to be sensitive to the meaning differences between different adjective classes (Foppolo & Panzeri, 2013; Syrett et al., 2006; Syrett et al., 2010). At the same time, three-year-olds have been reported to have not yet mastered the meaning of gradable adjectives (see Section 3). Thus, investigating children's interpretation of gradable adjectives between the ages of 3 and 5 can uncover at which point in the acquisition process the adult interpretation pattern emerges and how children arrive at this interpretation.

4.1. Participants

Forty-three monolingual German-speaking children participated in the experiment: 11 three-year-olds (5 girls, 6 boys, age range = 3;2 to 3;11 years, mean age = 3;7 years), 15 four-year-olds (7 girls, 8 boys, age range = 4;1 to 4;11 years, mean age = 4;6 years) and 17 five-year-olds (9 girls, 8 boys, age range = 5;0 to 5;9 years, mean age = 5;4 years). All children in our study were typically-developing, as ensured via a standardized language test (SETK 3-5, Grimm, 2001). The children were tested at their day-care centers in the Frankfurt area. Additionally, 26 undergraduate students of Goethe University with little or no background in linguistics (22 female, 4 male), all native speakers of German, took part in the experiment. They received compensation for participation.

4.2. Method

4.2.1. Materials

Like Barner and Snedeker (2008) and Solt and Gotzner (2010) in their studies of relative gradable adjectives, we employed the method of picture selection. In our task participants saw eight picture cards (14x14 cm) simultaneously. On each picture card a single object, a ball or a teddy, was depicted. The objects differed in color and changed from dirty to clean (see Figure 1). Different from previous studies investigating the interpretation of absolute gradable adjectives (Foppolo and Panzeri, 2013; Syrett et al., 2006), the objects were presented in random fashion (see Figure 1). This way, participants had to make their picture selection based on their own ordering rather than from a given order. All test prompts had the form 'Please give me the $ADJ_{Plural} N_{Plural}$', e.g., *Gib mir bitte die dreckigen Bälle* 'Please give me the dirty balls'. The test prompts were uttered with a non-contrastive intonation. The adjectives tested were the maximum absolute gradable adjective *sauber* ('clean') and the minimum absolute gradable adjective *dreckig* ('dirty'); the nouns were *Teddy* ('teddy') and *Ball* ('ball'). The participants received a total of four test items, two trials per adjective.
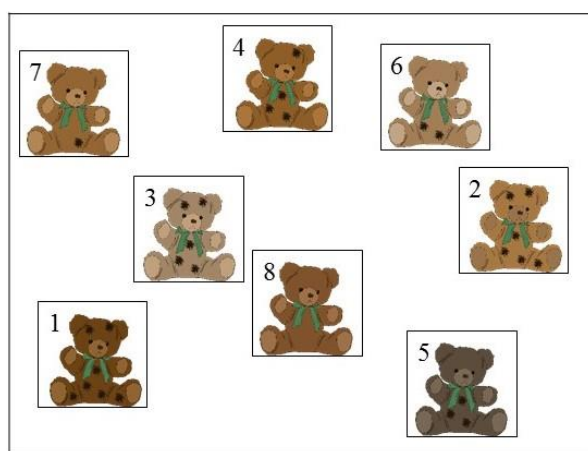

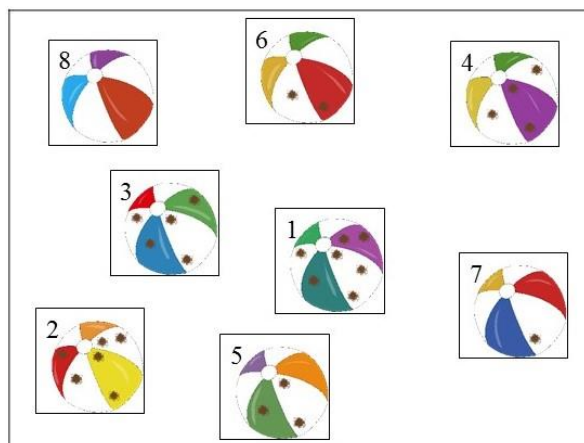Figure 1a. Visual array for the test prompt *Please give me the clean/dirty teddies*.

Figure 1b. Visual array for the test prompt *Please give me the clean/dirty balls*.
Figure 1. Test items.[3]

Crucially, the visual set up for *clean* and *dirty* was identical (see Figure 1). This way, it was possible to determine whether participants interpret these antonyms as complementary, i.e., as direct opposites. In this case, objects that are selected when asked for *clean balls/teddies* will not be selected when asked for *dirty balls/teddies* and *vice versa*. In other words, objects that were judged *clean* were judged *not dirty* and, conversely, objects that were judged *dirty* were judged *not clean* (see example (5)). Alternatively, if participants interpret antonyms as non-complementary, some objects will be neither selected when asked for *clean balls/teddies* nor when asked for *dirty balls/teddies*. That is, there are objects that count neither as *clean* nor as *dirty* (see example (6)).

In addition to the test trials, we added a total of eight filler items to the experiment for two reasons. First, the filler items increased the distance between test trials including the same adjectives to minimize potential influences from the prior presentation. Second, the filler items served to implement the overall set up of the experiment as a game, which requires taking turns. In the test trials a puppet, played by the experimenter, made a request and the participant had to select the matching objects; in the filler trials, the participant had to make a request and the puppet selected the matching picture cards (see Section 4.2.2. for details of the procedure). As illustrated in Figure 2, the filler trials also consisted of eight picture cards, but the picture cards in each trial showed toys from two different basic-level categories (book, bucket, dice, Lego® brick, soccer ball). In each trial, the objects differed in color and shape. The requests the participant was expected to make were similar to the test prompts. Due to the different objects and different properties (red, blue, round, square) displayed on the cards, the expected requests were of the form *Gib mir bitte die roten Legosteine* 'Please give me the red Lego bricks'.

---

[3] The numbers on the picture cards were not present in the actual experiment. They are added here for easier reference to the objects in Section 4.3. ('Results').
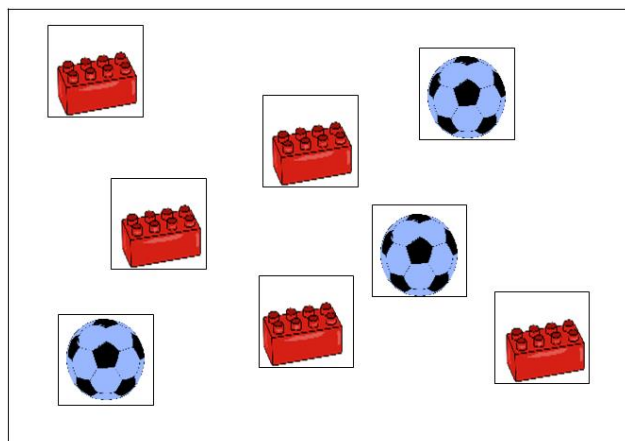
Figure 2. Example filler trial.

Before the test and the filler trials were administered, the participants were introduced to the objects and their names and were familiarized with the task. First, the experimenter showed a single exemplar of each object on a picture card and asked the participant to label them. Then participants received three practice trials, which did not contain any adjective (e.g., *Gib mir bitte die Puppen* 'Please give me the dolls'). Recall that for the test trials containing *clean*, only one object in the visual set up showed this property (i.e., is maximally clean) and for the test trials containing *dirty,* several objects showed that property (see Figure 1). Nevertheless, all requests were of the same form: they all contained a plural DP (e.g., *clean teddies*). Therefore, in the practice trials the number of objects showing a property (e.g., being a doll) varied; for one trial only one object fitted the description. This way, we illustrated that selecting only one object was a licit choice. For the practice trials participants received feedback: if they noticed that only one object matched the description, the experimenter emphasized that the puppet's request was the same independent of the number of matching objects. The experimenter also corrected the participant if she forgot to select one of the required objects.

4.2.2. Procedure

The testing was conducted in two separate sessions. In the first session, participants received the test trials containing the maximum absolute gradable adjective *clean*; in the second session, participants received the trials containing the minimum gradable adjective *dirty*. The visual displays and order of presentation were the same in both sessions. This way, we were able to compare the object choices for *clean* and *dirty* directly. To minimize possible influences from participants' choices in the first session to their choices in the second session, we ran the sessions about twelve days apart. The testing took place for all participants individually in a quiet room: for the children at their day-care centers and for the adults at their university. The participants and the experimenter sat next to each other on the floor or at a table large enough to display the eight picture cards per trial simultaneously.

At the beginning of each session the experimenter introduced the participant to a hand puppet and explained that the puppet wanted to play a game. The puppet and the participant each received a special dice that they had to use, keeping the outcome a secret. This set-up served to engage the children in a situation that naturally requires taking turns. The participants' dice

showed four options: square, circle, blue dot, red dot. If it was the participant's turn, the experimenter distributed the eight picture cards on the table and the participant rolled her dice. When the dice showed 'blue', for example, she had to ask the puppet to hand her the blue toys, etc. The participant's requests corresponded to the filler trials. If it was the puppet's turn, the experimenter distributed the picture cards on the table as well, and the puppet rolled her dice and made her request to the participant. The puppet's requests corresponded to the test trials. The participant's task was to select those objects that in her opinion matched the test prompt.

4.2.3. Data analysis

The analysis is based on the participants' object choices when asked to pick the clean or dirty objects, respectively. We analyzed the data in two ways. First, for the group analysis we calculated the percentage of selections for each of the eight objects across the two trials per adjective. Second, for each participant and for each of the four test trials we defined the cut-off point for *clean* and *dirty*, i.e., the dirtiest object selected as *clean* (together with all cleaner objects) and the cleanest object selected as *dirty* (together with all dirtier objects). The group and the individual analyses allow us determine the standard of comparison for *clean* and *dirty*. By directly comparing participants' cut-off points for both antonyms we are able to assess whether they were complementary or non-complementary. From this relation we can infer whether borderline cases exist for *clean* and *dirty*.

4.3. Results

To find out whether children and adults have an endpoint-standard for absolute gradable adjectives (Q1), we calculated how often participants selected the respective objects 1 to 8 (see Figure 1a and b). The results for *clean* are given in Figure 3 and the results for *dirty* are given in Figure 4. For each of the objects, 22 responses per adjective were analyzed in the group of 3-year-olds (2 trials x 11 participants), 30 responses per adjective in the group of 4-year-olds (2 trials x 15 participants) and 34 responses per adjective in the group of 5-year-olds (2 trials x 17 participants). A total of 52 responses per adjective were analyzed in the adult group (2 trials x 26 participants). In Figure 3 and Figure 4, the eight objects are plotted by number on the x-axis (1 = dirtiest object, 8 = cleanest object). Note that they are ordered linearly only for the ease of illustration. The y-axis displays how often the respective objects were selected. The raw numbers are given in the Appendix.

When asked for *clean*, the adults only selected object 8, which is the maximally clean object, i.e., the object that was without any dirt. The choice patterns for the three child groups is similar to the adults'. However, all child groups also sometimes selected object 7, which had one spot of dirt on it, as clean. The number of choices for object 7 decreased with age. Note that the patterns for the 3-year-olds and for the 4-year-olds differed slightly from the 5-year-olds: in total, in 7.5% of the *clean*-trials all objects were selected; this was due to three children.[4]

---

[4] It is unlikely that these three children simply did not understand the task, because they did not select all objects across all trials. Therefore, we did not remove these children from the analysis. For the question of why some children responded this way, see the discussion in Weicker (2019).

Although the patterns differed slightly between the groups, the median[5] cut-off point for *clean* was identical across age groups and trials, namely 8.0. That is, the distribution of cut-off points did not differ significantly across age groups, *clean teddies*: $X^2(3) = 5.51$, $p = .138$; *clean balls*: $X^2(3) = 7.64$, $p = .054$. Importantly, the standard of comparison was the maximal endpoint of the scale.
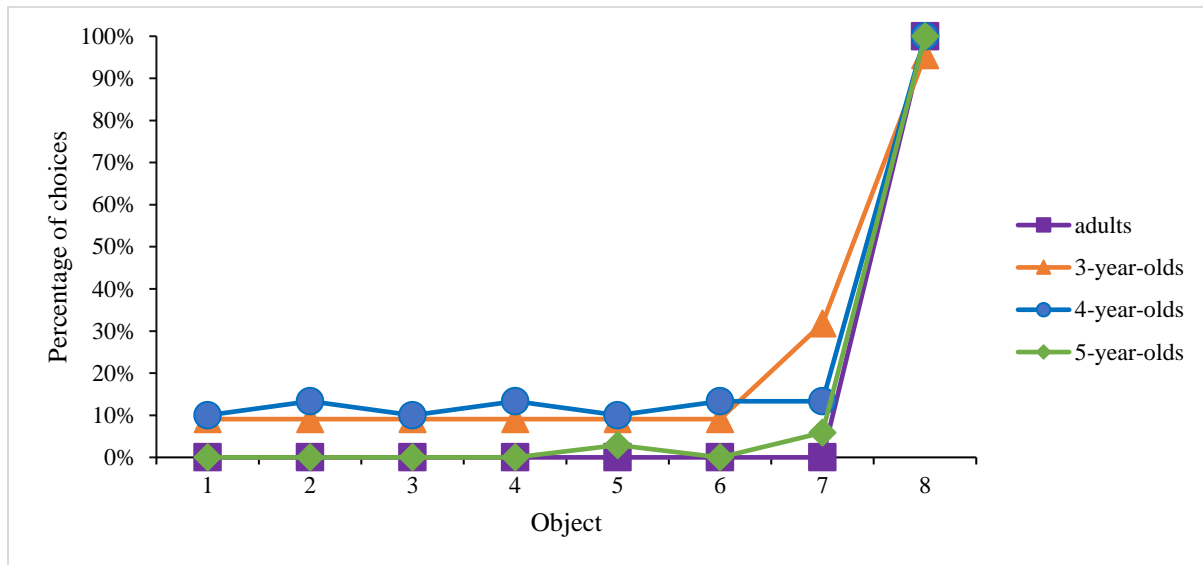
Figure 3. Percentage of choices for *clean*-trials per age group. Number 1 displays the dirtiest object, number 8 the cleanest object.

The adults' choices for *dirty* show that every object that contained some amount of dirt (objects 1-7) was selected as dirty. The clean object (object 8) was never selected. Again, the children's choices differed slightly because either not all dirty objects were selected or the clean object was selected as well. The 3-year-olds selected the dirty objects less often: objects that were only slightly dirty were sometimes not considered *dirty*. However, as for *clean*, the median cut-off point for *dirty* was identical across age groups and trials, namely 7.0. That is, the distribution of cut-off points did not differ significantly across age groups, *dirty teddies*: $X^2(3) = 3.47$, $p = .325$; *dirty balls*: $X^2(3) = 2.84$, $p = .417$. Here, the standard of comparison was the minimal endpoint of the scale. As for (Q1) we can conclude that adults and children have an endpoint-standard for absolute gradable adjectives.

---

[5] The median was chosen as the measure of central tendency because for the present purposes it is more suitable than the arithmetic mean. When talking about objects a median cut-off point of 8.0, for instance, is more meaningful than a mean cut-off point of 7.9. In addition, the median is the 'typical' cut-off point in the distribution because it is less affected by extreme scores than the arithmetic mean (Field, 2009), e.g., when children selected all objects.
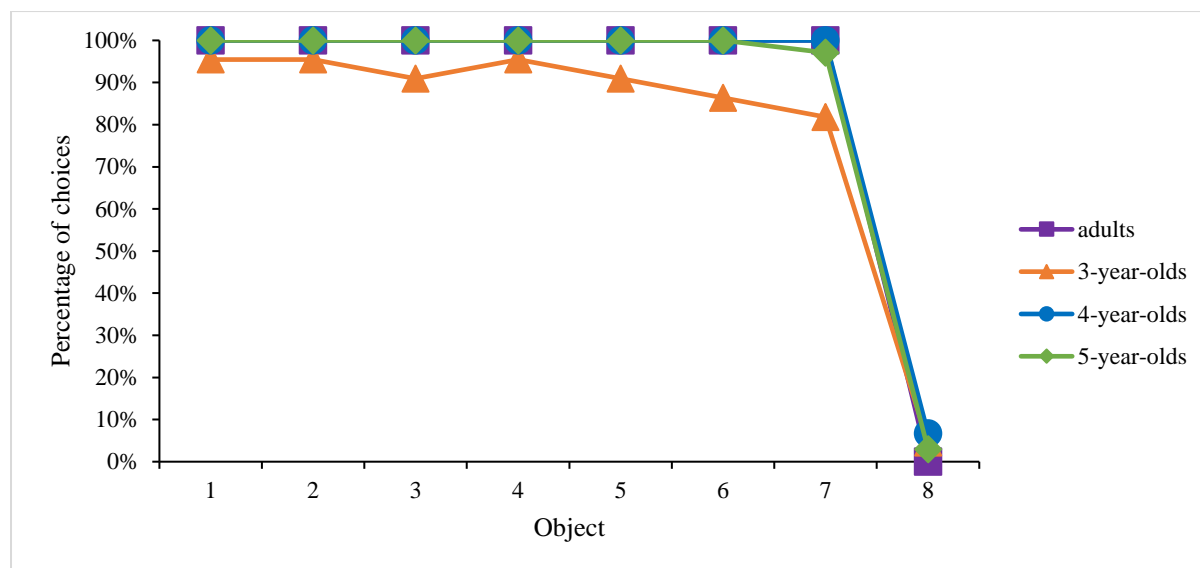
Figure 4. Percentage of choices for *dirty*-trials per age group. Number 1 displays the dirtiest object, number 8 the cleanest object.

Comparing the cut-off points for *clean* and *dirty* based on the group data indicates that adults and children interpret absolute gradable adjectives as complementary (Q2): participants' object selections for *clean* are the mirror image of their object selections for *dirty* and *vice versa*. Further evidence for this interpretation is provided by the individual analysis. For the majority of participants (59 out of 69, 86%), objects are either *clean* or *dirty* across trials. Among these are 6 of the 11 3-year-olds, 13 of the 15 4-year-olds, 14 of the 17 5-year-olds, and all 26 adults. Put differently, these 59 participants judged every object that they did not judge *dirty* as *clean* and *vice versa* (see example (5)). Interestingly, only one 3-year-old child considered some objects as neither *clean* nor *dirty,* which indicates a non-complementary interpretation. How can we classify the remaining nine participants that show neither a non-complementary nor a consistent complementary interpretation of absolute gradable adjectives? Four of them (one 3-year-old and three 5-year-olds) had an inconsistent complementary interpretation, i.e., only for one visual set up (e.g., only for clean and dirty balls). For one 3-year-old and one 5-year-old with an inconsistent complementary interpretation and one additional 3-year-old the object with one spot of dirt was considered both dirty and clean. Possibly, children selected this object when asked for *clean* and when asked for *dirty* because they judge it neither clearly dirty nor clearly clean. Hence, this specific object may constitute a borderline case, which suggests that three participants assume that borderline cases exist for absolute gradable adjectives (Q3). The responses of four children were unanalyzable. As for (Q4), interpretation patterns changed only slightly across age, as demonstrated by the fact that the cut off-point was the same for all age groups across trials.

5. Discussion and conclusion

The present study used a picture-selection task that was implemented as a game. The objects on the pictures all belong to the same superordinate-level category, namely toys. Participants had to make their choices based on their own ordering of the objects presented. Initial practice trials served to illustrate that the number of objects matching the request could vary. Accordingly, in the test trials participants could in principle select as many or as few objects

matching their interpretation of the test prompt as they thought appropriate. This way, participants' interpretation of absolute gradable adjectives could be inferred from their object choices in a natural setting. Our findings suggest that adults interpret the absolute gradable adjectives *clean* and *dirty* as not vague – at least in the contexts we created. Children as young as age 3 share this interpretation overall. (Q1) asked whether participants have an endpoint-standard for absolute gradable adjectives. The data of the adult group revealed that the standard of comparison for *clean* and *dirty* is the maximal and minimal endpoint of the scale, respectively. This data confirms the judgements in examples (3a) and (3b): if something has some dirt on it, it must be dirty; if something is clean, it cannot be cleaner. The data of the 3-, 4-, and 5-year-old groups show a very similar interpretation pattern. (Q2) asked whether participants interpret absolute gradable adjectives as complementary. The group data as well as the individual data of adults and children support the interpretation of absolute gradable adjectives as complementary, as illustrated in example (5). A comparison of the object selections for *clean* and *dirty* showed that the majority of adult and child participants considered objects either *clean* or *dirty*. Put differently, for them every object that is clean is not dirty and conversely, every object that is dirty is not clean. Hence, there were no objects that were considered neither clean nor dirty, which is in contrast to non-complementary adjectives such as *big* (see example (4)). Given this finding, we conclude that borderline cases do not seem to exist for absolute gradable adjectives (Q3). (Q4) addressed the question of whether interpretation patterns change with age. We found that the adult interpretation pattern is already present in 3-year-old children and that across age interpretation patterns change only slightly. For instance, a few children did not interpret absolute gradable adjectives as complementary across the visual contexts; they sometimes had a non-maximal standard for *clean* or a non-minimal standard for *dirty,* or they selected the object with one spot of dirt both when asked for *dirty* and when asked for *clean.*

The picture-selection task used in the present study allowed us to extrapolate participants' interpretation of absolute gradable adjectives from their object choices. To probe whether children's rare deviations from the adult pattern indeed reflect a non-adult-like meaning of absolute gradable adjectives, future studies could include truth-value judgements, which require explicit decisions on the part of the participants (for this method in adult studies: see Alxatib and Pelletier, 2011; Égré and Zehr, 2018). Truth-value judgements could assess directly whether children accept or reject descriptions containing an absolute gradable adjective and its negation for specific objects (e.g., *This teddy is dirty*, *This teddy is clean*, *This teddy is not dirty*, *This teddy is dirty and clean* or *This teddy is neither dirty nor clean* as descriptions of object 7 in Figure 1).

In summary, our results from a picture-selection task in German with 26 adults and 43 three- to five-year-old children provide first empirical evidence that not all gradable adjectives are interpreted as vague. For the contexts we created, adults consistently opted for a non-vague, i.e., precise, interpretation of absolute gradable adjectives: absolute gradable adjectives did not have context-sensitive interpretations and did not give rise to borderline cases. What is more, the child data revealed that this interpretation of absolute gradable adjectives as not vague is already present by age 3. To evaluate the robustness of this pattern, future studies should vary the contexts in a way that makes the imprecise reading of absolute gradable adjectives more readily available.

We conclude by suggesting that theories on vagueness and gradable adjectives can benefit from taking into account the finding that young language learners are sensitive to the differences between absolute and relative gradable adjectives regarding vagueness. Two questions remain open. First, it is unclear whether children arrive at the non-vague interpretation of absolute gradable adjectives because they are sensitive to semantic differences between absolute and relative gradable adjectives or because they have not yet acquired the pragmatic mechanisms necessary to interpret absolute gradable adjectives as vague. Second, it is unclear in which contexts both adults and children may opt for a vague reading of absolute gradable adjectives.

Appendix

Table 1. Number of choices for each object in *clean*-trials per age group.

Age	Object							
	1	2	3	4	5	6	7	8
3	2	2	2	2	2	2	7	21
4	3	4	3	4	3	4	4	30
5	0	0	0	0	1	0	2	34
Adults	0	0	0	0	0	0	0	52

Table 2. Number of choices for each object in *dirty*-trials per age group.

Age	Object							
	1	2	3	4	5	6	7	8
3	21	21	20	21	20	19	18	1
4	30	30	30	30	30	30	30	2
5	34	34	34	34	34	34	34	1
Adults	52	52	52	52	52	52	52	0

References

Alxatib, S. and F. J. Pelletier (2011). The psychology of vagueness: borderline cases and contradictions. *Mind & Language, 26*(3), 287–326.

Barner, D. and J. Snedeker (2008). Compositionality and statistics in adjective acquisition: 4-year-olds interpret 'tall' and 'short' based on the size distributions of novel noun referents. *Child Development, 79*(3), 594-608.

Blackwell, A. (2005). Acquiring the English adjective lexicon: relationships with input properties and adjectival semantic typology. *Journal of Child Language, 32*(3), 535-562.

Burnett, H. (2012). The puzzle(s) of absolute adjectives. On vagueness, comparison, and the origin of scale structure. In D. Paperno (Ed.), *UCLA Working Papers in Linguistics, Papers in Semantics Vol. 16*, p. 1-50.

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

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

Ebeling, K. and S. Gelman (1994). Children's use of context in interpreting 'big' and 'little'. *Child Development, 65*(4), 1178-1192.

Égré, P. and J. Zehr (2018). Are gaps preferred to gluts? A closer look at borderline contradictions. In E. Castroviejo, L. McNally, and G. Weidman Sassoon (Eds.), *The Semantics of Gradability, Vagueness, and Scale Structure*, pp. 25-58. Berlin: Springer.

Field, A. (2009). *Discovering statistics using SPSS*. London: SAGE.

Foppolo, F. and F. Panzeri (2013). Do children know when their room counts as clean? In S. Kan, C. Moore Cantwell, and R. Staubs (Eds.), *NELS 40: Proceedings of the 40th Annual Meeting of the North East Linguistic Society: Volume One,* p. 205-218. Amherst, MA: GLSA.

Grimm, H. (2001). *SETK 3-5. Sprachentwicklungstest für 3- bis 5-jährige Kinder*. Göttingen: Hogrefe.

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

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

McNally, L. (2011). The relative role of property type and scale structure in explaining the behavior of gradable adjectives. In R. Nouwen, R. van Rooij, U. Sauerland, and H.-C. Schmitz (Eds.), *Vagueness in Communication*, p. 151-168. Berlin: Springer.

Rotstein, C. and Y. Winter (2004). Total adjectives vs. partial adjectives: scale structure and higher-order modifiers. *Natural Language Semantics*, *12*(3), 259-288.

Sera, M. and L. B. Smith (1987). Big and little: 'nominal' and relative uses. *Cognitive Development, 2*, 89-111.

Solt, S. and N. Gotzner (2010). *Expensive, not expensive or cheap? An experimental investigation of vague predicates*. Paper presented at the 11th Szklarska Poreba Workshop on Experimental Pragmasemantics, Szklarska Poreba.

Syrett, K., Kennedy, C., and J. Lidz (2010). Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics, 27*, 1-35.

Syrett, K., Bradley, E., Kennedy, C., and J. Lidz (2006). Shifting standards: children's understanding of gradable adjectives. In K. Ud Deen, J. Nomura, B. Schulz, and B. D. Schwartz (Eds.), *Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition – North America, Honolulu, HI, Volume 2*, pp. 353-364. Cambridge, MA: UConn Occasional Papers in Linguistics 4.

Toledo, A. and G. W. Sassoon (2011). Absolute vs. relative adjectives – Variance within vs. between individuals. In N. Ashton, A. Chereches, and D. Lutz. (Eds.), *Proceedings of SALT 21*, pp. 135-154. Linguistic Society of America.

Tribushinina, E., van den Bergh, H., Ravid, D., Aksu-Koç, A., Kilani-Schoch, M., Korecky-Kröll, K., Leibovitch-Cohen, I., Laaha, S., Nir, B., Dressler, W., and S. Gillis (2014). Development of adjective frequencies across semantic classes. *Language, Interaction and Acquisition, 5*(2), 185-226.

von Stechow, A. (1984). Comparing semantic theories of comparison. *Journal of Semantics, 3*, 1-77.

Weicker, M. (2019). *The role of semantic complexity for the acquisition of adjectives* (Doctoral dissertation, Goethe University Frankfurt, Germany). Retrieved from http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/50979.

Weicker, M. and P. Schulz (2020). Not everything needs to be big or small: evidence from children's interpretation of vague adjectives. In M. M. Brown and A. Kohut (Eds.), *Proceedings of the 44th Annual Boston University Conference on Language Development,* pp. 724-737. Somerville, MA: Cascadilla Press.

Weicker, M. and P. Schulz (2018). Is 'clean' the same as 'not dirty'? On the understanding of absolute gradable adjectives. In A. B. Bertolini and M. J. Kaplan (Eds.), *Proceedings of the*

42th Annual Boston University Conference on Language Development, pp. 790-802. Somerville, MA: Cascadilla Press.