Slightly coerced: Processing evidence for adjectival coercion by minimizers

Elizabeth Bogal-Allbritten
University of Massachusetts Amherst

1 Introduction
Minimizers (slightly, a little, etc.) have been reported to compose only with adjectives whose scales contain a minimum endpoint (Kennedy and McNally 2005). The authors predict, and report, minimizers to be infelicitous with adjectives whose scales lack a minimum endpoint (1a). By contrast, minimizers are reported to be felicitous with adjectives whose scales contain minimum endpoints (1b).

(1)
   a. Slightly {large, tall, short, narrow}
   b. Slightly {dirty, ugly, dangerous, impure}

However, the predicted infelicity of (1a) has been questioned (Paradis 1997, Kagan and Alexeyenko 2011, Bylinina 2011, Solt 2011, Sawada 2011, Sassoon 2012, inter alia). This paper continues this line of research. On the basis of offline and online experimental results, I argue that minimizers only compose with adjectives with minimum endpoints, but are capable of coercing an endpoint where one is otherwise unavailable:

(2) Endpoint Coercion Hypothesis: Minimizers can, and must, coerce a minimum endpoint ($d_{\min}$) on an open adjectival scale.

When an open scale (i.e., dimensional) adjective composes with a minimizer, the minimizer coerces a minimum endpoint by interpolating the \(\text{ENDPT}\) function (3).

(3) \[
\text{ENDPT} = \lambda G \cdot \lambda d_0 \lambda x_0 . G(d_0) \wedge d \geq d_s
\]

\(\text{ENDPT}\) takes an adjective \(G\) and returns \(G\) with a limited scale (domain of degrees) such that all degrees \(d\) on the scale associated with \(G\) exceed some contextually determined \(d_s\).

The paper proceeds as follows. In §2, I summarize diagnostics for open vs. closed scale structure and revisit the problem posed by dimensional adjectives in greater depth, citing the results of two offline rating studies that confirm the felicity of collocations like (1a). I then present two hypotheses for how

\* This work has benefitted greatly from discussion with Lyn Frazier, Rajesh Bhatt, Charles E. Clifton Jr., and the audience at CLS 48. Any remaining errors are mine. This research was supported by a National Science Foundation Graduate Research Fellowship.
minimizer-modified dimensional adjectives can be incorporated into our theory of adjectival and minimizer meaning. In addition to the Endpoint Coercion Hypothesis, I explore a second possibility from Sawada (2011), in which minimizers are redefined such that we predict them to compose with open scale adjectives. In §3, I present the results of an online self-paced reading study showing that minimizer-modified dimensional adjectives incur longer reading times than modified negative evaluative adjectives. Following previous work on coercion, I take longer reading times to indicate increased processing costs, which in turn suggest coercion has taken place. In §4, I address variation between two paraphrases (goal-based standard, group-based standard) of minimizer-modified dimensional adjectives, supporting the claims with data from a paraphrase task, and relating these paraphrases to the semantics of ENDPT. §5 concludes.

2 Diagnosing scale structure

2.1 Open vs. closed scales
Following much previous work, I assume that adjectives are functions that relate individuals to degrees (e.g., Cresswell 1976, Heim 2001).

\[
\text{[[tall]]} = \lambda d \lambda x. \text{tall}'(x) \geq d
\]

I follow Kennedy and McNally (2005) in classifying scales as either ‘open’ or ‘closed.’ Closed scales have a minimum \(d_{\text{min}}\) and/or maximum \(d_{\text{max}}\) endpoint. Open scales lack both endpoints.\(^1\)

<table>
<thead>
<tr>
<th>Open Scale (no endpoint)</th>
<th>Closed Scale ((d_{\text{min}}, d_{\text{max}}))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No endpoints:</strong></td>
<td><strong>Maximum endpoint:</strong></td>
</tr>
<tr>
<td>Dimensional: Large, tall, short, narrow</td>
<td>Absolute: Straight, safe, pure, clean</td>
</tr>
<tr>
<td></td>
<td><strong>Minimum endpoint:</strong></td>
</tr>
<tr>
<td></td>
<td>Absolute: Bent, dangerous, impure, dirty</td>
</tr>
<tr>
<td></td>
<td>Negative evaluative: Rude, ugly, boring</td>
</tr>
</tbody>
</table>

**Figure A:** Types of adjectival scales

A widely used diagnostic for minimum endpoints \(d_{\text{min}}\) is the existence of an entailment relation from the comparative form of an adjective \((\text{adj}-\text{er}, \text{more adj})\) to the POS-marked form of the adjective \((\text{POS-adj})\) (Rotstein and Winter 2004, Kennedy and McNally 2005). When POS-marked, POS selects the endpoint of an absolute adjective rather than determine one based on context Kennedy 2007).

---

\(^1\) Within the class of closed scale adjectives, absolute adjectives are those adjectives with an endpoint whose value is physically ‘real’ (e.g., 0° bend, no dirt) (Rotstein and Winter 2004, Kennedy and McNally 2005). When POS-marked, POS selects the endpoint of an absolute adjective rather than determine one based on context Kennedy 2007).
Kennedy and McNally 2005). The sentence in (5) shows the entailment relation arising for minimum endpoint absolute adjectives (*bent*). The endpoint on the scale of *bent* is ‘0° bent’ such that all degrees on the scale exceed this degree.

(5) The red pipe is more bent than the blue pipe. ⊨ The red pipe is bent.

The set of minimum endpoint adjectives in Figure A includes negative evaluative adjectives, which denote descriptive, and undesirable attributes (Bierwisch 1989). Rett (2008) and Bierwisch (1989) observe that negative evaluative adjectives exhibit the entailment relation described above:

(6) Sandy’s hat is uglier than Mary’s hat. ⊨ Sandy’s hat is ugly.

The scale of *ugly* contains only degrees exceeding the contextual standard of comparison for *ugliness* (Bierwisch 1989): an object that is *uglier* also counts as *ugly*. Henceforth, I refer collectively to minimum endpoint absolute and negative evaluative adjectives as ‘minimum endpoint adjectives.’

In contrast with minimum endpoint adjectives, open scale (i.e., dimensional) adjectives fail to exhibit the entailment relation:

(7) The blue box is larger than the red box. ⊭ The blue box is large.

The absence of an entailment relation in (7) suggests that the scale of *large* does not have a minimum endpoint which all other degrees on the scale exceed.

Figure B compares minimum endpoint and dimensional adjectival scales.

<table>
<thead>
<tr>
<th>Minimum endpoint adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>bent</em></td>
</tr>
<tr>
<td>•----------------------------</td>
</tr>
<tr>
<td>[d_{min} = 0^\circ\text{ of bend}]</td>
</tr>
<tr>
<td><em>ugly</em></td>
</tr>
<tr>
<td>•----------------------------</td>
</tr>
<tr>
<td>[d_{min} = \text{contextual standard of ugly}]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Open scale (dimensional) adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>large</em></td>
</tr>
<tr>
<td>…</td>
</tr>
<tr>
<td>•----------------------------</td>
</tr>
</tbody>
</table>

Figure B: Scale structures of minimum endpoint and open scale (dimensional) adjectives

Linguistic evidence for scale endpoints is supported by evidence from acquisition (Syrett et al. 2010) and processing studies (Frazier et al. 2008). I follow Frazier et al. in classifying endpoints on the scales of minimum endpoint

---

2 POS is a null morpheme that selects an appropriate standard of comparison (STND(G)) for a given adjective G. A representative denotation is given in (i).

(i) \[[[\text{POS}]] = \lambda G. \exists d. G(d)(x) > \text{STND(G)}\] (Cresswell 1976, *inter alia*)
adjectives as part of the lexically supplied and obligatorily processed meaning of an adjective. In other words, when bent or ugly exit the lexicon, the minimum endpoint is already part of adjective’s meaning.

2.2 Minimizers as diagnostics of $d_{\text{min}}$

2.2.1 Introducing minimizers

The distribution of minimizers (slightly, a little) is taken by Kennedy and McNally (2005) as an additional diagnostic for minimum endpoints $d_{\text{min}}$. The denotation of slightly in (8) is extrapolated from Kennedy and McNally (2005).

\[(8) \quad [[\text{slightly}]] = \lambda G \cdot \lambda x \cdot \exists d : d \geq \text{MIN}(G) \land G(d)(x),\]

where MIN returns the minimum degree on adjective $G$’s scale.\(^3\)

Minimizers are felicitous with minimum endpoint absolute adjectives but not maximum endpoint absolute adjectives. The paradigm in (9) holds both in English and in other languages (Klein 1998, Kagan and Alexeyenko 2011, Sawada 2011). Frazier et al. (2008) confirm the English intuitions with a rating task.

\[(9)\]

a. Slightly {bent, dangerous, dirty, impure} \quad Minimum endpt.
b. ?? Slightly {straight, safe, pure, clean} \quad Maximum endpt

As in the comparative entailment diagnostic, negative evaluative adjectives pattern like minimum endpoint absolute adjectives:

\[(10)\quad \text{Slightly} \{\text{rude, ugly, boring}\} \quad \text{Negative evaluative}\]

Judgments like (10) are reported by Paradis (1997) and are also found in earlier work (e.g., Bolinger 1972). Similar effects obtain cross-linguistically, as reported by Kagan and Alexeyenko (2011) for Russian and Sawada (2011) for Japanese.

Given the semantics in (8) for minimizers, we predict that only adjectives with a minimum endpoint can felicitionally compose with minimizers. However, this generalization is challenged by (11):

\[(11)\quad \text{Slightly} \{\text{tall, large, short, narrow}\}\]


\(^3\) The relation $\geq$ indicates “exceeds to a small degree” (Sawada 2011).
2.2.2 Rating task confirmation of intuitions
Two rating tasks were conducted to confirm the felicity of various combinations of minimizers and adjectives. In both tasks, materials were presented in Linger. Participants were University of Massachusetts Amherst undergraduates receiving course extra credit in return for participation. Only responses from native English speakers were considered.

In the first rating task, participants (n=24) rated sentences containing (minimizer-modified and unmodified) negative evaluative, positive evaluative, maximum endpoint absolute, and minimum endpoint absolute adjectives.\(^4\)\(^5\) The materials and task were modeled on the materials of Frazier et al. (2008), who compared ratings of (modified and unmodified) maximum endpoint and minimum endpoint absolute adjectives. I added negative evaluative adjectives in order to probe the intuition from §2.2.1 that they also are felicitous with minimizers. Sentences like (12) and (13) were rated on a scale from 1 to 5, where ‘5’ was defined as “completely acceptable, sounds like something a native English speaker might say.” Participants saw only one sentence from each set to avoid minimal pair comparison effects (Hirotani 2004).

(12) a. During the talk on ecosystems, the students looked uninterested.
   b. ...slightly uninterested.

(13) a. The camp leader said that the pole he found to pitch the tent was straight.
   b. ...slightly straight. c. ...curved. d. ...slightly curved.

In the second rating task, participants (n=32) rated sentences containing (minimizer-modified and unmodified) negative evaluative and dimensional adjectives. Sentences like (14) and (15) were rated using the same scale and methodology as above. Participants had not taken part in the first rating task.

(14) a. Maxine thought that the dress Molly bought was ugly.
   b. ...slightly ugly.

(15) a. Marsha said that the hallway in her apartment was narrow.
   b. ...slightly narrow.

\(^4\) In both tasks, half of all modified items used slightly and the other half used a little. No difference due to choice of minimizer was found.

\(^5\) Materials for both tasks included equal numbers of positive and negative evaluative adjectives appeared (selected to be antonymic pairs). Given space restrictions, I do not discuss the findings for positive evaluative adjectives in depth here. Minimizer-modified positive evaluative adjectives received acceptability ratings higher than maximum endpoint adjectives but lower than minimum endpoint, negative evaluative, and dimensional adjectives: 3.58 in the first task, 3.71 in the second task. For discussion, see §5.2 and especially Bogal-Allbritten (2011).
Table 1 presents the results of the first task.

<table>
<thead>
<tr>
<th>Adjective type:</th>
<th>Max. endpt. abs.</th>
<th>Min. endpt. abs.</th>
<th>Neg. evaluative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modifier:</td>
<td>Ø</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>Mean rating:</td>
<td>4.60</td>
<td>4.47</td>
<td>4.59</td>
</tr>
<tr>
<td>Penalty:</td>
<td>-1.53</td>
<td>-.16</td>
<td>-.27</td>
</tr>
</tbody>
</table>

Table 1: Mean ratings from first offline rating task

A penalty is found in the mean ratings of maximum endpoint absolute adjectives when unmodified and minimizer-modified conditions are compared (4.35 vs. 3.10). The penalty was statistically significant (pMCMC < 0.001). Comparable penalties were not found for modified minimum endpoint absolute adjectives (4.52 vs. 4.30) or negative evaluative adjectives (4.59 vs. 4.21). Ratings given to modified negative evaluative and modified minimum endpoint absolute adjectives were not significantly different: 4.32 vs. 4.31 (pMCMC = 0.2782).

The significant penalty observed for maximum endpoint absolute adjectives – but not their minimum endpoint counterparts – agrees with the findings of Frazier et al. (2008). In addition, the results of the first task show that negative evaluative and minimum endpoint absolute adjectives do not pattern significantly differently. This result supports inclusion of both in the set of minimum endpoint adjectives.

Table 2 presents the results of the second task. Mean ratings for minimizer-modified negative evaluative and dimensional adjectives were basically equivalent (4.21, 4.18). The difference in penalties can be attributed to unmodified dimensional adjectives receiving a lower baseline rating than negative evaluative adjectives.

<table>
<thead>
<tr>
<th>Adjective type:</th>
<th>Neg. evaluative</th>
<th>Dimensional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modifier:</td>
<td>Ø</td>
<td>Ø</td>
</tr>
<tr>
<td>Mean rating:</td>
<td>4.66</td>
<td>4.41</td>
</tr>
<tr>
<td>Penalty:</td>
<td>-0.47</td>
<td>-0.23</td>
</tr>
</tbody>
</table>

Table 2: Mean ratings from second offline rating task

The high – and nearly equivalent – ratings given to minimizer-modified negative evaluative and dimensional adjectives are the result that motivate the remainder of the study. The result supports the intuitions reported in §2.2.1, challenging Kennedy and McNally’s (2005) analysis of minimizers as a diagnostic of minimum endpoints. In the next section, I consider two possible hypotheses to address this challenge.

---

6 Statistical analysis for both rating tasks is found in Bogal-Allbritten (2011, §3.1).
2.3 Standard Calculation vs. Endpoint Coercion Hypotheses

One possible response to the felicity of minimizer-modified dimensional adjectives is to discard minimizers as a diagnostic for minimum endpoints. This is the strategy adopted by Sawada (2011). I characterize the class of accounts represented by Sawada as the ‘Standard Calculation Hypothesis.’

(16) **Standard Calculation Hypothesis:** We modify the denotation of minimizers so they can compose with open scale adjectives.

Sawada considered Japanese data in which the minimizer *chotto* modifies dimensional adjectives. Sawada defines *chotto* as in (17).

(17) \[[chotto] = \lambda G <\lambda x \exists d. G(d)(x) \land d > \text{STND}(G)\]

The minimizer no longer makes reference to a minimum endpoint \((d_{\text{min}})\) through application of the function \(\text{MIN}\). Instead, comparison between the degree \(d\) is made with the degree returned by \(\text{STND}\), the standard calculating function found in the POS morpheme (fn. 2).

A second possible response is to claim that minimizers require a minimum endpoint on the scale of the adjectives with which they compose, but that an endpoint can be coerced if one is not otherwise available:

(18) **Endpoint Coercion Hypothesis:** We maintain the original denotation of minimizers but argue that minimizers can coerce \(d_{\text{min}}\) on otherwise open scales.

Following Traxler et al. (2002), I define coercion as enrichment of a lexical item’s interpretation through interpolation of (non-syntactically realized) semantic structure. When an open scale (i.e., dimensional) adjective composes with a minimizer, the minimizer coerces a minimum endpoint by requiring interpolation of the \(\text{ENDPT}\) function (19). Once the adjective’s scale has been modified by interpolation of \(\text{ENDPT}\), composition between the adjective and the minimizer proceeds as usual.

(19) \[[\text{ENDPT}] = \lambda G <\lambda d, \lambda x \exists d. G(d)(x) \land d > d_{\text{S}}\]

\(\text{ENDPT}\) takes an adjective \(G\) and returns \(G\) with a limited scale (domain of degrees) such that all degrees \(d\) on the scale associated with \(G\) exceed some contextually determined \(d_{\text{S}}\).

---

7 The analysis of Bylinina (2011) discusses possibility of coercion by a minimizer, but locates it in a different part of the derivation: coercion applies to POS rather than to the scale. In depth comparison of the two views is left to future work.
Both hypotheses predict the results from the offline rating tasks reported in §2.2. However, there are other ways in which the two hypotheses may differ in terms of behavior predicted for minimum endpoint vs. dimensional adjectives. Under the Standard Calculation Hypothesis, all adjectives undergo the same steps of composition with a minimizer regardless of whether they exit the lexicon with a minimum endpoint, or not. The function STND applies and returns a standard of comparison corresponding either to an endpoint or to a standard of comparison.

By contrast, the Endpoint Coercion Hypothesis predicts that open and minimum endpoint adjectives will behave differently with minimizers. Coercion through interpolation of ENDPT is necessary for open scale (dimensional) adjectives, but not for adjectives with a lexically-supplied minimum endpoint.

The processing literature discusses online evidence of coercion as attested in other semantic domains. For instance, online studies of complement coercion report reading time and eye movement increases for coerced conditions relative to control conditions (McElree et al. 2001, Traxler et al. 2002, Traxler et al. 2005). McElree et al. compared sentences like (20) in a self-paced reading study.

(20) a. The author was starting the book in his house on the island.
    b. The author was writing the book in his house on the island.

(McElree et al. 2001: B19)

Sentences like (20a) contained verbs (start) that select an event complement but were given an entity-denoting complement (the book) These sentences required coercion of the complement into an event-denoting interpretation (i.e., ‘start to write the book’). In the control condition (20b), coercion was not necessary since an entity-selecting verb (write) was given an entity-denoting complement (the book). An increase in reading time was reported for the coerced condition relative to the control condition. I refer the reader to Pylkkänen and McElree (2006) for more discussion of coercion costs.

If the Endpoint Coercion Hypothesis is correct, we predict that sentences with minimizer-modified dimensional adjectives will incur greater processing costs than sentences with minimizer-modified minimum endpoint adjectives. This prediction is borne out in the results of the experiment presented in §3.

3 Online evidence of coercion

As discussed above, online experiments have shown that coercion leads to increased processing costs (Pylkkänen and McElree 2006). The Endpoint Coercion Hypothesis says that only open scale (dimensional) adjectives require scale coercion in order for composition with a minimizer to take place. Thus, we predict that collocations like slightly long or slightly tall will incur greater processing costs than collocations like slightly ugly or slightly boring: ugly and boring already have scales compatible with the semantics of a minimizer. If no increase is found, then the Standard Calculation Hypothesis is supported.
A moving window self-paced reading study was conducted. Participants (n=36) had not participated in any other experiments. Participants saw 48 contexts followed by target sentences of the form in (21). Participants saw only sentence from each set of four, again in order to avoid minimal pair comparison. Sentences were counterbalanced across the four conditions (slightly vs. too, negative evaluative vs. dimensional). 48 filler sentences were included. All test items were normed by 24 individuals who did not take part in the study.\(^8\)

Contexts and target sentences were identical except for the critical region.

(21)  
\[\text{Context: Mr. Richards suggested that} | \text{Ms. Smith show a film on the Massachusetts railway system} | \text{to her students during history class.} \]
\[\text{a. Ms. Smith thought that} | \text{the film was} | \text{slightly boring} | \text{to interest} | \text{her students,} | \text{so she showed} | \text{another film.} \]
\[\text{b. Ms. Smith thought that} | \text{the film was} | \text{too boring} | \text{to interest} | \ldots \]
\[\text{c. Ms. Smith thought that} | \text{the film was} | \text{slightly long} | \text{to interest} | \ldots \]
\[\text{d. Ms. Smith thought that} | \text{the film was} | \text{too long} | \text{to interest} | \ldots \]

A note on methodology is in order. I did not directly compare condition (a) to condition (c) but instead compared the difference between (a) and a control condition (b) to the difference between (c) and a control condition (d). The control conditions were included to account for lexical effects. On average, evaluative adjectives are longer than dimensional adjectives in both character and syllable count. In addition, evaluative adjectives were found to be uniformly less frequent in a survey of the Corpus of Contemporary American English (COCA) as compared to dimensional adjectives. Thus, even if an increase in reading time were incurred by slightly-modified dimensional adjectives, it could be obscured by the length and frequency-induced increase in reading time for negative evaluative adjectives.

The too...to condition was the control condition. A to-phrase is a felicitous continuation both for adjectives headed by too, and for adjectives modified by slightly. In §4, I call this the ‘goal-based standard’ reading.\(^9,10\)

---

\(^8\)In the norming study, there was no difference between the too...to conditions (b) and (d), indicating that the difference in reading times given in Table 3 was not due to condition (b) receiving an exceptionally long reading time due to implausibility.

\(^9\) We will also see in §4 that the goal-based standard is dispreferred for negative evaluative adjectives. Thus, the use of a to-phrase as a continuation (forcing a goal-based standard) gives a maximal ‘advantage’ to slightly-modified dimensional adjectives in the continuation region.

\(^10\) I chose to use negative evaluative adjectives as opposed to minimum endpoint absolute adjectives in order to avoid a possible confound. A supporter of the Standard Calculation Hypothesis might counter that the longer reading times for dimensional adjectives were due to STND having to calculate a context-determined standard of comparison rather than simply select the minimum endpoint. The minimum endpoints of negative evaluative adjectives, however, seem to be more easily contextually manipulated (compare: ‘Mary is short / ugly for a model’ vs. #’This
It was important that the target and control conditions had identical continuation regions, since previous research suggests that reading time effects are likely to appear in the continuation region. McElree et al. (2001) report a delay in reading time effects to the continuation region for complement coercion, and the continuation region was the locus for increased processing costs in the Frazier et al. (2008) study.

Returning now to the predictions and results of the experiment, if minimizers require coercion of only dimensional adjectives, then the difference between (a) \((slightly\ boring)\) and (b) \((too\ boring)\) should be less than the difference between (c) \((slightly\ long)\) and (d) \((too\ long)\). If there is no difference, then the Standard Calculation Hypothesis seems more suitable:

\[
(slightly\ NegEval\ -\ too\ NegEval)\ <\ (slightly\ Dimen\ -\ too\ Dimen)
\]

**Figure C**: Predicted Critical Interaction under the Endpoint Coercion Hypothesis

The inequality represented in Figure C was attested. Table 3 shows the mean reading times (ms) for the continuation region by condition.

<table>
<thead>
<tr>
<th>Condition:</th>
<th>Sl. NegEval (a)</th>
<th>Too NegEval (b)</th>
<th>Sl. Dimen (c)</th>
<th>Too Dimen (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean RT:</td>
<td>548.7</td>
<td>517.2</td>
<td>548.9</td>
<td>472.5</td>
</tr>
<tr>
<td>Difference in RT:</td>
<td>31.5 ms</td>
<td>&lt; 76.4 ms</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3**: Mean readings times (ms) in the continuation region

The increase in reading times can be phrased in terms of an interaction between the factors of adjective type and modified vs. unmodified status. A Linear Mixed Effects Model was used. Fixed effects were adjective type (dimensional vs. negative evaluative) and modification type \((too\ vs.\ slightly)\). Random effects were subjects and items. The base line condition was dimensional adjectives modified by \(too\). The interaction of adjective type and modification type relative to this base line was found to be fully significant \((pMCMC = 0.0346)\). That is, the increase observed for collocations like \(slightly\ long\) was significantly larger than predicted given the model.\(^{11}\)

The results of the experiment provide support for the Endpoint Coercion Hypothesis. A plausible source of the increased processing costs is minimizer-driven coercion of otherwise open adjectival scales. The absence of a similar

\(^{11}\) The data were pruned to eliminate reading times < 200 ms and > 1500 ms. Elimination of improbably short and long reading times removed 12 responses from a total of 864 (1.39%).
increase in processing costs for negative evaluative adjectives suggests that they
do not have to undergo scale coercion but already have a minimum endpoint.

Only the Endpoint Coercion Hypothesis predicts this difference between
negative evaluative and dimensional adjectives. Under the Standard Calculation
Hypothesis, minimizers are semantically compatible with both negative
evaluative and dimensional adjectives, so no increase in reading time for
minimizer-modified dimensional adjectives is predicted.

4 The meaning(s) of slightly tall
I have proposed that scale coercion occurs through interpolation of semantic
structure that creates a minimum endpoint on an otherwise open scale. I propose
the following semantics for the interpolated structure, referred to as ENDPT.

\[
[[\text{ENDPT}]] = \lambda G_{<d, \text{et}} \lambda d_{d} \lambda x_{e}. G(d)(x) \land d \geq d_{s}
\]

ENDPT takes an adjective \( G \) and returns \( G \) with a delimited scale such that all
degrees \( d \) on the scale associated with \( G \) exceed a contextually determined degree
\( d_{s} \). Delimitation of the scale creates a minimum endpoint scale identical to the
(lexically-supplied) scales of negative evaluative and minimum endpoint absolute
adjectives. In this section, I address the possible values of \( d_{s} \) as evidenced by the
readings of minimizer-modified dimensional adjectives.

Minimizer-modified dimensional adjectives receive two different paraphrases:

\[
(23) \quad \text{a. Group-based standard: The Olympic commentator observed that the top-ranked US gymnast was slightly tall (for a gymnast).}
\]

\[
(23) \quad \text{b. Goal-based standard: The auditioning coach denied Gabby a spot on the team because she thought that Gabby was slightly tall (to be on the gymnastics team).}
\]

The paraphrase labels in (23) refer to the value of \( d_{s} \) and use the terminology
of Sassoon (2011). In the group-based standard paraphrase, \( d_{s} \) is the standard of
tallness for a gymnast in the context. In the goal-based standard paraphrase, \( d_{s} \)
is the degree of height that a person cannot exceed in order to be a gymnast.
Discussion of each type of standard is found in Kennedy (2007) and Heim (2001).

Participants (n=32) who completed the second rating task reported above were
given an exit poll containing 6 items like (24) taken from experimental materials
containing negative evaluative and dimensional adjectives. For each item, 12
participants were asked to provide a paraphrase. Paraphrases corresponding to the
group-based (a) and goal-based (b) standards were only shown after the
participant had been given an opportunity to come up with her own paraphrase.

---

12 Three forms of the exit poll were constructed and distributed evenly among participants. Items
were counterbalanced to include equal numbers of each type of adjective.
Christopher read reviews that said that the tent was slightly big. How would you paraphrase this sentence?

a. The reviews said the tent was big, but not extremely so.

b. The reviews said the tent was excessively/too big.

Table 4 shows both the totals by paraphrase and adjective type, as well as the percentage comprised by each total. Where participants were indecisive, both responses were included in the total.\(^{13}\)

<table>
<thead>
<tr>
<th>Adjective type: Standard:</th>
<th>Neg. evaluative</th>
<th>Dimensional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group-based</td>
<td>54 (90%)</td>
<td>70 (58.8%)</td>
</tr>
<tr>
<td>Goal-based</td>
<td>4 (6.7%)</td>
<td>47 (39.4%)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (3.3%)</td>
<td>2 (1.7%)</td>
</tr>
</tbody>
</table>

Table 4: Raw total and percentages of paraphrases permitted by adjective type

The group-based standard interpretation was overwhelmingly preferred by participants for negative evaluative adjectives (90% vs. 6.7%). By contrast, responses were split more evenly for dimensional adjectives (58.8% vs. 39.4%). In addition, both paraphrases were more frequently permitted for dimensional adjectives. These findings are consistent with an endpoint coercion account, in which coercion gives rise to an endpoint whose value can vary depending on context. By contrast, the absence of variation in preferred paraphrase for negative evaluative adjectives can be attributed to their lexically-supplied minimum endpoint. The minimum endpoint of a negative evaluative adjective is the contextual standard (Bierwisch 1989), and this is the degree that the subject must exceed. This configuration corresponds to the group-based standard paraphrase.

This experiment highlights two ways in which the present proposal differs from the proposals of Bylinina (2011), Solt (2011), and Sassoon (2011). First, Bylinina and Solt only recognize goal-based readings in English. The results of the paraphrase task presented here counter this generalization. Second, while Sassoon recognizes both paraphrases, she proposes that the standard of comparison must be overtly named by a to- or for-phrase. While overt specification of the standard disambiguates \(d_s\), absence of a standard phrase did not prevent participants from assigning an interpretation to minimizer-modified dimensional adjectives. Furthermore, the high ratings reported in Table 2 were given to minimizer-modified dimensional adjectives without an overt standard.

I defined \(\text{ENDPT} (22)\) such that the coerced minimum endpoint (\(d_s\)) is contextually valued as either a goal- or group-based standard. It is interesting that the value of \(d_s\) can apparently only correspond to these two types of standard. In

\(^{13}\) Again, the experimental materials included positive evaluative adjectives which I do not discuss due to space limitations. I refer the reader to Bogal-Allbritten (2011).
future work, we should ask why other conceivable standards (e.g., than-phrase) are excluded. While I leave concrete proposals to future research, I note that work on other adjectival constructions reports group- and goal-based standards to arise to the exclusion of standards. For instance, pos-marked adjectives. Fleisher (2011) discusses sentences like ‘Middlemarch is a long book to assign,’ where the pos-marked adjective long receives a goal-based standard interpretation rather than its usual group-based standard interpretation.

5 Conclusions and further work

5.1 Conclusions
The central problem of the paper was that in offline ratings, minimizers were as felicitous with open scale (dimensional) adjectives as they were with adjectives with minimum endpoints (minimum endpoint absolute, negative evaluative). I considered two hypotheses that address this challenge: the Standard Calculation Hypothesis and the Endpoint Coercion Hypothesis. Citing increased processing costs observed in an online self-paced reading experiment, I argued that the Endpoint Coercion Hypothesis best accounts for the felicity of minimizer-modified dimensional adjectives. I formalized the account of coerced semantic structure (ENDPT) such that we predict the availability of the attested goal- and group-based standard paraphrases.

In addition to adding processing evidence to the growing body of literature on the semantics of minimizers and their interaction with dimensional adjectives, this paper adds to the broader goal of bringing experimental evidence to bear on theories of adjectival meaning (Frazier et al. 2008, Sassoon and Zavakhina 2012). The studies reported here also support the theory that negative evaluative adjectives have minimum endpoints (Bierwisch 1989, Rett 2008).

5.2 Directions of further work
There are number of issues that deserve further exploration. First, the proposal for ENDPT was made such that the coerced endpoint can either be valued as a goal- or group-based standard. We might predict similar interpretations to arise for minimizer-modified dimensional adjectives cross-linguistically. Given the denotation of ENDPT, we expect to find both goal- and group-based standards serving as the value of the coerced minimum endpoint. However, it seems that the possible effects of coercion are more restricted in other languages. For Russian, Kagan and Alexeyenko (2011) only identify the goal-based reading for dimensional adjectives modified by the minimizer affix –ovat:

\[
(26) \quad \text{Velik-ovat-yj} \\
\text{big-OVAT}
\]

‘A little bit too big.’ \textbf{Not}: ‘Big, but not excessively so.’

By contrast, Sawada (2011) only identifies the group-based reading for Japanese
If endpoint coercion is indeed at work in both languages, then it appears that languages may impose different restrictions on the contribution of endpoint coercion. How can we encode language-specific restrictions on coercion?

Second, as noted in the discussion of the offline rating studies (§2.2.2), I also tested the felicity of minimizer-modified positive evaluative adjectives (e.g., pretty, friendly, smart). These results are reported in Bogal-Allbritten (2011). Collocations like *slightly pretty received lower ratings than collocations like *slightly tall. The proposal does not obviously preclude application of ENDPT to the pretty scale, but the low offline ratings suggest that something still prevents felicitous composition of the minimizer and pretty. I present suggestions in Bogal-Allbritten (2011), but leave focused study to future work.

Finally, how does adjectival coercion relate to other types of scalar coercion? For example, verbal coercion: adjectives and verbs have been associated with degree and event scales, respectively (Kennedy and Levin 2008). Piñango et al. (1999) consider the coercion of iterative meanings from semelfactives. More recently, Brennan and Pylkkänen (2010) discussed the coercion of inchoative meanings from stative verbs (28):

(28) a. The boy was asleep. Stative
    b. Within two minutes, the boy was asleep. Coerced Inchoative
       (Brennan & Pylkkänen 2010: 782)

The coercion of a stative verb’s scale (an open scale) to a delimited portion of the scale (i.e., inchoative meaning) is at least impressionistically relatable to the conversion of open adjectival scales by ENDPT into scales with minimum endpoints. As research in both domains continues, it may be interesting to compare the processing of adjectival scales to the processing of aspectual scales.

References