Negating Gradable Adjectives

Abstract: In this short paper, I analyze the syntax and semantics of the prefix *un-* with gradable adjectives like *unhappy* and compare it to the syntax and semantics of *not*. Within the framework of Collins and Postal 2014, I propose that *un-* and *not* have the same semantics but negate different constituents, accounting for differences in interpretation.

Key Words: gradable adjectives, negation, degree quantification, quantifier domain restriction, litotes, Klima tests

1.   Introduction
Consider the following sentences:

(1)   a.   John is unhappy.
        b.   John is not happy.

In both cases there is a negative morpheme, *un-* in (1a) and *not* in (1b). And furthermore, the two sentences overlap in truth conditions. It seems that (1a) entails (1b) but not vice a versa. Certainly, if John is unhappy, we can conclude that he is not happy. But if he is not happy, he may not be unhappy either (but somewhere in the middle of the scale of happiness which ranges from very unhappy to very happy).

I propose that both *un-* and *not* are negative morphemes of the category NEG (see Collins and Postal 2014: 19 for a slightly different analysis). But in (1a) *un-* modifies the adjective while in (1b) *not* modifies a covert degree quantifier phrase. I show how the difference in interpretation between (1a) and (1b) follows from this structural assumption.

The main goal of this paper is to present a formal compositional analysis of *un-* within the general framework for the analysis of the syntax and semantics of negation in Collins and Postal 2014. For the sake of brevity, I leave out a discussion of many related issues (e.g., scalar antonyms). It is no doubt true that my proposals bear on the analysis of scalar antonyms, but it would increase the length of the paper greatly to try to analyze them here. Nor am I trying to argue for the superiority of my particular syntactic or semantic framework to other frameworks. It would be of great interest to see whether others could take the central insights of this paper and apply them in different frameworks.

This paper focusses on gradable adjectives like *happy* (e.g., *kind*, *friendly*, *expensive*, *attractive, pleasant, tidy, successful, intelligent, qualified, selfish, helpful, cooperative, clear, healthy, comfortable*). The class of adjectives I am looking at have the following properties: (a) Adj can be modified by *very*, (b) Adj can be negated with *un-*, (c) *un-*Adj can be modified by *very*. I assume my analysis also covers similar forms such as *in-* (e.g., *expensive/inexpensive, accurate/inaccurate*) (see Zimmer 1964 on the range of negative affixes in English). Whether its conclusions extend to the use of *un-* with non-gradable adjectives is left for future research (on the range of adjectives taking *un-* prefixation see Horn 2001 and Zimmer 1964), see section 11 for some preliminary remarks. Furthermore, for the sake of brevity, the paper focuses exclusively on the predicate uses of such adjectives (*John is unhappy*), even though attributive uses are also possible (*the unhappy person, a not unhappy person*), on which see Langendoen and Bever 1973. For more on the syntax of *un-*, see Kayne 2017 and De Clerq and Vanden Wyngaerd 2018.

Other approaches to *un*-prefixation in the literature include Horn 2017, Krifka 2007 and Blutner 2004. These works focus for the most part on pragmatics (e.g., the issue of pragmatic strengthening of a contradictory negation to a contrary, or the issue of the mitigating effect of double negation). The focus of my paper is rather on the semantic values of *un-* and *not*, the truth conditions of sentences involving *un-* and *not* and entailment relations between the various sentences. As far as I know, the present analysis is the first attempt to give a compositional analysis of the semantics of sentences involving *un*-prefixation.

In section 2, I introduce the Scale of Happiness which is partitioned by the two predicates *happy* and *unhappy*. Section 3 addresses a compositionality issue that arises in introducing the external argument. Section 4 discusses an alternative set of syntactic assumptions for resolving the issues brought up in section 3. Section 5 motivates a covert degree quantifier. Section 6 discusses the difference between *happy* and *unhappy*. Section 7 discusses the difference between *unhappy* and *not happy*. Section 8 shows how the truth conditions of litotes are calculated (e.g., *not unhappy*). Section 9 explains why *happy* and *not happy* are not always contradictory. Section 10 discusses double negation: *not not happy* (versus *happy*). Section 11 shows how the Klima tests provide support for the proposed analysis. Section 12 is the conclusion.

2.   Scale of Happiness
A standard way to define the semantic value of a gradable adjective is as follows (see Kennedy and McNally 2007: 349):

(2)   ⟦happy⟧   = $\lambda d \lambda x . x$ is happy to degree d

This means that *happy* takes two arguments, a degree and an individual, and is true if the individual x is happy to degree d. I will modify this semantic value below in light of facts concerning *un-* modification.

Whether *un-* combines with an adjective or a whole adjective phrase is not relevant to the present paper:

(3)   a.   [NEG ADJ]
       b.   [NEG ADJP]

Since *un-* is NEG, I assume that its semantics is given by the semantics of negation in Collins and Postal 2014:

(4)   If X has a semantic type ending in <t>, then
       NEG takes X with semantic value: $\lambda P_1 \ldots \lambda P_n$ […]
       And returns Y with semantic value: $\lambda P_1 \ldots \lambda P_n \neg$[…]

For propositional variables p (no predicate abstraction), the negation is simply ¬p. On this view, negation can combine with constituents of different types, parallel to the analysis of conjunction given in Partee and Rooth (1983). One way to think of it is that the different types partition the domain of the negation function.

Assuming the structure (3a) for convenience, and applying (4), the semantics of *un-* are given below:

(5) a. ⟦un-⟧ = λX.λd.λx.¬X(d)(x)
     b. ⟦un-happy⟧ = λdλx.¬happy(d)(x)

The problem with (5b) is that it claims that x is unhappy to degree d if it is not the case that x is happy to degree d, and hence does not distinguish (1a) and (1b). In other words, according to (5), the phrase *John is unhappy* would not actually entail that John is unhappy (but only not happy to a certain degree).
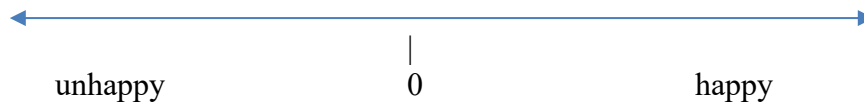
I propose instead that *happy* should be defined as a predicate of degrees. Then the semantic value of *unhappy* will be given in (6b):

(6) a. ⟦happy⟧ = λd.happy(d)
     b. ⟦un-happy⟧ = λd.¬happy(d)

Defining *happy* and *unhappy* in this manner avoids the problem noted following (5). Suppose John's degree of happiness is d, then John is unhappy iff ¬happy(d), which accurately yields the truth conditions (the truth conditions are fully spelled out in sections 5). Of course, it remains to see how the external argument is linked to particular degrees. I address that issue in the next section.

The range of degree values can be diagrammed as follows:

(7) **Scale of Happiness**

         unhappy   0   happy

Any position to the right of 0 is happy and any position to the left of 0 or equal to 0 is unhappy. Therefore, *un*-prefixation yields a contradictory, not a contrary negation (for a similar assumption, see Krifka 2007: 170 and De Clercq and Vanden Wyngaerd 2018, for *un*-prefixation as contrary negation at least in some cases, see Horn 2001: 275, 2017: 86, Sproat 1992: 349, Bierwisch 1989: 209, Jespersen 1917: 274).

This concept of unified Scale of Happiness resembles Bierwisch's (1989: section 10.2) analysis of joining scales for Pos/Neg pairs of gradable evaluative adjectives: "The intuitive notion that Pos/Neg pairs induce gradation, but in opposite directions, can now be conceived in such a way that the two scales are in a way joined at their zero point. As a condition for this, we can assume that the joined scales represent gradations regarding the same predicational domains, in other words they connect Pos-A with the appropriate Neg-A…".

For the analyses in this paper, I need to be more specific about what degrees are. Following Kennedy 2001: 52, I define a scale S to be "…a linearly ordered, infinite set of points, associated with a dimension that indicates the type of measurement that the scale represents."

Deviating from Kennedy (who does not discuss *un*-prefixation), I will define a degree as any subset of S of the following form [0, x] (x greater than or equal to 0 on the real number line) or [x, 0] (x less than or equal to zero on the real number line). Furthermore, ordering among the degrees is defined as follows (see Kennedy 2001: 54):

(8) a. d1 > d2 iff d1 ∩ d2 = d2 ∧ d1 ≠ d2
     b. d1 < d2 iff d1 ∩ d2 = d1 ∧ d1 ≠ d2

3

Concretely consider the following example:

(9)

According to the definitions in (8), d1 is greater than d2 and d4 is greater than d3. Degrees on the opposite sides of the scale are not ordered. For example, for d2 and d3, neither of the ordering relations in (8a,b) holds.

A clarification is in order. I am representing the scale of happiness on the real number line, with a 0 element. I chose this representation for three reasons. First, it allows a definition of degrees on the happy and unhappy part of the scale and how they are ordered. Second, it provides a single set to which all the degrees belong. Third, it provides an intuitive graphic representation, which I will employ throughout the paper. But the orderings referred to in the rest of the paper are not the orderings of real numbers, but rather orderings of degrees as defined in (8).

This scale makes sense of comparative constructions involving *un*-prefixation (for similar data see Kang 1993: 789, and Bierwisch 1989: 213 on "scale adjustment"). First, if both John and Bill are happy, one can say (10):

(10)   John is happier than Bill
        "The degree to which John is happy exceeds the degree to which Bill is happy."

But in the same situation, it is not possible to say (11) (# means that it is not felicitous):

(11)   #Bill is unhappier than John.
        "The degree to which Bill is unhappy exceeds the degree to which John is unhappy."

The reason for the infelicity of (11) is that neither Bill nor John is unhappy, so they are not associated with degrees in the left part of the happiness scale in (7). This pattern holds for all the gradable adjectives that can be modified with *un*- (compare to the discussion of *tall* versus *short* in Kennedy 2001: (49)).

A complication arises in that the contrast between (10) and (11) is not symmetric. Suppose that both John and Bill are unhappy, one can say:

(12)   Bill is unhappier than John.

But (12) seems to be consistent with the following:

(13)   John is happier than Bill.

In other words, if both John and Bill are unhappy, but John is less unhappy than Bill, it is possible to compare them using the adjective *happy*. It seems this pattern is fairly productive. Furthermore, suppose that Bill is unhappy and John is happy, then clearly (13) describes that situation accurately as well.

In order to account for the data in (12) and (13), I will modify the account. I suggest that *happy* has two related uses. In the first use, it characterizes degrees in the happy (positive) part of the scale. Call this use *happy*[+]. In the second use, it characterizes degrees anywhere on the scale of happiness. Call this use *happy*[*]. Degrees for *happy*[*] are defined as $(-\infty,x]$ and they are ordered according to (8). The prefix *un-* modifies only *happy*[+]. In fact, if *un-* modified *happy*[*], the result would be the empty set (since *happy*[*] is true of all degrees on the scale of happiness). A natural constraint is that *un-* cannot attach to adjectives where the resulting semantic value would be the empty set. From now on, I ignore the distinction between *happy*[+] and *happy*[*], focusing only on *happy*[+] (which I will continue to write as *happy*).

With this background, return to the semantic values of *happy* and *unhappy*. First, *happy* can be defined in terms of degrees as follows, where DEG_h is a predicate that is true of degrees on the happiness scale.

(14) a. ⟦happy⟧ = $\lambda d: DEG_h(d).\ d = [0,x]$ where $x > 0$
  b. ⟦un-⟧ = $\lambda P.\lambda d.\neg P(d)$
  c. ⟦un-happy⟧ = $\lambda d: DEG_h(d).\ d = [x,0]$ where $x \leq 0$

In this paper, I will write (14a,c) as $\lambda d.happy(d)$ and as $\lambda d.\neg happy(d)$. The following condition holds as a theorem: $\forall d[DEG_h(d) \rightarrow happy(d) \vee \neg happy(d)]$.
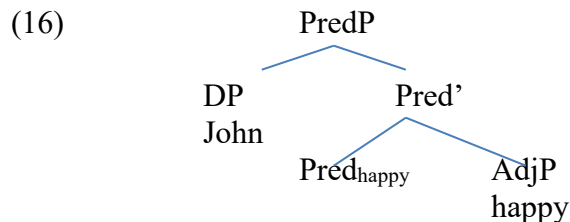
Heim 2006: (9) defines the word *little* as the negation of degree predicates, in a way very similar to my definition of the semantic value of *un-* in (14b). However, Heim focusses on the scope of *little* with respect to modal operators. She does not discuss the morpheme *un-* or any of the kinds of data discussed in this paper (e.g., litotes).

3. Adding the External Argument

Given the semantics in (14), *happy* is a predicate of a degree variable. So the question is how to incorporate an external argument (the experiencer of the emotion) as in sentences like (15):

(15) John is happy.

Following Bowers 1993 (see also Kratzer 1996 for related ideas), I propose that part of the structure of sentences like (15) is (leaving out the copula verb and TP):
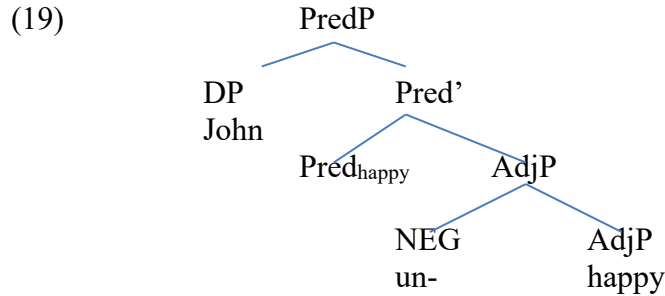
(16)

      PredP

   DP   Pred'
   John

     Pred$_{happy}$ AdjP
          happy

Given this tree, I define the semantic value of Pred in (17). degree$_h$ is a function which takes an individual and returns its position on the scale of happiness. I assume that the degree$_h$ function is defined for all sentient entities (humans, dogs, bears, martians): Everybody is happy or unhappy to some degree. That is, everybody has unique position on the happiness scale (at a particular point in time, which is not represented in the semantic values):

(17) $[\![Pred_{happy}]\!]$ = $\lambda P.\lambda x.\lambda d.degree_h(x) = d \wedge P(d)$

For example, the semantic value of the structure in (16) is calculated as follows:

(18) $[\![(16)]\!]$ = $[\lambda P.\lambda x.\lambda d.degree_h(x) = d \wedge P(d)]([\![happy]\!])([\![John]\!])$
        = $\lambda d.degree_h(John) = d \wedge happy(d)$

Consider now *unhappy* in the sentence *John is unhappy*. The tree below gives part of the structure of the sentence:

(19)

PredP
    DP   Pred'
    John
        Pred_{happy}   AdjP
            NEG   AdjP
            un-   happy

The semantic value of (19) is the following:

(20) $[\![(19)]\!]$ = $\lambda d. Degree_h(John) = d \wedge \neg happy(d)$

A reviewer notes that the definition in (17) requires a different Pred head for each gradable adjective. One possible response is to define a general $Pred_C$ where C is a contextually given scale. If the wrong scale is chosen, the result will be gibberish. On this definition, there is a just a single Pred head, not one for every gradable adjective.

4.   Alternative Syntactic Assumptions

In the above two sections, I have defined the syntax and semantics such that *un-happy* is interpreted compositionally, using the semantic value for negation from Collins and Postal 2014. My approach entailed severing the external argument from the adjective, and introducing it with a Pred head.

A reviewer suggests an alternative that does not involve severing the external argument from the adjective, while maintaining the analysis based on the definition of negation in Collins and Postal 2014. In this section, I will discuss the reviewer's proposal.

First, the semantic value of the lexical item *happy* is defined as follows:

(21) $[\![happy]\!]$ $= \lambda y \lambda d.degree_h(y) = d$

Then to create a version of this function that is only applicable to the positive portion of the Scale of Happiness, define a null functional head Pol (polarity) with the following semantic value:

(22) $[\![+Pol]\!] = \lambda P \lambda y \lambda d. \exists x > 0[d = [0,x] \wedge P(y)(d)]$

Applied to ⟦happy⟧, this yields:

(23)   ⟦+Pol⟧(⟦happy⟧)   =   λyλd.∃x>0[d = [0,x] ∧ happy(y)(d)]

Applying *un-* to the result yields:

(24)   ⟦un⟧(⟦+Pol⟧(⟦happy⟧))   =   λyλd.¬∃x>0[d = [0,x] ∧ happy(y)(d)]

Assuming that all individuals x have a position on the Scale of Happiness, this yields:

(25)   ⟦un⟧(⟦+Pol⟧(⟦happy⟧))   =   λyλd.∃x≤0[d = [x,0] ∧ happy(y)(d)]

This semantic value is equivalent to what I gave in sections 2 and 3 above. Compare them side by side:

(26)   Sections 2 and 3:
        a.   λy.λd.[degree_h(y) = d ∧ d = [0,x] where x > 0]
        b.   λy.λd.[degree_h(y) = d ∧ d = [x,0] where x ≤ 0]

(27)   Alternative:
        a.   λy.λd.∃x>0[degree_h(y) = d ∧ d = [0,x]]
        b.   λy.λd.∃x≤0[degree_h(y) = d ∧ d = [x,0]]

It is unclear to me what sort of facts, syntactic, semantic or pragmatic, would distinguish these two alternatives. I leave this issue to further work for now. For the rest of the paper, I will use the theory from sections 2 and 3.

5.   Covert Degree Quantifier
Consider first (28a) which can be paraphrased as (28b):

(28)   a.   John is happy.
        b.   John is happy to some degree.

I suggest that (28a) involves existential quantification over degrees, as with the overt degree phrase in (28b), so that (28a) has the syntactic structure in (29a) or (29b), where caps indicate non-pronunciation.

(29)   a.   John is happy [TO [SOME DEGREE]]
        b.   John is [SOME DEGREE] happy.

See (38) and (40) below for evidence supporting the order in (29b), where the covert degree quantifier phrase precedes the adjective.
     Furthermore, I suggest that this existential quantification has a restricted domain. Normally, when one says (28a) one does not mean that John is happy to some small or insignificant degree, but rather that he is happy to some significant extent. This range of degrees of happiness is seen in expressions such as the following:

7

(30) a.   John is a tiny bit happy.
      b.   John is sort of happy.
      c.   John is somewhat happy.
      d.   John is reasonably happy.
      e.   John is happy.
      f.   John is quite happy.
      g.   John is very/really happy.
      h.   John is really very happy.
      i.   John is extremely happy.

Because of scalar implicatures, sentences like (30a-d) also imply that John is not any happier than the amount specified in the degree expression. Without any degree modification, and minimal context, (30e) falls in the middle of the range of possibilities. So I suggest that the existential quantifier in (29) is the following:

(31)   some degree greater than or equal to a contextually given degree n_1

I use n_1 to distinguish it from n_2 which will be introduced below (when talking about the unhappy part of the scale of happiness). The existential quantifier and its quantifier domain restriction can be represented as follows:
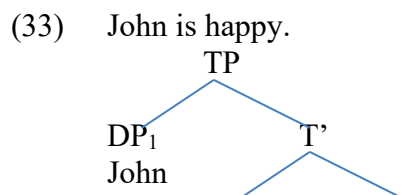
(32) $\llbracket$SOME DEG$_h\rrbracket$ = $\lambda P[\exists d\ (P(d) \wedge DEG_h(d) \wedge d \geq n_1)]$
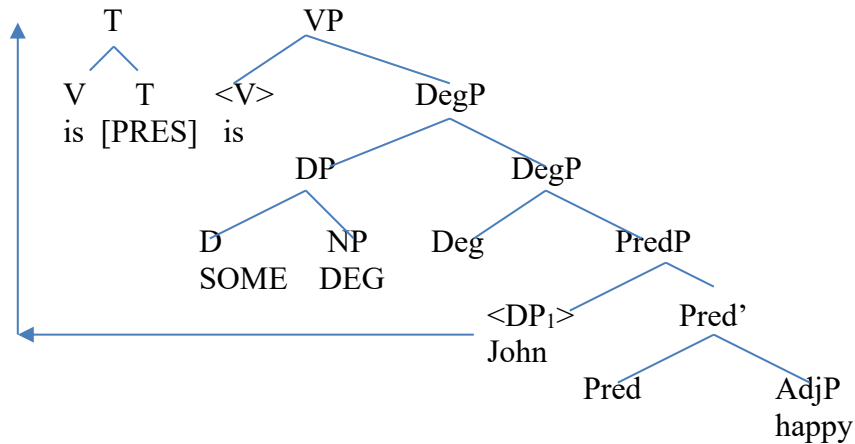
The DEG$_h$ predicate is true or false of degrees falling on the scale of happiness (see (7)), and undefined otherwise. Since all quantification in this paper is over the degrees on the scale of happiness, I will henceforth leave DEG$_h$ out of the semantic values.

An additional question, which I will not deal with here, is how the domain restriction ($d \geq n_1$) in (32) is syntactically represented. There is a large literature on this topic, and it is not relevant to the analysis in this paper (see Elbourne 2021 for discussion and references).

In the literature on gradable adjectives, the function of (32) is attributed to POS, as in Kennedy and McNally (2005: 350): "…unmodified APs actually contain a null degree morpheme *pos* (for positive form) whose function is to relate the degree argument of the adjective to an appropriate standard of comparison…". From this point of view, (32) is not an innovation, but rather corresponds to what is assumed quite generally in the literature on the semantics of adjectives.

Putting the assumptions in (16) (about adjectives) and (32) (about degree quantification) together, a tree diagram for (28a) is as follows. I assume that Spec DegP is occupied by the degree expression (and that Deg itself denotes the identity function). As usual, V to T is obligatory with finite auxiliaries. The subject DP *John* raises from Spec PredP to Spec TP.

(33)   John is happy.

T   VP

V   T   <V>   DegP

is [PRES] is

DP   DegP

D   NP   Deg   PredP

SOME   DEG

<DP₁>   Pred'

John

Pred   AdjP

happy

Given (33), the truth conditions of (28a) are as follows:

(34)   John is happy.
$[\lambda P.\exists d\ (P(d) \wedge d \geq n_1)]\ (\lambda d.\ \text{degree}_h(\text{John}) = d \wedge \text{happy}(d))$
$=\qquad \exists d(\text{degree}_h(\text{John}) = d \wedge \text{happy}(d) \wedge d \geq n_1)$

This can be paraphrased as follows: there is some degree d on the happiness scale, greater than or equal to a contextually given degree n_1, such that John is happy to degree d. These are intuitively the right truth conditions.

Quantifier domain restriction is influenced by contextual factors. For example, consider the following exchange:

(35)   a.   Is John happy with the hiring decision?
        b.   I guess you could say that, but he is not overjoyed.

(35b) says that John is happy, but n_1 is being pushed down from its usual standard position by use of the phrase 'I guess' and by the continuation, 'but he is not overjoyed'. This is a completely different value of n_1 than if the sentence *John is happy with the hiring decision* were used out of the blue. So this example shows that there is flexibility in the contextually determined value of n_1.

Krifka 2007: 172 suggests that uncertainty might also play a role in determining the use of *happy* and *unhappy*: "As a consequence of this uncertainty about the location of the border between happiness and unhappiness, the use of *unhappy* and *happy* is pragmatically restricted to those areas for which the interlocutors can assume to be in mutual agreement, to ensure safe communication." In my theory, such pragmatic factors could play a role in setting n_1. In thinking how to describe a situation, I choose a value of n_1 for which I know that my interlocutor will agree that John is happy (some robust level of happiness).

Consider now sentence (1b). Once again following Collins and Postal 2014, the negation of (32) is as follows.

(36) $[\![[\text{NEG SOME}]\ \text{DEG}_h]\!]$ $=$ $\lambda P\ \neg[\exists d\ (P(d) \wedge d \geq n_1)]$

Note that I am assuming (as in CP2014) that negation negates quantifiers directly. It would not change the analysis very much to assume a clausal NEGP. I leave it to the reader to show that the two analyses (clausal NEGP versus NEG modifying SOME) yield equivalent truth conditions for (1b).

Then (1b) has the following truth conditions (which are the negation of those in (34)):

(37) $\neg \exists d(\text{degree}_h(\text{John}) = d \wedge \text{happy}(d) \wedge d \geq n_1)$

Some syntactic evidence for a negative degree quantifier phrase in examples like (1b) is provided by the following sentence:

(38)   John is not at all happy.

Note that *at all* modifies negative DPs, free choice and negative polarity items, but not existential or universal quantifier DPs (see Horn 2000 for more on the distribution of *at all*):

(39)   a.   Nobody at all was there.
        b.   Anybody at all could do that.
        c.   I didn't see anybody at all.
        d.   *Every person at all showed up.
        e.   *Some people at all showed up.

In this paper, I do not attempt to account for the distribution of *at all*, but when it is present it has the effect of strengthening the quantification by lifting domain restrictions. So, the claim is that (38) involves a negative existential degree quantifier modified by *at all*, which has the effect of lowering n_1 to nearly 0 (see the diagram in (45)).

Another piece of evidence for a degree quantifier phrase in examples like (1b) is that the posited degree quantifier phrase sometimes appears overtly:

(40)   a.   John is not a bit happy.
        b.   John is not the least bit happy.
        c.   John is not the slightest bit happy.

In these examples, the expressions [not a bit], [not the least bit] and [not the slightest bit happy] seem to be negated occurrences of the degree quantifier phrase postulated in (33).

In the remainder of the paper I will apply the theory developed in sections 1-5 to the phrases *happy*, *unhappy*, *not happy*, *not unhappy* and *not not happy*, and the semantic differences between them.

6.   *happy* versus *unhappy*

Consider adding an existential degree quantifier to examples with *unhappy*. Since John is on the unhappy portion of the scale of happiness, I assume that n_2 (the standard value) will also be on that portion of the scale (guaranteeing that $d \geq n_2$ is defined). Since unhappy(d) holds for $d = 0$, it follows that n_2 can be zero as well. The resulting semantic value for (1a) (derived compositionally) is given below:
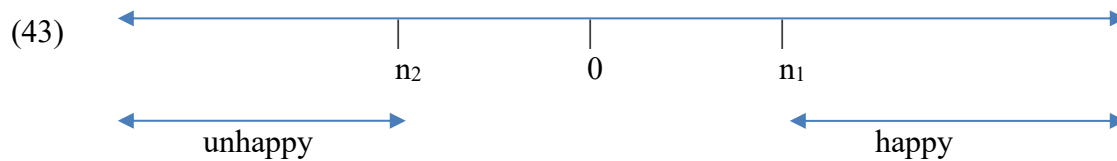
(41) $\exists d(\text{degree}_h(\text{John}) = d \wedge \neg \text{happy}(d) \wedge d \geq n_2)$

Recall that $d \geq n_2$ means that d and n_2 are in the unhappy part of the scale in (9), and that d is to the left of n_2 (that is, further away from zero, see the definitions of inequality in (8)).
        Now compare *John is happy* and *John is unhappy* in terms of these truth conditions:

(42)   a.   John is happy.
                $\exists d(\text{degree}_h(\text{John}) = d \wedge \text{happy}(d) \wedge d \geq n_1)$
        b.   John is unhappy.
                $\exists d(\text{degree}_h(\text{John}) = d \wedge \neg \text{happy}(d) \wedge d \geq n_2)$

        These truth conditions can be illustrated on the scale of happiness as follows:
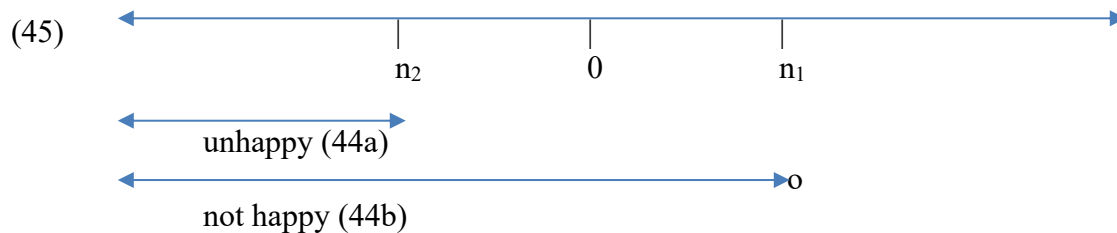
(43)

        Note that the sentences with *happy* and *unhappy* come out as contraries, since the relevant sentences cannot be true at the same time, but they can be false at the same time (where John is neither happy nor unhappy). However, the fact that sentences with *happy* and *unhappy* are contraries is not stipulated on my account, but rather derives from the presence of an existential quantifier with a domain restriction (as opposed to defining a special contrary negation, as in Horn 2017).

7.   *unhappy* versus *not happy*
        Now compare the sentences in (1) and their semantic values, repeated below:

(44)   a.   John is unhappy.
                $\exists d(\text{degree}_h(\text{John}) = d \wedge \neg \text{happy}(d) \wedge d \geq n_2)$
        b.   John is not happy.
                $\neg \exists d(\text{degree}_h(\text{John}) = d \wedge \text{happy}(d) \wedge d \geq n_1)$

        One can now see what accounts for the difference between (44a) and (44b). Consider the diagram in (45) (the o at the end of the second horizontal arrow represents not equal to):

(45)

11

The horizontal arrows under the scale in (45) are to be interpreted as follows: (44a) would be true if $degree_h$(John) were greater than or equal to n_2 (that is, to the left of n_2). Similarly, (44b) would be true if $degree_h$(John) were a member of the unhappy part of the scale, or less than n_1 (that is, to the left of n_1) on the happy part of the scale.

The diagram in (45) makes it clear that (44a) entails (44b), since the range of degrees of (44a) is a subset of the range of degrees of (44b). Furthermore, it is clear that (44b) does not entail (44a), since John might have some intermediate (not happy, not unhappy) state. The entailment relations are represented as follows:

(46)　a.　　John is unhappy ⊨ John is not happy
　　　b.　　John is not happy ⊭ John is unhappy

　　　Consider now the following sentence from Krifka 2007: 167:

(47)　a.　　I am not happy at all, in fact, I am quite unhappy.
　　　b.　　*I am quite unhappy, in fact, I am not happy at all.

The theory proposed here has no problem accounting for this contrast. The intensifier *at all* in (47a) lifts domain restrictions (see the discussion following (39)) and has the effect of reducing n_1 to 0 (or very close to 0). It says that my degree of happiness does not occupy the happy part of the scale, therefore by implication I only occupy the unhappy part of the scale. The *quite* (which is a degree expression) before *unhappy* has the effect of pushing n_2 further to the left than the standard value (see (30f) for an analogous effect on the happy part of the scale). And so the second half of (47a) is a strengthening of the first half, which is consistent with the use of the phrase *in fact*. This also explains the unacceptability of (47b), since normally the stronger sentence needs to follow *in fact*, not precede it.

Horn 2017: 87 notes that contradictory negation is often strengthened to a contrary. For example, (48a) can be continued with (48b). But spoken out of the blue, (48a) seems to imply (48c).

(48)　a.　　Alex isn't happy.
　　　b.　　…but he isn't unhappy either.
　　　c.　　Alex is unhappy.

As noted above, (48c) entails (48a). So (48c) is stronger than (48a). As discussed by Horn and others, there are many ways a weaker statement can imply a stronger one (via R-implicature). The current theory does not shed any new light on this issue, but it is consistent with the observations.

8.　　Litotes: *not unhappy*
Horn 2017 cites the OED definition of litotes as 'a figure of speech in which an affirmative is expressed by the negative of a contrary.' A typical example of litotes is an expression like that in (49):

(49)　John is not unhappy.

If both *not* and *un-* modified the adjective *happy*, then under the semantics of negation in Collins and Postal 2014, (49) should be equivalent to 'John is happy', but it is not. Rather, (49) is weaker than the positive (without double negation) (see Horn 2001: 306, Horn 2017: 89).

(50)   a.   He was happy, or at least not unhappy.
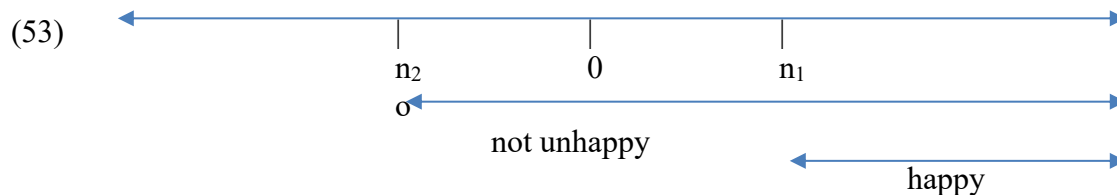   b.   #He was not unhappy, or at least happy.

In these examples, the *at least* phrase introduces the weaker alternative. If he was happy, then he was not unhappy. But if he was not unhappy, it does not follow that he was happy.
A similar contrast is the following:

(51)   a.   He was unhappy, or at least not happy.
       b.   #He was not happy, or at least unhappy.

According to the theory developed so far, (49) has the following truth conditions:

(52) $\neg\exists d(\text{degree}_h(\text{John}) = d \wedge \neg\text{happy}(d) \wedge d \geq n_2)$

This can be diagrammed as follows:

(53)

This diagram accounts for the fact that *John is happy* entails that *John is not unhappy*, since the interval covered by *happy* is a strict subset of the interval covered by *not unhappy*. It is also clearly the case that *John is not unhappy* does not entail that *John is happy* for the same reason.
I represent these entailment relations as follows:

(54)   a.   John is happy $\vDash$ John is not unhappy
       b.   John is not unhappy $\nvDash$ John is happy

Horn 2017: 82 and Krifka 2007: 164 note that the double negation has a mitigating effect: "…double negatives have the same interpretation as weak positives." While the present theory does not directly account for this observation, it is consistent with it. Even when double negation is interpreted as a weak positive, the entailment relations in (54) hold.

9.   *happy* versus *not happy*
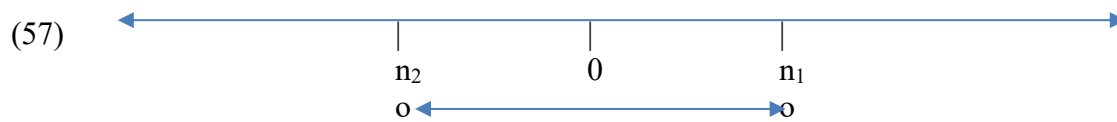
Consider (55a) and its negation (55b).

(55)   a.   John is happy.
       b.   John is not happy.

From the point of view of propositional logic, these sentences should be contradictory: HAPPY(John) is true iff ¬HAPPY(John) is not true. But they do not seem to be contradictory. They cannot both be true, but they can both be false, as shown in (56a) (thanks to a reviewer for pointing this out to me). I add (56b) for comparison.

(56) a.   John is neither happy nor not happy; he's just kind of blaah.
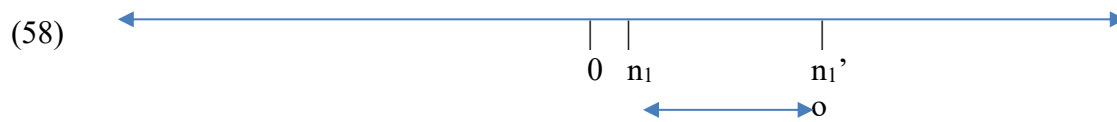   b.   John is neither happy nor unhappy; he's just kind of blaah.

I agree that (56a) is acceptable, and it seems to have roughly the same truth conditions as (56b). In my framework, these facts can be accounted for if the domain restriction on the positive degree quantifier is higher than the domain restriction on the negative degree quantifier, especially in comparisons such as (56a).
Consider first (56b), which has the following diagram:

(57)

$$\overset{\longleftarrow\ \ |\ \ \ \ \ \ \ |\ \ \ \ \ \ \ |\ \longrightarrow}{\ \ \ \ \ \ n_2\ \ \ \ \ \ \ 0\ \ \ \ \ \ \ n_1}$$
$$o \longleftarrow\ \ \ \ \ \ \ \ \ \ \ \ \longrightarrow o$$

This means that John is somewhere in the middle of the Scale of Happiness, not very happy but not very unhappy either. Right in the middle.
I analyze (56a) as involving two different domain restrictions for *happy* (both on the positive part of the Scale of Happiness):

(58)

$$\overset{\longleftarrow\ \ \ \ \ \ \ \ \ |\ |\ \ \ \ \ \ \ |\ \longrightarrow}{\ \ \ \ \ \ \ \ \ \ 0\ \ n_1\ \ \ \ \ \ n_1{}'}$$
$$\longleftarrow\ \ \ \ \ \ \longrightarrow o$$

If n_2 in (57) and n_1 in (58) are both very close to zero, it may not be possible to distinguish these two situations truth conditionally. In both cases the end result is that John's degree of happiness is somewhere in the middle between happy and unhappy.
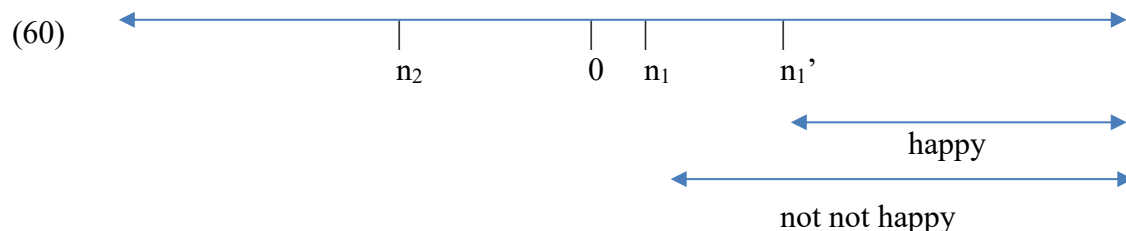
10.   *not not happy*

Given the analysis of the previous section, it is now possible to make sense of sentences like the following:

(59) a.   John is happy.
      b.   John isn't not happy.

The intuition here is that clearly (59a) is stronger than (59b), in that if (59a) is true, so is (59b), but not vice versa.
Assume that the inner *not* modifies the existential degree quantifier and the outer *not* is just sentential negation (heading NEGP), but that both have the semantics given in Collins and Postal 2014. We can understand the contrast between (59a,b) in terms of domain restriction. As noted above, the positive existential degree quantifier has a high n_1 and the negative existential degree quantifier has a low n_1. Given these assumptions, the diagrams for (59a,b) are given below:

14

(60)

In general, manipulations of n_1 and n_2 in the domain restrictions are in part syntactic, semantic and pragmatic. So, I am not claiming that my theory is a purely semantic or purely syntactic theory or purely pragmatic theory. Rather, it locates the nexus of these influences in the domain restrictions of the existential degree quantifiers.

11. Klima Tests

Another fortunate consequence of the truth conditions in (44) is that they explain why (44a) does not count as sentential negation in the sense of Klima (1964). Consider the following generalization (from Collins and Postal 2017):

(61)   A sentence S is an instance of sentential negation only if some NEG or negative quantifier DP takes widest scope in the matrix clause of S.

In (44a), the positive existential quantifier takes widest scope, and so (44a) does not count as sentential negation. In (44b), the negation takes widest scope, so (44b) counts as sentential negation. These predictions are correct, as shown below for tag-questions and negative parentheticals (see Klima 1964: 291, 316 for related data):

(62)   a.   John is unhappy, *is he?/isn't he?
         b.   John is unhappy, I think/*I don't think.
         c.   John is not happy, is he?/*isn't he?
         d.   John is not happy, I think/I don't think.

(62a,b) show that only the negative tag and positive parenthetical go with *unhappy*. (62c,d) show that only the positive tag goes with *not happy*. In contrast to (62b), both the positive and negative parenthetical are possible in (62d).

A reviewer points out that there is a similar contrast for non-gradable adjectives such as *indivisible*:

(63)   a.   73 is indivisible by 2, *is it?/isn't it?
         b.   73 is indivisible by 2, I think/*I don't think.
         c.   73 is not divisible by 2, is it?/*isn't it?
         d.   73 is not divisible by 2, I think/I don't think.

It is not the purpose of this paper to analyze non-gradable adjectives so I will not go into this contrast in depth. But an explanation similar to the one I gave for gradable adjectives can be given. Suppose that non-gradable adjectives have a state argument (parallel to the event argument, under Davidsonian views of verbal semantics, see Wellwood 2019). Existential quantification over

the state argument will take scope over negation for *un*-Adj forms, but under negation for the *not* Adj forms. I propose that this difference is what captures the difference between *un-* and *not* with respect to the Klima tests for non-gradable adjectives.

12.   Conclusion

I have analyzed *un-* and *not* as negative morphemes with the semantics of negation given in Collins and Postal 2014. The crucial difference between them is that *un-* modifies an adjective directly, whereas *not* modifies a covert degree quantifier (in the examples under consideration). I showed how the analysis accurately accounts for entailment relations between various sentences and how the Klima tests differentiate *un-* and *not*.

A crucial component of my analysis is the setting of the domain restrictions of the covert degree operator of gradable adjectives. I have shown the following factors to be relevant: (a) the speaker can adjust n by linguistic context (see (35)), (b) the expression *at all* lifts domain restrictions (see (38)) lowering n, (c) the positive existential degree quantifier has a higher n than the negative one (see sections 9 and 10).

Declarations
The author declares that he has no conflicts of interest.

References:
Bierwisch, Manfred. 1989. *The Semantics of Gradation*. In Manfred Bierwisch and Ewald Lang (eds.), Dimensional Adjectives, 71-261. Springer-Verlag, Berlin.

Blutner, Reinhard. 2004. Pragmatics and the Lexicon. In Laurence R. Horn and Gregory Ward (eds.), *The Handbook of Pragmatics*, 488-514.

Bowers, John. 1993. The Syntax of Predication. *Linguistic Inquiry* 24, 591-565.

Collins, Chris and Paul Postal. 2014. *Classical NEG Raising*. MIT Press, Cambridge.

Collins, Chris and Paul Postal. 2017. Interclausal NEG Raising and the Scope of Negation. *Glossa* 2.1, 1-29.

De Clercq, Karen and Guido Vanden Wyngaerd. 2019. Negation and the Functional Sequence. *Natural Language and Linguistic Theory* 37, 425-460.

Elbourne, Paul. 2021 Literal vs enriched meaning: It's raining. In Lisa Matthewson, Cécile Meier, Hotze Rullmann, and Thomas Ede Zimmermann (eds.), *The Wiley Companion to Semantics*, John Wiley and Sons.

Heim, Irene. 2006. *Little*. In M. Gibson and J. Howell (eds.), SALT XVI 35-58. Cornell University, Ithaca, NY.

Horn, Laurence. 2000. *any* and (-)*ever*: Free choice and free relatives. In Adam Zachary Wyner (ed.), *The Proceedings of the Fifteenth Annual Conference of the Israel Association for Theoretical Linguistics*.

Horn, Laurence. 2001. *A Natural History of Negation*. CSLI Publications.

Horn, Laurence. 2017. Lie-toe-tease: Double Negatives and Unexcluded Middles. *Philosophical Studies* 174, 79-103.

Jespersen, Otto. 1917. Negation in English and Other Languages. A. F. Høst, Copenhagen.

Kang, Beom-mo. 1993. Unhappier is Really a "Bracketing Paradox". *Linguistic Inquiry* 34, 788-794.

Kayne, Richard. 2017. Antisymmetry and Morphology. Prefixes vs. Suffixes. In Clemens Mayr and Edwin Wiliams (eds.) *Festschrift für Martin Prinzhorn, Wiener Linguistische Gazette*, 82, 145-161.

Kennedy, Christopher. 2001. Polar Opposition and the Ontology of 'Degrees'. *Linguistics and Philosophy* 24, 33-70.

Kennedy, Christopher and Louise McNally. 2005. Scale Structure, Degree Modification, and the Semantics of Gradable Predicates. *Language* 81, 345-381.

Klima, Edward. 1964. Negation in English. In *The structure of language*. Jerry A. Fodor and Jerrold J. Katz (eds.) 246-323. Englewood Cliffs, New Jersey: Prentice Hall.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In J. Rooryck and L. Zaring (eds.), *Phrase Structure and the Lexicon*, 109-137. Dordrecht: Kluwer.

Krifka, Manfred. 2007. Negated Antonyms: Creating and Filling the Gap. In Uli Sauerland and Penka Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*. Palgrave Macmillan.

Langendoen, Terence and Thomas G. Bever. 1973. Can a not unhappy person be called a not sad one? In Stephen R. Anderson and Paul Kiparsky (eds.) *A Festschrift for Morris Halle*, 392-409. New York: Holt, Rinehart and Winston.

Partee, Barbara and Mats Rooth. 1983. Generalized Conjunction and Type Ambiguity. In Paul Portner and Barbara H. Partee (eds.), *Formal Semantics: The Essential Readings*. Blackwell.

Sproat, Richard. 1992. *Unhappier* is Not a "Bracketing Paradox". *Linguistic Inquiry* 23, 347-352.

Zimmer, Karl E. 1964. *Affixal Negation in English and Other Languages.* Supplement to Word 20, Monograph No. 5.

Wellwood, Alexis. 2019. The Meaning of *More*. Oxford University Press, Oxford.