Is clean the Same as not dirty? On the Understanding of Absolute Gradable Adjectives

Merle Weicker and Petra Schulz

1. Theoretical background

In order to evaluate whether the property denoted by the adjective holds of an individual, the speaker has to interpret gradable adjectives relative to a standard of comparison. The standard of comparison is determined depending on the adjective’s scale type\(^1\) (Kennedy & McNally 2005, Kennedy 2007). For adjectives that correspond to open scales such as big or small (referred to as relative gradable adjectives), the standard is context-dependent. Here the standard of comparison depends on the comparison class the adjective refers to and is located around the midpoint of the scale. In contrast, for adjectives related to (partially) closed scales such as clean or dirty (referred to as absolute gradable adjectives), the standard is not context-dependent. Consequently, the adjective’s argument must either show a maximal degree of the property (clean) or a non-zero degree of the property (dirty) resulting in a standard at the endpoint of the scale. This is illustrated by Kennedy’s (2007) examples in (1), which are contradictory (marked as \(^\#\)) for absolute gradable adjectives as in (1a) and (1b), but not for relative gradable adjectives as in (1c) and (1d).

(1) 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.

\(^1\) For a detailed discussion of alternative approaches regarding gradable adjectives and the absolute-relative distinction see Weicker (in prep.).

\(^2\) The capitals signal focal stress. According to Kennedy (2007), this forces a precise interpretation of the maximal gradable adjective, which is necessary for this kind of entailment. For a discussion on imprecise meanings of maximal absolute gradable adjectives see Kennedy (2007).
According to Kennedy (2007), sentences with gradable adjectives are vague due to the adjectives’ context-sensitivity. In contrast to relative gradable adjectives that are context-sensitive, absolute gradable adjectives are not context-sensitive and thus do not give rise to vagueness. A second characteristic of vague sentences are so-called ‘borderline-cases’. Take for instance the relative gradable adjective *big*; in the context displayed in Figure 1 there may be a set of objects that can be clearly judged as big (e.g., water balloons 6, 7, and 8) and a set of objects that can be judged as not big (e.g., water balloons 1, 2, and 3). However, there may be also objects that are more difficult to judge as big or not big (e.g., water balloons 4 and 5).

![Image of water balloons](image)

**Figure 1. Example context for so-called borderline cases.**

As a result, the negation of one form does not entail the assertion of the other in antonym pairs as illustrated in (2):

(2) The door is not large. $\not\Rightarrow$ The door is small.

Following Kennedy (2007), the inference in (2) results from the context-sensitive standards of relative gradable adjectives: the standard for *big* and *small* need not to be the same degree (cf. also Solt 2011). To account for the intuitive ‘gap’ between positive and negative antonyms such as in example (2) and in sentences like *Fred isn’t tall, but he’s not short either* Solt (2011) argues that the standard of comparison should be defined as a range of degrees rather than as a single degree.
Unlike in relative gradable adjectives, borderline-cases do not exist for absolute gradable adjectives. Consequently, the negation of one form entails the assertion of the other in antonym pairs. Put differently, absolute gradable adjectives license these inferences, as illustrated in (3):

(3) The door is not open. ⇒ The door is closed.

Therefore, absolute gradable adjectives are expected to exhibit a standard of comparison that is a single degree on the scale shared by the positive and the negative antonym. However, Rotstein & Winter (2004) illustrate that not all absolute gradable adjectives must be interpreted as complementary. They mention for instance the antonym pair clean/dirty as illustrated in (4):

(4) 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.

There seems to be a degree of dirtiness for which both clean and dirty would be judged false. Nevertheless, Rotstein & Winter (2004) argue that the default interpretation of absolute gradable antonyms is complementary.

In summary, theoretical accounts of the semantics of gradable adjectives argue that relative gradable adjectives and absolute gradable adjectives differ in that the former but not the latter show the characteristics of vague predicates. Absolute and relative gradable adjectives both denote relations between individuals and degrees; they differ with respect to their standard of comparison and regarding their context-sensitivity.

2. Previous acquisition findings on gradable adjectives

Previous research on relative gradable adjectives by Barner & Snedeker (2008) indicates that English-speaking children as young as age 4 - just like adults - have different standards of comparison for positive and negative antonyms. They tested the adjectives tall vs. short. In their study, participants saw eight unknown objects called pimwits ranging in height from 1 to 8 inches and were asked to determine which pimwits were tall and which were short. The average minimum height of objects called tall and the average maximum height of objects called short differed significantly in both children and adults. Similarly, in a recent study by Tribushinina (2013), participants saw seven known objects, e.g., elephants, differing in size. In one trial they were asked which objects they find small, and in another trial which ones they find big. In this study, 4-year-old Dutch-speaking children had a cut-off point exactly in the middle of the scale, i.e., it was the same for big and small. From age 5, children showed more asymmetric cut-off points for big and small. These findings point to a gap between tall and short or big and small objects, which is expected for vague predicates (cf. Kennedy 2007). Negative and positive antonyms of relative
gradable adjectives evoke different standards of comparison, i.e., there are objects which are judged neither tall nor short for instance.

Little research has been conducted on absolute gradable adjectives (Syrett 2007, Syrett et al. 2006, Syrett et al. 2010 for English, Foppolo & Panzeri 2013 for Italian). Syrett (2007) and Syrett et al. (2006) used two different tasks to investigate children’s interpretation of relative and absolute gradable adjectives (big, long, spotted, full). In the Scalar Judgement Task 3- to 5-year-old children saw a series of seven objects that displayed the same property but to different extents. Participants had to judge whether the property denoted by the adjective in question is true of each of the objects in the series. For each object a child was asked ‘Is this ADJ’, e.g., ‘Is this full?’ Regarding the absolute gradable adjective with a maximal standard full, a “yes”-response was expected for the object with the maximal degree of the property (i.e. being full). For spotted, an absolute gradable adjective with a minimal standard, a “yes”-response was expected for all objects with a non-zero degree of the property (i.e. having at least one spot). Regarding full, contrary to prediction children accepted containers as full that were not maximally full, which suggests a non-endpoint standard. This finding was replicated in the second task that used Presupposition Assessment. In this task children saw two containers, which were filled to different degrees, but were not maximally full. When asked Please give me the full one some of the children selected the fuller container instead of rejecting the request. These findings across both tasks indicate that adjectives like full may receive a vague interpretation. Syrett et al. (2010) argue against a vague interpretation of full. They claim that the apparent context-sensitivity of absolute gradable adjectives does not result from vagueness, as it does in relative gradable adjectives; it is not part of the adjective’s meaning. Instead, absolute gradable adjectives can be used in a context-sensitive way, which results in imprecise meanings. As Syrett (2007) and Syrett et al. (2006) did not investigate antonym pairs, in their studies it remains open whether children and adults show a gap for absolute gradable antonym pairs.

First insights regarding the latter point come from Foppolo & Panzeri’s (2013) study. Extending the Scalar Judgement Task to antonym pairs such as full/empty and clean/dirty, their results regarding full and empty are comparable to Syrett et al.’s (2006). For clean and dirty, however, children showed more categorical judgements: only objects without any dirt were judged as clean, and conversely, all objects with some amount of dirt on them were judged as dirty. The latter results for clean and dirty suggest that there is no gap for absolute gradable adjectives that refer to partially closed scales. However, as participants seem to have been presented with either the negative or the positive antonym, it is not possible to compare their individual cut-off points for negative and positive adjectives. It remains open whether the judgements for clean and dirty are maintained, when both adjectives are presented together with the same visual array.

In short, previous empirical studies have shown that children between ages 4 and 5 already know that relative gradable adjectives give rise to vague interpretations. Regarding the question of whether children have an endpoint-
standard or a midpoint-standard for absolute gradable adjectives, the empirical evidence is mixed. Moreover, it is open whether children treat absolute gradable adjectives as vague or not with respect to for instance borderline-cases. That is, do children interpret antonymous absolute gradable adjectives as complementary or are there objects of which neither the positive nor the negative adjective is true? The aim of the present study is to extend the previous findings from English and Italian by using a modified experimental design for German (cf. Weicker, in prep. for an extension to the acquisition of relative gradable (big, small) and absolute gradable adjectives (clean, dirty)).

3. Study

The present study investigates whether monolingual German-speaking children and adults interpret absolute gradable adjectives as vague. If they take absolute gradable adjectives to be not vague, they should exhibit an endpoint-standard for them. Moreover, they should evaluate absolute gradable antonym pairs according to the same standard of comparison, i.e., objects in a given context should be assigned either to the positive or the negative antonym with no undefinable objects left.

3.1. Participants

Seventeen 5-year-old children (9 girls, 8 boys, age range: 5;0 to 5;9 years, mean age: 5;4 years) were tested. All children were typically-developing monolingual speakers attested via a standardized test (SETK 3-5, Grimm 2001). The children were tested at their day-care centers in the Frankfurt area. Twenty-six adult native speakers of German (22 female, 4 male) served as controls. The adults were undergraduate students of Goethe University Frankfurt with little or no linguistic background. They were compensated for participation.

3.2. Method

3.2.1. Materials

In a forced-choice picture task, which took Barner & Snedeker’s (2008) design as a starting point, participants saw eight picture cards simultaneously in unordered fashion (see Figure 1). In contrast to previous studies (Syrett 2007, Syrett et al. 2006, Foppolo & Panzeri 2013), the objects were shown in random fashion to invite participants to establish their own ordering. This way we could prevent participants from inferring their judgements from the given order.

In contrast to previous studies, participants were not asked to attest a respective property for every object individually (‘Is this ADJ?’). Instead the test prompts of the present study were always of the form ‘Please give me the $ADJ_{plural}$ $N_{plural}$’, e.g., Gib mir bitte die sauberen Teddys ‘Please give me the clean teddies’ (uttered with a neutral, i.e., non-contrastive, intonation). We used this kind of test
prompt for several reasons. First, the trials described here were part of a larger study. For other parts of this study the noun played a role as well and this way of asking yielded a consistent format of the test prompts. Second, pilot studies revealed that asking participants to attest a respective property for every object individually was tiring, given the fact that other trials of the experiment consisted of more than eight objects.

Two absolute gradable adjectives were tested: the maximum absolute gradable adjective *sauber* ‘clean’ and the minimum absolute gradable adjective *dreckig* ‘dirty’. There were two trials per adjective.\(^3\) The materials consisted of picture cards (14x14 cm) that each displayed one object. In each of the four trials, the participants saw eight picture cards with either teddy bears or balls. Every object was of a different color. The property denoted by the adjective was realized to different degrees across the objects, i.e., they changed from dirty to clean as illustrated in Figure 2.

\(^3\) The trials described here were part of a larger study. In total, the experiment consisted of 24 test items and 20 filler items.
Importantly, the visual set up for clean and dirty was identical. This way, it was possible to explore whether participants interpret the negative and the positive antonyms as direct opposites. Alternatively, they could interpret some objects as showing neither the property denoted by the positive nor the property denoted by the negative antonym, e.g., objects that count neither as clean nor as dirty.

A total of eight filler items were included in this experiment. The fillers served to control for the distance between trials with the same visual context. As mentioned before the experiment consisted of more trials than the ones described here. Since we used a within-subjects design, participants saw some visual contexts twice, but with different test prompts. The filler items enlarged the distance between the same visual contexts in order to minimize possible influences of the first presentation. Each filler trial consisted of eight picture cards displaying toys (buckets, dices, soccer balls, books, and Lego® bricks). In each trial, objects from two different basic-level categories were shown. The objects differed in shape (round or square) and color (blue or red). These filler items corresponded to the participants’ turns to make a request themselves, which was similar to the test prompts, but used different adjectives (shape or color), e.g., Bitte gib mir die blauen Fußbälle ‘Please give me the blue soccer balls’.

---

4 Note that the numbers on the picture cards were not present in the original experiment; they are added here for easier reference to them in the result section.
Preceding the main test, participants received three practice trials without adjectives (e.g., *Gib mir bitte die Teddys* ‘Please give me the teddies’) to familiarize them with the task. The inclusion of the practice trials also served to illustrate that participants were allowed to select as many or few objects as they wanted. In other words, selecting only one object was introduced as an appropriate response, although the noun in the test prompt contained a plural DP. If the participants noticed that only one object was available matching the test prompt, the experimenter explained that the puppet’s request sounded the same independent of the number of matching objects. If the participants forgot to select a required object, the experimenter pointed this out to the participant.

In order to introduce the participants to the objects used in the experiment and their names the experimenter presented single exemplars of the objects and asked the participants to label them.

### 3.2.2. Procedure

The testing was divided into two sessions, which were about 12 days apart to minimize the influence of the items in the first session on the items in the second session. In the first session, the positive pole adjective *clean* was tested and in the second session the negative pole adjective *dirty* was tested. The visual contexts and the order of presentation were kept identical for both sessions. This allowed us to directly compare the standards of comparison for positive and negative antonyms. The children were tested individually in a quiet room at their day-care centers. Adults were tested at university. The participants sat next to the experimenter on the floor or at a table large enough to display all picture cards at the same time.

At the beginning of each test session participants were introduced to a hand puppet. They were told that the puppet wants to play a card game. The puppet and the child were each given a special dice that they were to roll without letting the other one see the outcome. The puppet’s rolls corresponded to the test items, and the participants’ rolls to the filler items. This set-up was chosen to engage the children in a game situation that naturally requires taking turns. The participants’ dice showed either a square or a circle, or a blue color dot, or a red color dot. When it was their turn, the experimenter distributed the eight picture cards on the table and participants had to roll their dice. When the dice showed ‘blue’, for example, they had to ask the puppet to hand her the blue toys, etc. When it was the puppet's turn, the experimenter distributed the picture cards on the table and the puppet rolled the dice and made its request corresponding to the test items. The participants’ task was to select those objects that in their opinion matched the test prompt.

### 3.2.3. Data Analysis

The analysis is based on the choices participants made, i.e., which objects they considered clean or dirty, respectively. At the group level, the data were
analyzed in terms of the percentage of choices for each object across the two trials for each adjective. This analysis yielded the location of the standard of comparison on the scale for clean and dirty (cf. Figures 3 and 4). Extending previous studies, we analyzed the participants’ individual choices as well. At the individual level, the standard of comparison was coded as the so-called cut-off point. For clean this is the dirtiest object that was considered clean (together with all cleaner objects); for dirty this is the cleanest object considered dirty (together with all dirtier objects). If a participant for instance selected the teddies 1, 2, 3, 4, 5, 6, and 7 (cf. the display in Figure 2a), then the cut-off point for dirty teddies is 7. Following Kennedy (2007), this is considered a target-like response for minimum absolute gradable adjectives. In the same line of reasoning, the target-like cut-off point for maximum gradable adjectives such as clean is object number 8, which shows no dirt at all. No other object should be judged clean. By directly comparing the cut-off points for the positive and the negative adjective we were able to assess whether the negative and the positive adjectives are direct opposites or whether there are some objects that refer to neither the positive nor the negative adjective.

3.3. Results and Discussion

A significant difference in the cut-off points was found neither across the two trials for dirty nor across the two trials for clean (Wilcoxon-Test, all ps > .05). Accordingly, the two trials were analyzed together. Figures 3 and 4 illustrate how often the respective objects 1 to 8 (cf. Figure 2 for the objects the numbers refer to) were chosen by the participants when asked to pick the clean balls/teddies (Fig. 3) and to pick the dirty balls/teddies (Fig. 4). On the x-axis, the eight objects are plotted by number. Note that they are ordered linearly only for ease of illustration (1 = dirtiest object, 8 = cleanest object). The raw numbers of choices are given in the Appendix. A total of 34 responses were analyzed 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).

When asked to pick the clean objects, the 5-year-olds as well as the adults only selected object 8, which does not have any dirt spot (cf. Figure 3). These data indicate that the standard of comparison for clean is the maximal degree of the scale both for children and adults.

Interestingly, the choices for dirty objects are the mirror image of the choices for clean objects, as shown in Figure 4. Every object that contained some amount of dirt was selected as dirty.
The graphs in Figures 3 and 4 summarize the choices for both trials per adjective. Visual inspection already indicates which objects constituted the cut-off points for the respective adjectives in adults and children. Table 1 displays the statistical values for the cut-off points for each trial separately (mean and median). In both trials, the cut-off point for dirty is object number 7, i.e. the standard of comparison is a non-zero degree. In both clean trials, the cut-off point is object number 8, i.e. the standard of comparison is the maximal degree. Importantly, children and adults did not significantly differ in their cut-off points (Mann-Whitney-U).
Table 1. Mean (standard deviation) and median cut-off points for each trial per group, and significance according to Mann-Whitney-U tests.

<table>
<thead>
<tr>
<th>Trial</th>
<th>5-year-olds</th>
<th>Adults</th>
<th>Mann-Whitney-U</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Median</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>clean teddies</td>
<td>7.9 (.3)</td>
<td>8</td>
<td>8.0 (0)</td>
</tr>
<tr>
<td>clean balls</td>
<td>7.9 (.3)</td>
<td>8</td>
<td>8.0 (0)</td>
</tr>
<tr>
<td>dirty teddies</td>
<td>7.0 (0)</td>
<td>7</td>
<td>7.0 (0)</td>
</tr>
<tr>
<td>dirty balls</td>
<td>7.0 (.4)</td>
<td>7</td>
<td>7.0 (0)</td>
</tr>
</tbody>
</table>

A further individual analysis investigated whether participants exhaustively divided the range of objects between the negative (dirty) and positive (clean) antonym or whether there were any objects for which both the positive and the negative adjective are judged false. Recall that the participants saw the same arrays of objects for the trials with the negative antonym and for the trials with the positive antonym so that the cut-off points could be directly compared. All adult responses and 88.2% of the responses of the 5-year-olds (15/17) did not show a gap between the cut-off points for clean and dirty. This result indicates that children and adults exhaustively classified the objects as either clean or dirty. Put differently, every object that was not considered clean was considered dirty and vice versa (cf. example (3)).

The findings of the present study suggest that adults interpret absolute gradable adjectives not as vague and children at age 5 share this interpretation. In both test groups the standard of comparison for clean and dirty was the minimal or maximal endpoint of the corresponding scale, respectively. These data confirm the judgements in example (1): if something is not dirty, it is not possible that it has some dirt on it. In contrast, if something is clean, it cannot be cleaner. Moreover, the present data confirm the judgement in example (3). For absolute gradable adjectives, borderline-cases do not seem to exist. A series of objects can be exhaustively divided into a set of clean and a set of dirty objects, without any objects that are not attributed to one of the sets.

Our study extends previous work on the interpretation of absolute gradable adjectives in several ways. First, we used a different experimental design, which was used to investigate both positive and negative adjectives and invited the participants to establish their own ordering of the objects presented. Importantly, our data from German confirm the findings for English and Italian across different designs: absolute gradable adjectives are interpreted with endpoint standards. Second, by comparing positive and negative adjectives our findings provide new evidence that absolute gradable adjectives share the same standard degree and are not interpreted as vague predicates. That is: clean is the same as not dirty!
4. Conclusion

Semantic theory argues that gradable adjectives can be relative or absolute. The two classes differ regarding the location of the standard of comparison on the corresponding scale (endpoint for absolute gradable adjectives vs. midpoint for relative gradable adjectives) and regarding their context-sensitivity. Moreover, it has been claimed that relative gradable adjectives, but not absolute gradable adjectives, share the characteristics of vague predicates. Previous empirical studies, testing children’s and adults’ interpretation of relative and absolute gradable adjectives, partially confirmed this distinction. However, some absolute gradable adjectives (e.g., full) were not interpreted as expected. Moreover, negative adjectives were not always taken into account and children’s judgements were restricted to predefined object orderings.

Our results from a forced-choice picture task in German with 26 adults and 17 five-year-olds indicate that absolute gradable adjectives that refer to partially closed scales such as clean and dirty are not interpreted as vague predicates. Adults and children assign absolute gradable adjectives to an endpoint standard of comparison. The group as well as the individual data revealed that objects were considered as either clean or dirty, i.e. there were no borderline-cases, confirming the diagnostic criterion by Kennedy (2007). The findings suggest that, unlike relative gradable adjectives, absolute gradable adjectives share the same standard of comparison, which is defined as a single degree on the corresponding scale.

Currently, we are extending our study to include the interpretation of relative gradable adjectives and to test younger children. In a further step, the same design has been used to examine the context-(in)sensitivity of absolute and relative gradable adjectives (Weicker in prep.).

These findings from acquisition can contribute to further refinement of the semantics of adjectives. We argue that the default-interpretation of minimum and maximum absolute gradable adjectives is not vague, although they can receive a non-complementary interpretation as shown by Rotstein & Winter (2004). Moreover, taking findings from previous studies into account we propose that the relative-absolute distinction of gradable adjectives holds cross-linguistically.

Appendix

Table 2. Number of choices for each object across trials in the group of 5-year-olds. Total number of possible choices per object = 34.

<table>
<thead>
<tr>
<th>Adjective</th>
<th>Object</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>clean</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>34</td>
</tr>
<tr>
<td>dirty</td>
<td></td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>33</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 3. Number of choices for each object across trials in the adult group. Total number of possible choices per object $= 52$.

<table>
<thead>
<tr>
<th>Adjective</th>
<th>Object</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>clean</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>dirty</td>
<td></td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>0</td>
</tr>
</tbody>
</table>

References


