Classification strategies and the semantics of gradable adjectives

Louise McNally

Departament de Traducció i Ciències del Llenguatge
Universitat Pompeu Fabra

March 4, 2010
First, some data: Two kinds of gradable properties

(1) Marta is **tall**.
(2) The door is **closed**.

Differences between these properties:
- The truth of (1) is context-dependent; that of (2) isn’t.
- **closed** yields “crisp judgments”; **tall** does not.
- **tall** yields borderline cases; **closed** does not.
- **tall** gives rise to the Sorites paradox; **closed** does not.
The differences: Context dependence

(3) a. Compared to Andrea, Marta is tall.  [True]
b. Compared to the Torre Mapfre, Marta is tall.  [False]

(4) a. ??Compared to Door #1, Door #2 is closed.
b. ??That box of cookies is closed for a box my daughter has gotten into.
The differences: Crisp judgments

Assume (5) is true. It’s not obvious that (6) is false.

(5) The girl is tall.
(6) The boy is tall.
Crisp judgments, continued

In contrast, (7) is clearly true, and (8) clearly false.

(7) The door on the top is closed.
(8) The door on the bottom is closed.
The differences: Borderline cases

True or false??

(9) Leyre is tall.

A group of 11-year-olds:

<table>
<thead>
<tr>
<th>Kid</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marta</td>
<td>1.60m</td>
</tr>
<tr>
<td>Claudia</td>
<td>1.58m</td>
</tr>
<tr>
<td>Leyre</td>
<td>1.50m</td>
</tr>
<tr>
<td>Irene</td>
<td>1.41m</td>
</tr>
<tr>
<td>Carla</td>
<td>1.40m</td>
</tr>
</tbody>
</table>
The differences: The Sorites Paradox

- Exemplified with *tall*:
  
  **Premise 1:** A 1.65m tall kid is tall (for a kid).
  **Premise 2:** If $x$ is a tall kid and $y$ is a kid 1mm shorter than $x$, then $y$ is a tall kid.
  **Possible conclusion:** A 0.50m tall kid is a tall kid.

- Premise 2 fails if we try to recreate the paradox with *closed*. 
Two kinds of properties, two kinds of standards

Kennedy (2007b): The difference lies in the adjectives’ standards

- RELATIVE standards: determined relative to a comparison class; as a rule, not an endpoint on a scale.
- ABSOLUTE standards: a maximum or minimum value on a scale.

(see Kennedy and McNally 2005)
The question

- What explains the contrasts between these two kinds of gradable properties?
The question

- What explains the contrasts between these two kinds of gradable properties?
  - Kennedy’s answer: degree-based semantics + a principle of Interpretive Economy
The question

- What explains the contrasts between these two kinds of gradable properties?
  - Kennedy’s answer: degree-based semantics + a principle of Interpretive Economy

- But what if you want to eliminate degrees from the semantics of positive form adjectives (see e.g. Bale 2006; van Rooij 2007)?
The question

- What explains the contrasts between these two kinds of gradable properties?
  - Kennedy’s answer: degree-based semantics + a principle of Interpretive Economy

- But what if you want to eliminate degrees from the semantics of positive form adjectives (see e.g. Bale 2006; van Rooij 2007)?

- And can/should we rely on Interpretive Economy?
Goals of the talk

- To focus on some underappreciated facts about scales and standards:
Goals of the talk

- To focus on some underappreciated facts about scales and standards:
  - Some closed scale adjectives have relative standards.
Goals of the talk

- To focus on some underappreciated facts about scales and standards:
  - Some closed scale adjectives have relative standards.
  - Some cases of absolute standards are not best modeled as minimum/maximum values on a scale.
Goals of the talk

- To focus on some underappreciated facts about scales and standards:
  - Some closed scale adjectives have relative standards.
  - Some cases of absolute standards are not best modeled as minimum/maximum values on a scale.
  - Absolute standards can be context dependent, but are so in a crucially different way than are relative standards.
Goals of the talk

○ To focus on some underappreciated facts about scales and standards:
  ○ Some closed scale adjectives have relative standards.
  ○ Some cases of absolute standards are not best modeled as minimum/maximum values on a scale.
  ○ Absolute standards can be context dependent, but are so in a crucially different way than are relative standards.

○ To argue that a focus on the properties that adjectives contribute and the classification strategies underlying their use sheds light on the relative/absolute distinction.
Plan

- More on the data, scale structure, and the relative/absolute distinction
- Classification strategies and the relative/absolute distinction
- Implications for the semantics of the positive form
Gradable adjectives in Kennedy and McNally (2005)

All gradable adjectives denote measure functions.

(10) a. tall(Marta) = 1.65m  
b. closed(\(\forall x. \text{door}(x)\)) = 0 ÷ 1

But they can differ in their SCALE STRUCTURE:

- **Closed scale:** there are min./max. values in the codomain of the measure function.
- **Open scale:** there are no such min./max. values.
From measure function to property

- Measure functions turn into properties via degree morphology.
- Degree morphology introduces a STANDARD VALUE, which determines whether the adjective truthfully applies to its argument or not.
Standards

A standard can be:

- **Relative:** a context dependent value (e.g. *tall for a 12-year-old*)
- **Absolute:**
  - A minimum value: *The door is open.*
  - A maximum value: *The door is closed.*
The standard for a positive form adjective is contributed (in English and presumably most languages) by a null morpheme \( pos \):

\[
(11) \quad pos : \lambda g \lambda x. g(x) \succeq s(g)
\]

where \( s \) is a context-sensitive function that chooses a standard of comparison so as to ensure that the objects that the positive form is true of ‘stand out’ in the context of utterance, relative to the kind of measurement that the adjective encodes. (Kennedy, 2007b, p. 17)
Some general concerns about this semantics

- **Acquisition**: Children arguably learn adjectives like *big* before they understand numbers or orderings between them.
Some general concerns about this semantics

- **Acquisition:** Children arguably learn adjectives like *big* before they understand numbers or orderings between them.

- **Simplicity:**
  - Degree-based analyses depend on a null *pos* morpheme for which there is no obvious evidence.
Some general concerns about this semantics

- **Acquisition:** Children arguably learn adjectives like *big* before they understand numbers or orderings between them.

- **Simplicity:**
  - Degree-based analyses depend on a null *pos* morpheme for which there is no obvious evidence.
  - A degree semantics for the positive form is a kind of generalization to the worst case.
Some general concerns, continued

- **Typology:** Not all languages use degree expressions to express comparisons.

(12) Ua loa lenei va’a, ua puupuu lena
     is long this boat, is short that
     ’This boat is longer than that boat.’ (Samoan)

(see e.g. Klein 1980; Bale 2006; van Rooij 2007; Kennedy 2007a)
The crucial question:

What standard allows the objects that the positive form is true of to “stand out”?
Kennedy’s answer:

- **Relative adjectives**: A degree that depends on the individuals under consideration in the context.

- **Absolute adjectives**: The degree that marks the transition from a zero to a non-zero value on the scale, or from a non-maximal to a maximal value.
Relative adjectives

The intuition concerning “standing out” works well for sets of individuals with a discontinuous distribution:

<table>
<thead>
<tr>
<th>Kid</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marta</td>
<td>1.65m</td>
</tr>
<tr>
<td>Pau</td>
<td>1.10m</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kid</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marta</td>
<td>1.60m</td>
</tr>
<tr>
<td>Claudia</td>
<td>1.58m</td>
</tr>
<tr>
<td>Leyre</td>
<td>1.58m</td>
</tr>
<tr>
<td>Irene</td>
<td>1.41m</td>
</tr>
<tr>
<td>Carla</td>
<td>1.40m</td>
</tr>
</tbody>
</table>
Relative adjectives

It doesn’t work well for sets with a continuous distribution:

<table>
<thead>
<tr>
<th>Kid</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jordi</td>
<td>1.11m</td>
</tr>
<tr>
<td>Pau</td>
<td>1.10m</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Kid</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{25}$</td>
<td>1.25m</td>
</tr>
<tr>
<td>$k_{24}$</td>
<td>1.24m</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$k_3$</td>
<td>1.03m</td>
</tr>
<tr>
<td>$k_2$</td>
<td>1.02m</td>
</tr>
<tr>
<td>$k_1$</td>
<td>1.01m</td>
</tr>
</tbody>
</table>
Relative adjectives

$k_{25}$ clearly stands out from $k_1$, but there is no way to partition this set completely into two subsets such that all members of one subset stand out from all members of the other.
Relative adjectives

- \( k_{25} \) clearly stands out from \( k_1 \), but there is no way to partition this set completely into two subsets such that all members of one subset stand out from all members of the other.

- Intuitively, if we associate such adjectives with an extension gap we can create the necessary “space” between two such subsets.
Relative adjectives

- $k_{25}$ clearly stands out from $k_1$, but there is no way to partition this set completely into two subsets such that all members of one subset stand out from all members of the other.

- Intuitively, if we associate such adjectives with an extension gap we can create the necessary “space” between two such subsets.

- But what would the standard get fixed as so as to make those individuals that are Adj stand out from those in the extension gap?
Assessment

It’s not the standard that determines what stands out, but what stands out that determines the standard.

Compare Barker (2002) on *stupid* and related adjectives.
Absolute adjectives

Why should the degree that marks the transition from a zero to a non-zero value on the scale, or from a non-maximal to a maximal value, be the one that makes an object stand out?
Answer: It needn’t be

- Some closed scale adjectives allow for relative standards that behave like the standard for *tall*. 
Answer: It needn’t be

- Some closed scale adjectives allow for relative standards that behave like the standard for *tall*.
- Others have standards that behave like absolute standards without being maximum/minimum values.
Some closed scale adjectives allow for relative standards that behave like the standard for *tall*.

Others have standards that behave like absolute standards without being maximum/minimum values.

These facts suggest that the crucial difference between relative and absolute adjectives does not have to do with the properties of the measure function’s codomain (i.e. whether the scale is closed or not).
Closed scales with relative standards

(13) For a student who has just moved here, she is very familiar with the class routines and her teachers’ expectations. In fact, she’s completely familiar with them.

- Both instances involve the same sense of familiar.
- very indicates a relative standard for the first instance (Kennedy and McNally, 2005).
- completely indicates a closed scale (Hay et al., 1999).
Closed scales with relative standards

Nonetheless, this sort of manipulation of the standard does not always seem to be possible.
Syrett (2007)

Adult subjects consistently considered a request for “the full jar” infelicitous in this context:
But...

In contrast, audiences have consistently resisted our attempts to deny that in cases like this the glass is full.
Possible analyses

- The standard for *full* is relative.
- When we call the champagne flute “full”, we are speaking loosely (Lasersohn, 1999).
- The granularity of the degree of measurement is coarse enough that the flute can count as completely full.
The problem for loose talk or coarser granularity

Sometimes we use adjectives like *full* in cases where the degree of the property in question is so far from the maximum value that it seems implausible that we could be speaking loosely or have coarsened the granularity of measurement.
Example: full glasses

This photograph is entitled:

Wine Glass 2 (full)

A recalibrated maximum?

- This standard differs from relative standards: *If we know what the standard is*, crisp judgments remain, and Sorites Premise 2 is not valid.

- So why not say we simply ignore part of the volume and that the standard continues to be a maximum?
The problem with a recalibrated maximum analysis

In describing the wine glass, (14-a) sounds contradictory (or infelicitous at best); if we were simply ignoring part of the volume, we might expect (14-a) to be interpretable as (14-b).

(14)  
   a. The glass is completely full, but not completely full.  
   b. This glass is completely filled to the fill line, but not completely full.
Another candidate for a non-maximal absolute standard

The truth of a color term on a color-extension reading entails more than a minimum but less than a maximum:

(15) a. This shirt is blue, but not completely blue.
b. This shirt is not black.
Another candidate for a non-maximal absolute standard

But the standard is not like typical relative standards:

(16) a. This shirt is slightly/very blue.
    b. ??This shirt is blue for a patterned t-shirt.
    ? ??This shirt is blue compared to the one I wore yesterday.

Rather, it seems to be something like “predominantly COLOR.”
Comparison class and vagueness with these adjectives

Comparison class:
- The standard (in the case of *full*, though not with color extension) can depend on the class to which the object being described belongs.
- But intuitions about the standard do not depend on the distribution of the degrees to which the property is held within that class.
Comparison class and vagueness with these adjectives

Comparison class:
- The standard (in the case of *full*, though not with color extension) can depend on the class to which the object being described belongs.
- But intuitions about the standard do not depend on the distribution of the degrees to which the property is held within that class.

Vagueness (in the Sorites sense):
- My intuitions about “crisp judgments,” borderline cases, and the Sorites paradox are still a bit unclear.
- But if you push on your intuitions about other absolute adjectives, intuitions can become equally unclear.
Assessment

What matters is not “standing out” *per se* but the possibility of being sortable according to some property.
Assessment

- What matters is not “standing out” *per se* but the possibility of being sortable according to some property.
- Whether and how that can be done depends on what that property is.
The intuition

- Some (gradable) adjectives denote properties that reflect classifications by similarity.
- Others denote properties which reflect classifications on the basis of rule-like criteria.

(See e.g. Hahn and Chater 1998.)
**Similarity-based classification**

We cannot evaluate the truth of (17) without comparing the piece of wood to something else.

(17) This piece of wood is long.
Rule-like classification

(18) is clearly true of the depicted door.

(18) This door is closed.

We can check this without knowing anything about the characteristics of any other doors (or anything else).
What are your intuitions about (19) applied to the jars on the right?

(19) The jar is full.

What criterion/a are you using?
Absolute vs. relative standards

Though potentially vague and dependent on a comparison class, the truth conditions for absolute adjectives differ from those for relative adjectives in a crucial way:

Our intuitions about whether the former hold of their arguments are not influenced by random variation in the selected comparison class.
The role of the comparison class

- Relative standards depend crucially on a comparison class because the classification involves sorting the individuals in that class according to similarity.
  - What is similar to what depends on what is being considered.
The role of the comparison class

- Relative standards depend crucially on a comparison class because the classification involves sorting the individuals in that class according to similarity.
  - What is similar to what depends on what is being considered.

- Absolute standards do not: the role of the “comparison” class is simply to fix contextually variable parameters that will allow a rule-like classification to proceed.
More on similarity based classification

**Clustering** is a typical similarity-based classification. The typical clustering strategy:

- The existence of $n$ clusters is presupposed.
- Each cluster is associated with a centroid.
- Minimize within-cluster distance and maximize between-cluster distance.

(see also Gärdenfors 2000 on Conceptual Spaces)
Similarity based classification

Thinking in terms of similarity-based classification explains:

- The need for a comparison class: Follows from the need for two centroids to support clustering.
- The intuitive validity of Sorites Premise 2: The premise conflicts with maximizing within-cluster similarity and between-cluster distance.
- Indeterminacy/borderline cases: Arise because some individuals might be equidistant between the two cluster centroids.
Rule-like classification

If we classify an object according to “rule”:

- There is no need for a comparison class except to determine which variant of a rule might apply.
- Sorites Premise 2 need not be valid: If the rule has clear application criteria, there will be (relatively) crisp judgments.
- Indeterminacy/borderline cases: Will arise only if the rule lacks clear application criteria.
Conclusion

The relative-absolute distinction is real, but is less about scale structure than about the nature of the properties involved and our criteria for determining their application.

- Relative standards reflect similarity-based classification.
- Absolute standards reflect rule-like classification.
- But both kinds of standards can be non maximal/minimal and context dependent in some sense.
Degree-based semantics for the positive form

\[ pos(A)(x) \text{ is true iff } [A](x) \succeq s([A]) \]
Klein’s degreeless semantics for the positive form

\((A)(x)\) is true iff \(x\) is in the positive extension of \(A\), where:

- Adjectives denote functions which assign individuals to positive extensions, negative extensions, and extension gaps.
- Extensions are defined relative to a domain \(D\) (i.e. comparison class), \(|D| \geq 2\).
- At least two members of the comparison class are required to be distinct.
- If the comparison class for the adjective is defined, the positive and negative extensions of the adjective must be nonempty.
Problems for Klein’s analysis

Klein’s analysis overuses comparison classes to induce orderings in two different ways:

- For relative adjectives, the distinctness condition should rule out two-member comparison classes whose members differ only slightly on the property in question. But it doesn’t (and, indeed, cannot, because of the way Klein handles comparatives).
Problems for Klein’s analysis

Klein’s analysis overuses comparison classes to induce orderings in two different ways:

- For relative adjectives, the distinctness condition should rule out two-member comparison classes whose members differ only slightly on the property in question. But it doesn’t (and, indeed, cannot, because of the way Klein handles comparatives).

- For absolute adjectives, true comparison classes are irrelevant and thus it should not be possible for an object to be Adj relative to one class and not Adj relative to another. But on Klein’s analysis, it must be possible, again, in order to handle comparatives.
Additional considerations

- Removing degrees entirely from the basic semantics of gradable predicates carries a very high cost. In the worst case, it amounts to denying that we conceive of certain properties as gradable.
Additional considerations

- Removing degrees entirely from the basic semantics of gradable predicates carries a very high cost. In the worst case, it amounts to denying that we conceive of certain properties as gradable.

- But it also seems that the criticisms by advocates of degreeless semantics for the positive form have less to do with the “deep” semantics of the *adjective* than with the semantics of the *positive form*, namely, that the semantics of the positive form is characterized with reference to a degree that serves as the standard.
Towards an alternative semantics for the positive form

- Distinct sorts of satisfaction conditions for absolute and relative adjectives.
- Extensions of absolute adjectives in simple, absolute terms.
- Extensions of relative adjectives defined in terms of similarity to (pre-established) centroids for the adjective and its contrary.

NB: This does not force or exclude the elimination of measure functions from the semantics of the adjective. I will be neutral on this point here.
The positive form for absolute adjectives

\[ \text{AdjP}_{abs}(x) \text{ is true in context } C \text{ iff } x \text{ satisfies the satisfaction conditions imposed by AdjP in } C. \]
Calculating similarity for relative adjectives

Two approaches to classification (of 1-dimensional properties such as height) into contraries:

- Use two clusters and make the classification partial.
- Use three clusters (one for the adjective, one for its contrary, and one for things which fall in the middle) and make the classification total.
Calculating similarity for relative adjectives

Considerations:

- Using two clusters requires justifying a metric such as “classify $x$ as $Adj$ iff the degree of similarity to the centroid for $Adj$ exceeds the degree of similarity to $Adj$’s contrary by a factor of 2.”
  - This looks arbitrary.
Calculating similarity for relative adjectives

Considerations:

- Using two clusters requires justifying a metric such as “classify \( x \) as \( \text{Adj} \) iff the degree of similarity to the centroid for \( \text{Adj} \) exceeds the degree of similarity to \( \text{Adj} \)’s contrary by a factor of 2.”
  - This looks arbitrary.

- Using three clusters requires justifying the reason for a third cluster corresponding to the gap between \( \text{Adj} \) and its contrary.
  - Perhaps this makes sense if 1) properties are generally positively defined, not negatively defined, and 2) there has to be a certain density to the class in question in order for the property to be useful.
Calculating similarity for relative adjectives

$\text{AdjP}_{rel}(x)$ is true in context $C$ and relative to a comparison class $K$ iff $x$ is more similar to the centroid from $K$ for the property contributed by $\text{AdjP}$ than it is to the centroid for any other cluster under consideration for classifying $x$.

NB: Depending on intuitions about whether e.g. $\text{very Adj}$ entails $\text{Adj}$, this semantics could be generalized to cases where $>2$ adjectives are used to classify individuals according to a given property.
Potential concerns

- We have two different kinds of satisfaction conditions for adjectival predications.
- The conditions for relative adjectives make reference to contraries.
But perhaps this is as it should be

(From S. Boynton *Opposites* / From Rosa Mendoza, et al. *Colors and Shapes*)
Conclusion

Our thinking about gradable properties has been heavily conditioned by thinking in terms of orderings rather than (potentially multi-dimensional) similarity relations.
Conclusion

Our thinking about gradable properties has been heavily conditioned by thinking in terms of orderings rather than (potentially multi-dimensional) similarity relations.

While we cannot lose sight of the need to derive orderings, our semantics for adjectives should not lose sight of nature of the properties involved, either.
Conclusion

- Our thinking about gradable properties has been heavily conditioned by thinking in terms of orderings rather than (potentially multi-dimensional) similarity relations.

- While we cannot lose sight of the need to derive orderings, our semantics for adjectives should not lose sight of nature of the properties involved, either.

- Thinking about the classification strategies that might be involved in ascribing adjectival properties leads to a fresh perspective on adjectival semantics.


