Degree expressions[1]

AD NEELEMAN, HANS VAN DE KOOT, AND JENNY DOETJES

Abstract

This paper investigates the syntax and semantics of degree expressions. We argue that some degree items are functional heads: they project when merged with a gradable expression. Other degree items are modifiers: they do not project when merged. This simple hypothesis explains a range of data, including the distribution of much. *We also distinguish various semantic classes of degree expression, based on the operations they perform and on their evaluative or non-evaluative nature. Interestingly, there is no one-to-one correspondence between the proposed syntactic and semantic classifications. This suggests that the mapping between syntax and semantics must be partially arbitrary, at least where lexical items are involved. On the other hand, the extent to which degree expressions can be stacked follows from an LF mapping rule sensitive to specifier-head configurations. This suggests more regularity in the mapping between complex syntactic and semantic objects.*

1. The syntactic realization of degree

Natural languages have expressions that may accompany APs and that indicate the degree to which the property expressed by the AP holds of its subject. Some examples are given below:

---

1. Work on this paper was originally started by Neeleman and Van de Koot. As some of the main conclusions in the paper had independently been arrived at in Doetjes' (1997) dissertation (HIL/Leiden University), it was decided to join forces and complete the paper together. We are very grateful to Arnim von Stechow for help with the proposed semantic analysis and to Chris Kennedy for numerous perceptive comments on earlier versions of this paper. All remaining shortcomings are of course our own. We also thank Peter Ackema, Annabel Cormack, John Harris, Teun Hoekstra, Richard Hudson, Villy Rouchota, Neil Smith, Fred Weerman and Jonathan White. Jenny Doetjes gratefully acknowledges the support of the Netherlands Organisation for Scientific Research (project number 355-70-003).

The Linguistic Review 21 (2004), 1–66 0167–6318/04/021-01

(1) a. *John is **too** fond of Mary.*
 b. *The weather was **very** hot in Cairo.*
 c. *John is **as** fond of Mary as Bill.*
 d. *Noone is **more** careless than Bill.*
 e. *The police searched the small room **less** carefully.*
 f. *John is good **enough** at mathematics to enter our graduate program.*

Since the bold-faced items in (1) perform a similar semantic function, the question arises whether this apparent semantic class is associated with a uniform syntactic realization. The main claim of this paper is that this is not the case: in the domain of degree expressions, syntactic realization is at least partially independent of semantic content. We begin, however, by giving an overview of some of the questions that the literature on degree expressions has dealt with.

1.1. Indirect degree systems

Perhaps the most influential analysis of degree expressions in generative syntax is that of Bresnan (1973). Bresnan assumes that every comparative contains an underlying quantifier that specifies the AP and that is itself specified by a degree word. In English, the unmarked quantifier is *much*. Thus, (1a) is given the following partial structure:

(2) $[_{AP}$ $[_{QP}$ $[_{DegP}$ *too*] $[_{Q'}$ *much*]] $[_{A'}$ *fond of Mary*]].

We will henceforth refer to analyses that require the relation between degree words and adjectives to be mediated by a quantifier as *indirect degree systems*.

One may wonder why Bresnan assumes the underlying representation in (2), given that the quantifier *much* never appears in the surface form derived from it. The reason is that there are restricted contexts in which *much* must appear. For example, as pointed out by Corver (1997a), it is obligatorily spelled out if the AP in (2) is replaced by the proform *so*:

(3) *John is very* fond of Mary. *In fact, he is* $[_{XP}$ $[_{QP}$ $[_{DegP}$ *too*] $[_{Q'}$ *(*much*)]] $[_{X'}$ so]].

Bresnan assumes on the basis of comparable data that *much* is present uniformly, but deleted if it immediately precedes an adjective. *Much* deletion presumably does not apply in (3) because *so* is categorially underspecified; it also substitutes for verbal, prepositional and nominal predicates:

(4) a. *John* loves Mary*, and Peter does* so *too.*

b. *John is still very much* on drugs*, but he is less* so *than he was as a teenager.*
c. *John is* a real gentleman – *he has always been* so.

Still, the proposed deletion rule is not sufficient to account for the distribution of *much*. As observed by Corver (1997a), *so* substitution does not license realization of *much* following all degree expressions. *As* and *very* behave like *too*, but *more*, *less* and *enough* do not:

(5) a. *The weather was* hot *in Cairo. Indeed, it was very *(much)* so.
b. *John is very* fond of Mary. *Maybe he is as *(much)* so *as Bill.*
c. *Of all the* careless *people, no one is more (*much)* so *than Bill.*
d. *The police searched the big room* carefully*, but the small room less (*much)* so.
e. *John is* good at mathematics. *He seems enough (*much)* so *to enter our graduate program.*

We will refer to items that pattern with *too* as class-1 degree expressions and to those that pattern with *more* as class-2.

In Bresnan's theory, this complication in the distribution of *much* is accounted for by analyzing *more* and *less* as suppletive forms, derived from a comparative morpheme -*er* and the quantifiers *much* and *little*, whereas *enough* is analyzed as a quantifier accompanied by a null degree expression, as in (6) (suppletion precedes *much* deletion).

(6) a. *Noone is* [$_{\text{AP}}$ [$_{\text{QP}}$ [$_{\text{DegP}}$ -er] [$_{\text{Q}'}$ *much*]] [$_{\text{A}'}$ *careless*]] *than Bill.*
b. *The police searched the small room* [$_{\text{AP}}$ [$_{\text{QP}}$ [$_{\text{DegP}}$ -er] [$_{\text{Q}'}$ *little*]] [$_{\text{A}'}$ *carefully*]].
c. *John is* [$_{\text{AP}}$ [$_{\text{QP}}$ [$_{\text{DegP}}$ $\emptyset$] [$_{\text{Q}'}$ *enough*]] [$_{\text{A}'}$ *good at mathematics*]] *to enter our graduate program.*

Although descriptively adequate, this analysis raises the question why deletion is restricted to pre-adjectival contexts. Corver (1997a, b) develops an updated version of Bresnan's theory, which shares the key property that the relation between a degree word and an adjective is mediated by a quantifier, but which accounts for the distribution of *much* in an entirely different way. The basic assumption is that both Deg and Q are heads in the extended adjectival projection:

(7) [$_{\text{DegP}}$ Deg [$_{\text{QP}}$ Q AP]]

In this structure, Deg selects for a QP, while Q combines with an AP. Contrary to Bresnan, Corver assumes that class-2 degree expressions (such as *more*, *less* and *enough*) are simplex lexical items base-generated as heads of QP, whereas class-1 degree expressions (such as *too*, *as* and *very*) are analyzed as heading DegP. The different generation sites associated with the two classes of degree expression underlie Corver's explanation of the distribution of *much.*

The analysis is based on a possible parallel with English *do* support: *much* is taken to be a dummy inserted when Q would otherwise remain empty. In the normal case, A raises to Q, raising being more economical than the language-specific rule of *much* insertion; hence the contrast between (8a,b), which extends to *as* and *very*. Trivially, *more*, *less* and *enough* are incompatible with *much* insertion, since they themselves occupy the insertion site.

(8) a. **John is* $[_{\text{DegP}}$ *too* $[_{\text{QP}}$ *much* $[_{\text{AP}}$ *fond of Mary*]]].
b. *John is* $[_{\text{DegP}}$ *too* $[_{\text{QP}}$ $[_{\text{Q}}$ *fond* Q] $[_{\text{AP}}$ t_{A} *of Mary*]]].

In the case of *so* substitution, the AP complement of Q is replaced by a pro form underspecified for category. This pro form cannot undergo head movement, with the consequence that Q will remain empty unless *much* is inserted:

(9) *John is very* fond of Mary. *In fact, he is* $[_{\text{DegP}}$ *too* $[_{\text{QP}}$ *(*much*) $[_{\text{XP}}$ so]]].

Despite its initial attraction, this analysis faces complications. It must be assumed that *so* cannot replace QP, as in (10a), although it can replace AP, as in (9), or DegP, as in (10b). If it could stand in for QP, *much* support would not be triggered, since there would be no empty Q to be filled.

(10) a. *John is very* $[_{\text{QP}}$ fond $[_{\text{AP}}$ t_{A} of Mary]]. **In fact, he is* $[_{\text{DegP}}$ *too* $[_{\text{XP}}$ so]].
b. *John is* $[_{\text{DegP}}$ too $[_{\text{QP}}$ fond $[_{\text{AP}}$ t_{A} of Mary]]] *and Bill is* $[_{\text{XP}}$ so] *too*.

To complicate matters, there is evidence that *so* can in fact replace QP. On Corver's own assumptions, both Deg-heads and contentful Q-heads are interpreted by binding a degree variable in the adjective. Thus, **too more intelligent* is ruled out as a case of vacuous quantification. This implies that *more intelligent than Bill* is a QP, not a DegP. However, *so* can substitute for it:

(11) *John is* $[_{\text{QP}}$ more $[_{\text{AP}}$ intelligent than Bill]], *and Mary is* $[_{\text{XP}}$ so] *too*.

In short, Corver is forced to stipulate that *so* can replace AP, DegP and QP, but the latter only if not dominated by DegP. It is this that allows the environment of *much* support to be derived.

1.2. Direct Degree Systems

Alternatives to Bresnan's and Corver's indirect degree systems could be based on the assumption that degree expressions can apply to adjectives without the intervention of a quantifier. A first analysis of this type was proposed by Jackendoff (1977), who argues that degree expressions are generated as specifiers of adjectives. To put it differently, the base component is organized in such a way that A in effect selects a DegP as its specifier. This structure is assumed for both class-1 and class-2 degree expressions. Hence, in neither of the examples below does Jackendoff assume the presence of an underlying *much*, thus rejecting the decomposition analysis of *more* into a comparative morpheme and a quantifier:

(12) a. *John is* $[_{AP}$ $[_{DegP}$ *too*$]$ $[_{A'}$ *fond of Mary*$]]$.
b. *Noone is* $[_{AP}$ $[_{DegP}$ *more*$]$ $[_{A'}$ *careless than Bill*$]]$.

Although Jackendoff defends a direct degree system for APs, he follows Bresnan in taking other categories to require the mediation of a quantifier: the rewrite rules for these categories determine that their specifier is a QP, not a DegP. If QP is itself adjectival, its specifier may be a DegP.

The direct degree analysis for adjectives renders *much* deletion superfluous for this category. The ungrammaticality of *$[_{AP}$ $[_{QP}$ $[_{DegP}$ *too*$]$ $[_{Q'}$ *much*$]]$ $[_{A'}$ *fond of Mary*$]]$ (cf. 2) results from the base component's inability to generate QP as a specifier of A. Similarly, the obligatory presence of *much* in *so* substitution contexts would follow on this view from the fact that *so* is not an adjective, and hence selects a QP rather than an DegP as its specifier. Thus, in (13b) DegP is not a suitable specifier for a non-adjectival head.

(13) a. *John is very* fond of Mary. *In fact, he is* $[_{XP}$ $[_{QP}$ $[_{DegP}$ *too*$]$ $[_{Q'}$ *much*$]]$ $[_{X'}$ so$]]$.
b. *John is very* fond of Mary. **In fact, he is* $[_{XP}$ $[_{DegP}$ *too*$]$ $[_{X'}$ so$]]$.

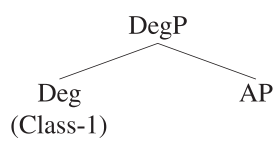
One of the attractions of Jackendoff's proposals is that the distribution of *much* is no longer accounted for by an ad hoc deletion rule, but is instead attributed to selection for category. Nevertheless, from a contemporary perspective, the analysis faces two problems, one related to the mechanism of selection, the other to the assumption that degree modification is indirect for non-adjectival categories.

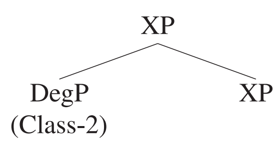
In modern terms, the process of selection, expressed in Jackendoff's framework by rewrite rules, is encoded in the lexical properties of heads. Since adjectives freely occur without degree expressions, while at least some degree expressions require the presence of an adjective, the relevant type of selection should be encoded in those degree words. But this implies that they cannot be specifiers of adjectives; rather they must be functional heads taking an AP as their complement.[2]

Moreover, since selection, as expressed by the base component, generalizes over all degree words, Jackendoff is forced to adopt an indirect degree system for non-adjectival categories. This explains why *much* shows up in (13a) but it fails to capture the fact that no overt quantifier is present when a class-2 degree expression modifies a non-adjective, as in (14a). Since *much* cannot be allowed to surface in such contexts, Jackendoff must assume a spell-out rule of the type in (14b), which is more limited in scope than Bresnan's deletion rule, but equally ad hoc.

(14) a. *John is very* fond of Mary. *In fact, he is* [$_{XP}$ [$_{QP}$ [$_{DegP}$ *more*] [$_{Q'}$ *much*]] [$_{X'}$ so] *than Bill*].
b. *more* + *much* ⇒ *more*

The present-day view of selection suggests a version of Jackendoff's approach that generalizes the direct degree system to all categories. Suppose class-1 degree expressions are functional heads selecting an adjectival complement, whereas class-2 expressions are modifiers, which do not have c-selectional requirements and therefore freely combine with any gradable predicate:

(15)


(16)


This allows both class-1 and class-2 degree expressions to combine directly with an AP:[3]

(17) a. *John is* [$_{DegP}$ [$_{Deg}$ *too*] [$_{AP}$ *fond of Mary*]].
b. *Noone is* [$_{AP}$ [$_{DegP}$ *more*] [$_{AP}$ *careless*]] *than Bill*.

---

2. Much recent work on conversion assumes that lexical categories lack categorial features (Marantz 1997, Borer 2000). Their extended projection acquires such features by containing the appropriate functional heads. This view is incompatible with our analysis, which relies on lexical heads being specified for the categorial features [±N, ±V]. Don (1993) provides, in our view, conclusive evidence for directionality in conversion and hence against the view that categorial specification is absent in lexical heads.
3. Nothing hinges on the DegP label assigned to class-2 expressions in (16). What is crucial is whether the degree word or the gradable expression projects.

Further, class-1 items select an AP and therefore cannot combine directly with the categorially underspecified pro form *so;* class-2 items lack c-selectional requirements and therefore can:

(18) a. *John is very* fond of Mary. **In fact, he is* $[_{\text{DegP}}\ [_{\text{Deg}}$ *too*$]\ [_{\text{XP}}$ so$]]$.
b. *John is very* fond of Mary. *In fact, he is* $[_{\text{XP}}\ [_{\text{DegP}}$ *more*$]\ [_{\text{XP}}$ so$]]$ *than Bill*.

The distribution of *much* can be accounted for if we follow Corver in assuming that *much* is a dummy adjective. In particular, class-1 expressions can only combine with a non-adjectival predicate if they first combine with *much*. The resulting DegP can then adjoin to this predicate in a structure identical to that of class-2 degree modification, as in (19a); *much* insertion is superfluous if a class-2 degree item attaches to a non-adjectival projection, as in (19b).

(19) a. *John is very* fond of Mary. *In fact, he is* $[_{\text{XP}}\ [_{\text{DegP}}\ [_{\text{Deg}}$ *too*$]\ [_{\text{AP}}$ *much*$]]\ [_{\text{XP}}$ so$]]$.
b. *John is very* fond of Mary. **In fact, he is* $[_{\text{XP}}\ [_{\text{AP}}\ [_{\text{DegP}}$ *more*$]\ [_{\text{AP}}$ *much*$]]\ [_{\text{XP}}$ so$]]$ *than Bill.*

In summary, the assumption that only class-1 degree words select an AP, in conjunction with an analysis of *much* as a dummy adjective, accounts for the data discussed so far without stipulations about the environments for *much* deletion or *so* substitution.

Although simple, the proposed analysis goes against an assumption shared by all previous accounts, namely that degree expressions share a uniform underlying syntax. For Bresnan and Jackendoff degree expressions are uniformly XPs that modify a gradable predicate, while for Corver they are uniformly functional heads. The present approach takes the view that some are functional heads and others XP modifiers. Let us consider whether this constitutes a shortcoming.

1.3. The issue of syntactic uniformity

The question whether there is a one-to-one mapping between semantic categories and their syntactic counterparts has been addressed in a variety of areas, perhaps most notably in θ-theory. Baker's (1988) UTAH, for example, states that every thematic relation (an essentially semantic notion) has a unique syntactic realization. The attraction of such proposals is that they radically simplify the syntax-semantics mapping.

However, in a modular theory the simplification of one component does not necessarily constitute a simplification of the theory as a whole. For example, a simplification of the transformational component may come at the expensive of a massive complication of the base component. Similarly, a simplification of the syntax-semantics mapping may require unacceptable complications of the syntax. In this respect the discussion surrounding generative semantics in the seventies is illustrative. Generative semantics assumed a completely transparent mapping between semantics and underlying syntactic structures. In fact, the underlying syntactic structure was taken to be identical to the semantic representation of a sentence. This trivial mapping led to a transformational component whose complexity was unmanageable.

There is a comparable trade-off in the domain of degree expressions. Holding on to a uniform mapping requires stipulations about *much* deletion or *so* substitution. This undermines any a priori argument for uniformity in this domain: the matter must be treated as empirical.

In what follows, we argue that a uniform analysis of class-1 and class-2 degree expressions is untenable (see also Doetjes 1997). Independently of this issue, however, the case for uniformity is severely undermined by the existence of combinations of an adjective and a degree expression that are morphological. The cases we have in mind are morphological comparatives and adjectival compounds whose lefthand part functions as a degree expression that modifies the righthand part. Dutch examples are *verliefder* 'more in love' and *straalverliefd* 'completely in love'; the latter contains the noun *straal* 'beam' as a lefthand part.

(20) a. *Fred is verliefder op Mona dan op Anne-Marie.*
Fred is in.love-er with Mona than with Anne-Marie
b. *Fred is straalverliefd op Mona.*
Fred is beam-in.love with Mona

Whereas it is possible to decompose morphological comparatives in syntax – although this is not an analysis we would advocate – morphological decomposition is highly problematic in the case of adjectival compounds. A uniform syntactic analysis of degree expressions would force one to assume that the noun *straal* is generated in the same position as other degree expressions and subsequently combined with the adjective through head-to-head movement. This is incompatible with an analysis of degree expressions as specifiers, since a specifier cannot incorporate into its head and a head cannot incorporate into its specifier. If degree expressions are taken to be heads, movement of the adjective to the noun is legitimate, but other problems arise.

First, the complex X^0 derived by incorporation is headed by the degree word, so that the resulting order should be the reverse of what is actually found, given that Dutch compounding strictly adheres to the right-hand head rule (Trommelen and Zonneveld 1986). Instead of *straalverliefd* one would expect **verliefd-straal*.

(21) $*[_{DegP}\ [_{Deg}\ \textit{verliefd}_{A}\ \textit{straal}_{Deg}]\ [_{AP} \ldots t_{A} \ldots]]$

Second, because *straal* is nominal, it cannot be part of an extended adjectival projection. But even if it could, the resulting expression would have the distribution of an NP, contrary to fact. *Straalverliefd* behaves like an adjective in every respect.

In conclusion, neither the assumption that degree expressions are uniformly specifiers, nor the assumption that they are uniformly functional heads can be straightforwardly extended to morphological degree expressions. Since some variation must apparently be allowed, nothing appears to stand in the way of a theory that extends non-uniformity to the syntactic cases.

The rest of the paper is organized as follows. In Section 2 we develop six empirical arguments for an analysis of class-1 and class-2 degree expressions as functional heads and XP modifiers, respectively. In doing so, we contrast our reinterpretation of Jackendoff's direct modification system with Corver's reinterpretation of the indirect system proposed by Bresnan. Once the syntax of class-1 and class-2 items has been established, two options present themselves. Either one adopts the non-uniform analysis proposed here or one tries to argue that a semantic distinction underlies the syntactic division. Section 3 explores the semantics of degree in some detail and concludes that there is no semantic basis for the distinction between class-1 and class-2 degree expressions. Although on the basis of a degree word's semantics no prediction can be made about its syntactic behaviour, we will show in Section 4 that there are regularities in the syntax-semantics mapping of more complex structures. In particular, stacking of degree expressions is partly conditioned by the option of deleting shared semantic material in a specifier-head configuration. Section 5 contains a summary and evaluation.

2. The syntax of degree expressions

We consider six further empirical arguments supporting the syntactic characterization of class-1 and class-2 degree expressions as heads and modifiers, respectively: (i) class-2, but not class-1, expressions attach to prepositional, verbal and nominal categories; (ii) class-2, but not class-1, expressions can have internal structure; (iii) class-2, but not class-1, expressions can appear without a gradable predicate; (iv) class-1 expressions must precede the AP they combine with; class-2 items are as a rule ordered more freely; (v) class-1 degree expressions cannot undergo topicalization but class-1 expressions can; and (vi) class-1 expressions block topicalization of the gradable predicate they combine with, whereas class-2 items do not. We also discuss two potential arguments in favour of Corver's analysis of class-2 degree items as Q heads, based on extraction and adjacency effects, and show that the assumptions underlying these arguments suffice to explain the crucial data even if no Q head is assumed.

We extend the discussion of class-1 items to *that* and interrogative *how*, which trigger *much*-support with *so*-substitution, and that of class-2 items to *a little* and *a good deal*, which do not.

(22) a. *John told me he was* afraid of spiders*, but I wonder* ***how*** *(*much*) so *he really is.*
b. *John is* fond of Mary*, but not* ***that*** *(*much*) so.
c. *John's shirt was* dirty*, but only* ***a little*** (**much*) so.
d. *John is* indebted to his colleagues*. In fact, he is* ***a good deal*** (**much*) so.

2.1. Attachment to non-adjectival projections

As argued in Grimshaw (1991) and subsequent work, all heads in an extended projection must have matching categorial features. In other words, a functional head in effect c-selects a complement of a particular category.[4] In contrast, modifiers form their own extended projection and therefore in principle freely attach to any semantically suitable category. If this view of functional heads holds true, the distinction drawn here between degree heads and degree modifiers predicts that class-1 items attach exclusively to AP, whereas class-2 items exhibit a less restricted distribution. This is in contrast to analyses along the lines of Abney (1987), Zwarts (1992), Corver (1997a,b) and Kennedy (1997), which treat all degree expressions as functional heads in the extended adjectival projection and hence predict that they all c-select an adjectival complement.

As expected, both types of degree expression attach to APs:

(23) a. *He is* [$_{\text{DegP}}$ *too* [$_{\text{AP}}$ *famous*]] *to leave town.*
b. *The door is* [$_{\text{DegP}}$ *very* [$_{\text{AP}}$ *red*]].
c. *He is* [$_{\text{DegP}}$ *as* [$_{\text{AP}}$ *intelligent*]] *as Bill.*
d. *I wonder* [$_{\text{DegP}}$ *how* [$_{\text{AP}}$ *rich*]] *he really is* t.
e. *I didn't know he was* [$_{\text{DegP}}$ *that* [$_{\text{AP}}$ *impatient*]].

(24) a. *He is* [$_{\text{AP}}$ *more* [$_{\text{AP}}$ *famous*]] *than I thought.*
b. *His paper is* [$_{\text{AP}}$ *less* [$_{\text{AP}}$ *interesting*]] *than I thought.*
c. *He is* [$_{\text{AP}}$ [$_{\text{AP}}$ *funny*] *enough*] *to be my buddy.*
d. *He is* [$_{\text{AP}}$ *a little* [$_{\text{AP}}$ *impatient*]].
e. *He is* [$_{\text{AP}}$ *a good deal* [$_{\text{AP}}$ *indebted to his colleagues*]].

---

4. Grimshaw argues that the relevant relation is not one of c-selection, but in what follows we will continue to use this term for presentational reasons.

Given that class-1 degree items can attach directly to adjectives, we expect *much*-support to be blocked by considerations of economy in examples like (23). Indeed all variants of (23) with *much*-insertion are ungrammatical. Class-2 degree expressions do not have c-selectional properties and therefore do not trigger *much*-insertion when combined with APs or with any other category. And indeed all variants of (24) with *much*-insertion are also ungrammatical.

However, when we turn to categories other than AP, we find the expected contrast between class-1 and class-2 expressions. The former require *much*-insertion when attached to PP, but economy disallows *much*-insertion with the latter:[5]

(25) a. *He is* [$_{PP}$ [$_{DegP}$ *too* [$_{AP}$ *(*much*)]] [$_{PP}$ *under scrutiny*]] *to be elected at this time*.
b. *He is* [$_{PP}$ [$_{DegP}$ *very* [$_{AP}$ *(*much*)]] [$_{PP}$ *on drugs*]] *indeed*.
c. *He is* [$_{PP}$ [$_{DegP}$ *as* [$_{AP}$ *(*much*)]] [$_{PP}$ *in the running*]] *as anyone I know.*
d. *I wonder* [$_{PP}$ [$_{DegP}$ *how* [$_{AP}$ *(*much*)]] [$_{PP}$ *into syntax*]] *he really is* t.
e. *I didn't realize that he was* [$_{PP}$ [$_{DegP}$ *that* [$_{AP}$ *(*much*)]] [$_{PP}$ *over the limit*]].

(26) a. *He is* [$_{PP}$ *more* (**much*) [$_{PP}$ *on drugs*]] *than any of his friends.*
b. *He is* [$_{PP}$ *less* (**much*) [$_{PP}$ *like his father*]] *than he used to be.*
c. *This candidate was not* [$_{PP}$ *enough* (**much*) [$_{PP}$ *in the running*]] *to stand a decent chance.*
d. *His behaviour was* [$_{PP}$ *a little* (**much*) [$_{PP}$ *over the top*]].
e. *He was* [$_{PP}$ *a good deal* (**much*) [$_{PP}$ *over the limit*]].

So, if a PP expresses a gradient property, a class-2 expression can be attached to it, but attachment of a class-1 expression is impossible. A class-1 expression has to combine with dummy *much* in order to satisfy its c-selectional requirements, after which the resulting complex expression can combine with a PP.

The contrast found with PPs expressing gradient properties can be replicated for DPs. Of course, in most of their usages DPs denote a set of individuals rather than a gradient property. However, certain DPs yield the relevant semantics if used in an appropriate context. The predicative DP *a linguist*, for

---

5. There are some PPs that may directly combine with a class-1 degree word. We follow Abney (1987: 190) in assuming that only a restricted set of idiomatic PPs can do so and that these have undergone zero-derivation to A. This process is similar to the one that derives nouns from certain idiomatic VPs, as in English *an Indian take-away* and Dutch *die irritante sta-in-de-weg* 'that irritating stand in the way' (that irritating obstacle).

instance, applies more to a person if that person has more characteristics prototypical of linguists. As (27) and (28) demonstrate, DPs used in this way cannot be selected by Deg heads but they can be modified by class-2 expressions. (There is one curious gap in the data: *a good deal* cannot modify DPs, probably for semantic reasons; see Note 23). The pattern of *much*-support found with PPs (and *so*-substitution) extends to gradable DPs: all variants of (27) with *much*-insertion are fine, whereas all such variants of (28) are ruled out by economy.

(27) a. *He is* [$_{DP}$ [$_{DegP}$ *too* [$_{AP}$ *(*much*)]] [$_{DP}$ *a scientist*]] *to care about such problems.*
b. *It's* [$_{DP}$ [$_{DegP}$ *very* [$_{AP}$ *(*much*)]] [$_{DP}$ *time for coffee*]] *now*.
c. *He is* [$_{DP}$ [$_{DegP}$ *as* [$_{AP}$ *(*much*)]] [$_{DP}$ *a typical Hollywood celebrity*]] *as Robin W*.
d. *I wonder* [$_{DP}$ [$_{DegP}$ *how* [$_{AP}$ *(*much*)]] [$_{DP}$ *the village idiot*]] *he is generally considered* t.
e. [$_{DP}$ [$_{DegP}$ *That* [$_{AP}$ *(*much*)]] [$_{DP}$ *the village idiot*]], *I never thought John was.*

(28) a. *He is* [$_{DP}$ *more* (**much*) [$_{DP}$ *a theoretical linguist*]] *than a psychologist.*
b. *This is* [$_{DP}$ *less* (**much*) [$_{DP}$ *a typical Italian opera*]] *than most of Puccini's.*
c. *He is* [$_{DP}$ [$_{DP}$ *man*] *enough* (**much*)]] *for Sue.*
d. *Always* [$_{DP}$ *a little* (**much*) [$_{DP}$ *the village idiot*]], *John was never taken very seriously.*

The action expressed by a VP can also be interpreted as gradable. When this is the case, the familiar pattern reappears: only class-2 modifiers can directly combine with a VP.

(29) a. **He* [$_{DegP}$ *too* [$_{VP}$ *likes venison*]] *for his own good.*
b. **He* [$_{DegP}$ *very* [$_{VP}$ *loves Mary*]] *indeed*.
c. **He* [$_{DegP}$ *as* [$_{VP}$ *lives like a typical Hollywood celebrity*]] *as Robin W.*
d. **I wonder how he* [$_{DegP}$ *t* [$_{VP}$ *really expects to be nominated*]. (*on a degree reading*)
e. **I didn't realize he* [$_{DegP}$ *that* [$_{VP}$ *enjoyed opera*]].

(30) a. *He* [$_{VP}$ [$_{VP}$ *likes venison*] *more* (**much*)] *than his family does.*
b. *He* [$_{VP}$ [$_{VP}$ *lives like a celebrity*] *less* (**much*)] *than he would like to.*
c. *He* [$_{VP}$ [$_{VP}$ *loves Mary*] *enough* (**much*)] *to marry her.*

d. *Peter* [$_{VP}$ [$_{VP}$ *enjoys reading*] *a little* (**much*)]; *his brother does not enjoy it at all.*

e. *He* [$_{VP}$ [$_{VP}$ *loves Mary*] *a good deal* (**much*)].

As before, the class-2 items in (30) do not tolerate *much*-insertion, while the c-selectional requirements of class-1 expressions must be satisfied by the dummy adjective:

(31) a. *He* [$_{VP}$ [$_{VP}$ *likes venison*] [$_{DegP}$ *too* [$_{AP}$ *much*]]] *for his own good.*

b. *He* [$_{VP}$ [$_{VP}$ *loves Mary*] [$_{DegP}$ *very* [$_{AP}$ *much*]]] *indeed*.

c. *He* [$_{VP}$ [$_{VP}$ *lives like a typical Hollywood celebrity*] [$_{DegP}$ *as* [$_{AP}$ *much*]]] *as Robin W.*

d. *I wonder* [$_{VP}$ [$_{DegP}$ *how* [$_{AP}$ *much*]] *he* [$_{VP}$ [$_{VP}$ *really expects to be nominated*] t].

e. *I didn't realize he* [$_{VP}$ [$_{VP}$ *enjoyed opera*] [$_{DegP}$ *that* [$_{AP}$ *much*]]].

It is generally assumed that modifiers do not impose c-selectional requirements. Hence, the data discussed in this section cannot be captured if degree expressions are uniformly modifiers (Bresnan 1973, Jackendoff 1977). The situation is slightly more complex for the proposal that all such expressions are heads (Abney 1987, Zwarts 1992, Corver 1997a,b, Kennedy 1997). If heads in an extended projection must have the same categorial features, an analysis of all degree expressions as functional heads incorrectly predicts that no degree expression can attach to a PP, DP or VP.

Alternatively, one could assume, following Marantz (1997) and others, that some functional heads may combine with complements of various categories. Corver's proposals could be interpreted along these lines if Deg heads (class-1 degree expressions) uniformly select a QP, while Q (class-2 degree expressions) may select a complement of any category:

(32) [$_{DegP}$ Deg [$_{QP}$ Q XP]]

In order to account for the data discussed above, a further assumption is necessary: Q must be filled by an adjectival category. If XP is adjectival, this is achieved by head movement of X to Q. If XP is non-adjectival, Q is filled by dummy *much*.

The proposal in (32) is less parsimonious than the non-uniform analysis pursued here. First, since it is based on the idea that functional heads may combine with complements of various categories, it is a stipulation that Deg heads select a QP whereas Q heads impose no selectional requirement. In our proposal this distinction follows directly from the syntactic status of class-1 and class-2 degree expressions, as the analysis is based on categorial uniformity in extended projections.

Second, the proposal relies on a further, apparently irreducible, stipulation that only adjectival heads may fill Q. Since Q is taken not to have any c-selectional properties, it is unclear why heads other than adjectives cannot raise to it. Yet, it is crucial that they cannot, since otherwise *much*-support would not be triggered in non-adjectival contexts.

There is also an empirical problem with the structure in (32). It implies that if *much*-support applies, *much* forms a constituent with XP rather than with Deg. In this respect, Corver's analysis differs from both the present proposal and from the analyses in Bresnan (1973) and Jackendoff (1977), all of which assume a left-branching structure. Such a structure is corroborated by the fact that the Deg head and dummy *much* can form a coordinate in a structure of coordination (see (33)) and by the fact that, at least in certain cases, they can be moved as a constituent (see (34)). The extractions in (34) are somewhat marginal, but much better than attempted movement of a non-constituent, as in **Which book about did he read Chomsky?*.

(33) a. *It is inadvisable to be* [$_{PP}$ [$_{DegP}$ [$_{DegP}$ *very* [$_{AP}$ *much*]] *or* [$_{DegP}$ *too* [$_{AP}$ *much*]]] [$_{PP}$ *into syntax*]].
b. *He is considered by most to be* [$_{DP}$ [$_{DegP}$ [$_{DegP}$ *very* [$_{AP}$ *much*]] *or* [$_{DegP}$ *too* [$_{AP}$ *much*]]] [$_{DP}$ *the village idiot*]].
c. *Does he* [$_{VP}$ *love her* [$_{DegP}$ [$_{DegP}$ *very* [$_{AP}$ *much*]] *or* [$_{DegP}$ *too* [$_{AP}$ *much*]]]]?

(34) a. *I wonder* [$_{DegP}$ *how* [$_{AP}$ *much*]] *he really is* [$_{PP}$ *t* [$_{PP}$ *into syntax*]] *as compared to phonology.*
b. *I know what they call him, but only the size of his bowie-knife will tell* [$_{DegP}$ *how* [$_{AP}$ *much*]] *he really is* [$_{DP}$ t [$_{DP}$ *the King of the Jungle*]].
c. *I wonder* [$_{DegP}$ *how* [$_{AP}$ *much*]] *he is* [$_{VP}$ t [$_{VP}$ *looking forward to his retirement*]].

Corver does not discuss data like (34), but he does note similar examples involving substitution of an AP by *so*. A relevant case is given in (35).

(35) *They say John is* fond of Mary. *But nobody knows* [$_{DegP}$ *how* [$_{AP}$ *much*]] *he really is* [$_{XP}$ t [$_{XP}$ so]].

Corver accounts for examples like (35) by assuming that they do not involve dummy *much* but rather its semantically charged variant, which forms a constituent with *how* in the specifier of QP. We think that this assumption is problematic: the embedded questions are interpreted as "to which extent …" rather than "to which high extent …". Hence, it seems that we are dealing with dummy *much*, a conclusion confirmed by the fact that the predicates in (34) and (35) do not tolerate modification by semantically charged *much*, as opposed to, say, *different* in *John is much different from his twin brother*:

(36) a. *He is *much fond of Mary / *much in the running / *much the King of the Jungle.*
b. **He much lives like a celebrity / *He lives like a celebrity much.*

In sum, (34) and (35) probably involve dummy *much* in a left-branching structure, contra the proposal in (32).

2.2. *The projection of degree expressions*

We have argued in the preceding section that the difference in distribution between class-1 and class-2 degree expression can be straightforwardly accounted for by assuming that the former are functional heads selecting an AP, while the latter are modifiers attaching to any gradable predicate. Two predictions follow from this. The first is that class-1 expressions cannot have any internal syntactic structure. The second is that class-2 expressions may or may not have internal structure.

The first prediction is clearly correct for the degree expressions under discussion. The second prediction is confirmed by at least three observations. First, two of the class-2 degree expressions we have considered, namely *a little* and *a good deal* are syntactically complex. Cross-linguistically there are numerous examples of complex class-2 expressions. Further examples from English include *a bit*, *somewhat* and *above average*. Consider *a bit*. This expression has the internal syntax of a DP. It can, for example, host prenominal modifiers such as *tiny* and *little* (*a tiny bit*, *a little bit*). As one would expect, it attaches freely to adjectival and non-adjectival categories:

(37) a. *After all these activities, his face was* [$_{AP}$ *a bit* [$_{AP}$ *red*]].
b. [$_{PP}$ *A bit* [$_{PP}$ *over the top*]], *his remarks certainly were.*
c. *Always* [$_{DP}$ *a bit* [$_{DP}$ *the village idiot*]], *John was not taken very seriously.*
d. *I'm sure John* [$_{VP}$ [$_{VP}$ *loves Mary*] *a bit*].

Theories that uniformly analyze degree expressions as functional heads cannot allow degree expressions to be complex. Hence, such analyses must deny that *a little*, *a good deal*, *a bit*, etc., are degree expressions. One could instead claim that they are measure phrases. One would then be compelled to justify a semantic division which coincides with the syntactic division between phrasal and non-phrasal modifiers, in order to avoid adopting a variant of the proposal made here. But such a semantic characterization of phrasal modifiers is unlikely to succeed, in view of the degree expressions used in dictionary entries for class-1 items. The Concise Oxford Dictionary (sixth edition), for example, contains these descriptions of the semantics for *too*, *as*, *very*, *how* and *that*:

(38) a. *too:* in a higher degree than is admissible or desirable
b. very: in a high degree
c. *as:* in the same degree
d. *how:* to what extent
e. *that:* to such a degree

The crucial observation is that the descriptions are themselves PPs, which means that they should behave as class-2 degree expressions. Indeed, as opposed to *as*, its description can attach to APs, PPs, DPs and VPs. We demonstrate this using topicalization structures to ensure that the PP is attached to the intended predicate:

(39) a. *Fond of opera in the same degree as Bill, only John can claim to be.*
b. *Into opera in the same degree as Bill, only John can claim to be.*
c. *An opera afficionado in the same degree as Bill, only John can claim to be.*
d. *Love opera in the same degree as Bill, only John does.*

This point is reinforced by our account of *much*-support, which derives a complex degree expression from a degree head (see the discussion surrounding (34) and (35)). It is therefore hardly surprising that the resulting complex expression behaves like a class-2 degree modifier, in that it can attach to PP, DP and VP:

(40) a. *As much into opera as Bill, only John can claim to be.*
b. *As much an opera afficionado as Bill, only John can claim to be.*
c. *Love opera as much as Bill, only John does.*

2.3. The position of degree items

In English, functional heads obligatorily precede their complement. Modifiers, however, show variation in their positioning vis-à-vis the category they modify. This is apparent from the distribution of VP adjuncts: some can occur preverbally and postverbally, whereas others must precede or follow the VP. The examples below are taken from Jackendoff 1972.

(41) a. *Stanley easily ate his Wheaties.*
a′. *Stanley ate his Wheaties easily.*

b. **Sam well did his work.*
b′. *Sam did his work well.*
c. *Albert is merely being a fool.*
c′. **Albert is being a fool merely.*

If it is indeed the case that class-1 items are functional heads and that class-2 items are modifiers, we predict that the former will precede AP without exception, whereas the latter will exhibit variable distribution. It is indeed the case that class-1 items cannot follow AP:

(42) a. **John is fond of Mary too for his own good.*
b. **John is indebted to his colleagues as as Bill.*
c. **John is unreliable very for a priest.*
d. **I wonder intelligent how John really is.*
e. **I didn't know John was rich that.*

The second prediction is corroborated by the variation found with class-2 degree expressions when attached to APs. Some of them may both precede and follow the predicate (*more*, *less* and *a good deal*), while others must precede (*a little*) or follow it (*enough*).

(43) a. *More fond of Mary than Bill, only John can claim to be.*
a′. *Fond of Mary more than Bill, only John can claim to be.*
b. *Less proud of his achievements than Bill, only John is.*
b′. *Proud of his achievements less than Bill, only John is.*
c. ??*Enough reliable to be a priest, only John is.*
c′. *Reliable enough to be a priest, only John is.*
d. *A little mad about his children, John certainly is.*
d′. ??*Mad about his children a little, John certainly is.*
e. *John – a good deal indebted to his colleagues – was being overpolite.*
e′. *John – indebted to his colleagues a good deal – was being overpolite.*

The distribution of class-2 degree expressions is also sensitive to the category of the modified expression. For instance, all class-2 items must follow VP, even *a little* (cf. (44a, a′)). In contrast, all class-2 items can precede PP, even *enough* (cf. (44b, b′)).

(44) a. **John a little loves Mary.*
a′. *John loves Mary a little.*
b. *John is enough into syntax to enter our PhD program.*
b′. *John is into syntax enough to enter our PhD program.*

The variable position of class-2 expressions is confirmed by the fact that the 'dictionary descriptions' of class-1 items may precede or follow AP, PP and DP but must follow VP. Compare (45) with (39).

(45) a. *⟨In the same degree⟩ fond of opera ⟨in the same degree⟩ as Bill, only John can claim to be.*
b. *⟨In the same degree⟩ into opera ⟨in the same degree⟩ as Bill, only John can claim to be.*
c. *⟨In the same degree⟩ an opera afficionado ⟨in the same degree⟩ as Bill, only John can claim to be.*
d. *⟨*In the same degree⟩ love opera ⟨in the same degree⟩ as Bill, only John does.*

The same pattern is found with class-2 degree expressions derived by *much*-support. These may either precede or follow PP and DP, but must appear to the right of VP. Compare (46) with (40).

(46) a. *⟨As much⟩ into opera ⟨as much⟩ as Bill, only John can claim to be.*
b. *⟨As much⟩ an opera afficionado ⟨as much⟩ as Bill, only John can claim to be.*
c. *⟨*As much⟩ love opera ⟨as much⟩ as Bill, only John does.*

The fact that uncontroversially phrasal degree expressions, such as those in (45), show behaviour comparable to *more*, *less* and *enough* supports an analysis of the latter as modifiers. Similarly, the fact that *as much* can appear on either side of PPs and DPs corroborates its status as a constituent, a conclusion which extends to other cases of *much*-support in prepositional and nominal contexts. It is not clear how the variable behaviour of class-2 items can be accounted for in a theory that takes them to be functional heads.

2.4. Omission of the adjective

A further prediction of the approach taken here concerns the extent to which a degree item can occur without an AP. One would not expect this possibility to be available to functional heads, as these c-select a particular lexical complement. Modifiers, however, are maximal projections and may hence be used on their own. This prediction can be tested if we turn to expressions that generally allow non-nominal subjects, such as *Black is beautiful* or *In the corner would be fine*. It turns out that class-1 degree items cannot be used in this environment, as expected:

(47) a. **Too is inadvisable.*

b. **Very might be offensive.*
c. **Half as would be acceptable.*
d. **How would be acceptable?*
e. **That would be horrible.* (on a degree reading of *that*)

Class-2 degree expressions, however, are acceptable in the same context:

(48) a. *More is not always better.*
b. *In fact, less is more.*
c. *Enough is enough.*
d. *A little is enough.*
e. *But a good deal is better.*

The same is true of complex degree expressions derived from degree heads by *much*-support or the 'dictionary descriptions' of class-1 degree heads:

(49) a. *Too much/in a higher degree is inadvisable.*
b. *Very much/in a high degree might be offensive.*
c. *Half as much/in the same degree would be acceptable.*
d. *How much/to what extent would be acceptable?*
e. *That much/to such a degree would be horrible.*

On an analysis of *more*, *less* and *enough* as functional heads, it is unclear why they pattern with clearly phrasal degree expressions rather than with degree heads.

2.5. Topicalization

Two further arguments supporting a distinction between class-1 and class-2 elements relate to Abney's (1987: 44) generalization that functional heads cannot be separated from their complement by movement. The data we use are based on topicalization in Dutch. The landing site of this movement is the specifier position of CP and hence it affects maximal projections only. It is therefore predicted that topicalization of class-2 expressions is possible, but topicalization of Deg heads is not. (50) and (51) show that there is indeed a sharp contrast in the expected direction.

(50) a. *Ik acht hem* $[_{\text{DegP}}$ *te* $[_{\text{AP}}$ *afhankelijk van zijn*
I consider him too dependent on his
vader]] *om een eigen zaak te beginnen.*
father for a own business to start

b. **Te* *acht* *ik* *hem* [$_{\text{DegP}}$ *t* [$_{\text{AP}}$ *afhankelijk* *van* *zijn*
too consider I him dependent on his
vader]] *om* *een* *eigen* *zaak* *te* *beginnen.*
father for a own business to start

(51) a. *Ik* *acht* *hem* [$_{\text{AP}}$ *minder* [$_{\text{AP}}$ *afhankelijk* *van*
I consider him less dependent on
alcohol]] *dan* *van* *andere* *drugs.*
alcohol than on other drugs

b. ?*Minder* *acht* *ik* *hem* [$_{\text{AP}}$ *t* [$_{\text{AP}}$ *afhankelijk* *van*
less consider I him dependent on
alcohol]] *dan* *van* *andere* *drugs.*
alcohol than on other drugs

In principle, *minder* in (51b) could have been moved from a VP-adjoined position. This would imply that the examples in (50) and (51) do not form a minimal pair, although (51b) would still show that *minder* is a maximal projection. We can control for the origin of *minder* by providing a context which forces the AP-adjoined reading. Suppose that we are working in a detoxification clinic for alcohol and drugs dependencies and that this clinic has developed a scale from 1 to 10 to express a patient's dependence on a particular drug. We could then felicitously say:

(52) *Hoeveel* *punten* *minder* *acht* *je* *hem* [$_{\text{AP}}$ *t* [$_{\text{AP}}$
how many points less consider you him
afhankelijk *van* *alcohol*]] *dan* *van* *andere* *drugs?*
dependent on alcohol than on other drugs

This example, and in particular the sharp contrast with (50b), shows that the movement under discussion is indeed allowed.[6]

6. The grammaticality of examples like (51) and (52) may seem surprising since they involve left branch extractions. It has often been claimed that such extractions are uniformly ruled out (cf. Ross 1967 and Emonds 1976, 1985). However, Corver 1990 argues convincingly that left branch extractions from APs (and other categories) *are* allowed, at least in principle. Some representative examples from Dutch and English are given below:

(i) a. [*Hoe* *nauw*] *zijn* *wij* [$_{\text{AP}}$ *t* *verwant* *aan* *de* *aap*]*?*
How closely are we related to the monkey

b. [*Hoe* *lang*] *is* *die* *melk* [$_{\text{AP}}$ *t* *houdbaar*]*?*
How long is that milk keepable

c. [*How badly*] *was John* [$_{\text{AP}}$ t *short of funds*]*?*

d. [*How easily*] *are these things* [$_{\text{AP}}$ t *obtainable*]*?*

Corver shows that in these examples the preposed phrase originates in the AP on a left branch. Those cases of left branch extraction that uniformly yield unacceptable results all involve

A final consequence of the proposed analysis of class-1 and class-2 expressions concerns topicalization of AP. Suppose that traces must be properly head-governed and that, in general, functional heads do not qualify as proper head governors (although they do, of course, create a minimality barrier for external government).[7] It should then be impossible to front the complement of a functional head. In other words, we expect that class-2 degree items can be stranded by topicalization of AP, whereas class-1 degree items cannot. This is true:[8]

(53) a. **Intelligent is hij* [$_{\text{DegP}}$ *te* [$_{\text{AP}}$ *t*]] *om enigszins*
intelligent is he too for more-or-less
normaal te functioneren.
normally to function

b. *Intelligent is hij* [$_{\text{AP}}$ *minder* [$_{\text{AP}}$ *t*]] *dan de*
intelligent is he less than the
gemiddelde Nederlander.
average Dutchman

We have now seen seven differences between class-1 and class-2 degree expressions. (i) Class-1 items require *much*-support when they attach to pro-forms that replace AP; class-2 items do not. (ii) Class-1 items select an AP; class-2 items can be combined with any category of the appropriate semantic type. (iii) Class-2 degree expressions can have internal syntactic structure; class-1 items cannot. (iv) Class-1 degree items precede AP; class-2 items either precede or follow the category they modify. (v) Class-2 expressions can occur as nonnominal subject in the absence of an AP; class-1 expressions cannot. (vi) Class-1 items cannot be topicalized; topicalization of class-2 items is allowed. (vii) Topicalization of AP cannot strand a class-1 item, but it can strand an item of class-2. These differences can all be related to a single factor, namely the syntactic status of the two types of degree expression as heads and modifiers, respectively.

---

extraction of a head, as in the examples in (50). There is a small residue of modifiers that cannot be extracted for ill-understood reasons, but these do not constitute more than a handful of cases.

7. For reasons we do not understand this generalization does not extend to modals and auxiliaries, which license VP-preposing.

8. Again one may wonder whether *minder* is attached to the AP or the VP in (53b). We can control for this by adding the modifier *zeker tien IQ punten* 'at least ten IQ points' which forces an AP-adjoined reading:

(i) *Intelligent is hij* [$_{\text{AP}}$ *zeker tien IQ punten minder* [$_{\text{AP}}$ *t*]] *dan de*
intelligent is he at least ten IQ points less than the
gemiddelde Nederlander.
average Dutchman

2.6. Complements of adjectives in Dutch

A compelling argument for the extended adjectival projection proposed by Corver (1997a, b) can be based on the distribution of PP-complements to adjectives in Dutch and the extractions they allow. We will now show that the same facts follow from our analysis.

Corver's (1997b) proposal, slightly adapted for reasons of presentation, is based on the following main assumptions. (i) The extended projection of AP can contain QP and DegP, with Deg selecting QP and Q selecting AP. (ii) A functional head must be filled. (iii) There are two types of adjectives in Dutch: type-A adjectives take their complement to the right only, while type-B adjectives, which Corver characterizes as typically deverbal, also allow base-generated complements to the left.[9] (iv) PPs can undergo leftward movement; this requires a contrastive or emphatic reading.

Let us now see how this theory accounts for the distribution of complements to adjectives. The adjective *trots* 'proud' is of type A. As a result, its complement can only be base-generated to its right, as in (54a); the PP therefore requires focus when it appears in a derived position to the left of the adjective, as in (54b). (The PP complement is bold faced if in a derived position.) The focus requirement is confirmed by the fact that a PP with a weak pronoun is ruled out in this position.

(54) a. *dat Jan trots op Marie / op d'r is.*
that John proud on Mary / on her is
b. *dat Jan* ***op Marie*** */ *****op d'r*** *trots is.*
that John on Mary / on her proud is

The situation is different with the type-B adjective *verliefd* 'in love', whose complement can be base-generated to its left or its right: the complement can appear to the left without focal stress.

(55) a. *dat Jan verliefd op Marie / op d'r is.*
that John in-love on Mary / on her is
b. *dat Jan op Marie / op d'r verliefd is.*
that John on Mary / on her in-love is

9. Corver assumes that type-B adjectives are systematically ambiguous between a type-A adjective and a type of adjective that also takes its complement to its right but in addition moves rightwards to an agreement head in the extended adjectival projection. This alternative is compatible with the analysis developed here, but we abstract away from it for the sake of presentational convenience.

The next set of examples involves structures in which AP is dominated by QP. The first case is that of morphological comparatives. Corver assumes that adjectives ending in the comparative morpheme must raise to Q. This implies that the adjective will surface to the left of a base-generated complement, irrespective of the adjective's type. The complement can only precede the adjective if it moves. Indeed, complements preceding morphological comparatives must be focussed:

(56) a. *dat Jan trotser / verliefder op Marie / op d'r is.*
that John more-proud / more-in-love on Mary / on her is

b. *dat Jan* ***op Marie*** */ ****op d'r*** *trotser / verliefder is.*
that John on Mary / on her more-proud / more-in-love is

In structures containing a class-2 degree expression, Q is filled, so that the adjective will remain in situ. The prediction for complements of type-A adjectives is trivial: they will be focussed, unless they follow the adjective (cf. 57a, b, c). The prediction for type-B adjectives is more complex. Since their complement can be base-generated to their left, but not external to the AP, it will require focus only if it precedes the class-2 degree expression located in Q (cf. 57a, b′, c).

(57) a. *dat Jan minder trots / verliefd op Marie / op d'r is.*
that John less proud / in-love on Mary / on her is

b. *dat Jan minder* ***op Marie*** */ ****op d'r*** *trots is.*
that John less on Mary / on her proud is

b′. *dat Jan minder op Marie / op d'r verliefd is.*
that John less on Mary / on her in-love is

c. *dat Jan op* ***Marie*** */ ****op d'r*** *minder trots / verliefd is.*
that John on Mary / on her less proud / in-love is

In structures containing a class-1 degree expression, both DegP and QP must be present. In the absence of a class-2 expression in Q, the adjective must raise to fill this position. It follows that a PP complement can only precede the raised adjective if it is focussed. It is indeed true that a PP which precedes a class-1 item must bear stress (cf. 58c). However, there is a complication in the data: whether it bears stress or not, the PP-complement of neither type of adjective can move to spec-QP or adjoin to QP (cf. 58b). This restriction cannot be due to a general prohibition against this type of movement, witness the grammaticality of examples like (57c). Rather, the relevant restriction appears to be that Deg and the head of its complement must be adjacent. This condition is not made explicit in Corver (1997b), but it seems unavoidable.

(58) a. *dat Jan te trots / verliefd op Marie / op d'r is.*
that John too proud / in-love on Mary / on her is
b. **dat Jan te* ***op Marie*** *trots / verliefd is.*
that John too on Mary proud / in-love is
c. *dat Jan* ***op Marie*** / ****op d'r*** *te trots / verliefd is.*
that John on Mary / on her too proud / in-love is

The word order in the adjectival projection can be derived on much the same assumptions even if class-2 degree expressions are modifiers rather than heads. Suppose we combine our classification of degree expressions with the following three assumptions, also required under Corver's analysis. (i) Dutch type-A adjectives take their complement to the right only, while type-B adjectives also allow complements to the left. (ii) PPs can undergo leftward movement; this requires a contrastive or emphatic reading. (iii) Deg must be adjacent to the head of its complement.

The data in (54) and (55), which do not involve degree expressions, receive the same account as under Corver's approach: type-A adjectives can only follow a moved complement, which consequently requires stress; type-B adjectives can also select a complement to the left.

The data in (56) are explained as follows. The comparative suffix *-er* is not deverbal and we therefore expect it to derive adjectives of type A. If so, it follows that its complement can only precede it if it is moved and hence focussed.

The data in (57) fall out from the assumption that class-2 degree expressions are adjoined to AP. Any complement surfacing to the left of the class-2 expression must have moved there and must therefore be focussed. When a complement surfaces between the class-2 expression and the adjective, its status depends on the nature of the adjective in question. For a class-A adjective, which selects its complement to the right, the relevant word order must be the result of movement and hence requires focus. A class-B adjective can take its complement on either side, so that it can surface between the adjoined degree phrase and the adjective without focus.

The data in (58) follow if class-1 degree expressions must be adjacent to the heads of their complements, in our analysis the head of AP. Hence, in the relevant context, a complement can only precede a type-A or type-B adjective if moved to the left of the degree expression.

To summarize, the assumptions underlying the distribution of focussed and unfocussed PP-complements to adjectives in Dutch are compatible with both Corver's and our classification of degree expressions. Hence, these data cannot provide an argument for either approach.

2.7. Extraction

A further potential argument in favour of Corver's analysis of the extended adjectival projection is that it neatly explains under which circumstances extraction is possible from the PP-complement of an adjective in Dutch. The account is based on the assumption that phrases which have undergone A'-movement are islands for extraction. On Corver's analysis, any PP to the left of a type-A adjective occupies a derived position (in the QP or DegP domain). Hence, extraction from the complement of a type-A adjective is correctly predicted to be possible only if the PP follows the adjective (in the examples below, the position of the PP complement is bold-faced if derived):

(59)	a.	*Dit*	*is*	*het*	*artikel*	*waar*	*Jan*	*trots*	t	*op*	*is.*	
		this	is	the	article	where	John	proud		on	is	
	b.	**Dit*	*is*	*het*	*artikel*	*waar*	*Jan*	**t**	***op***	*trots*	*is.*	
		this	is	the	article	where	John		on	proud	is	
(60)	a.	*Dit*	*is*	*het*	*artikel*	*waar*	*Jan*	*minder*	*trots*	t	*op*	*is.*
		this	is	the	article	where	John	less	proud		on	is
	b.	**Dit*	*is*	*het*	*artikel*	*waar*	*Jan*	*minder*	**t**	***op***	*trots*	*is.*
		this	is	the	article	where	John	less		on	proud	is
	c.	**Dit*	*is*	*het*	*artikel*	*waar*	*Jan*	**t**	***op***	*minder*	*trots*	*is.*
		this	is	the	article	where	John		on	less	proud	is
(61)	a.	*Dit*	*is*	*het*	*artikel*	*waar*	*Jan*	*te*	*trots*	t	*op*	*is.*
		this	is	the	article	where	John	too	proud		on	is
	b.	**Dit*	*is*	*het*	*artikel*	*waar*	*Jan*	**t**	***op***	*te*	*trots*	*is.*
		this	is	the	article	where	John		on	too	proud	is

The pattern is more complex with type-B adjectives, which may also take a complement to their left. Hence, the complement of this adjective is in derived position only if it precedes a degree expression. It is therefore correctly predicted to resist subextraction in exactly this environment:

(62) a. *Dit is het meisje waar Jan verliefd* t *op is.*
this is the girl where John in-love on is

b. *Dit is het meisje waar Jan* t *op verliefd is.*
this is the girl where John on in-love is

(63) a. *Dit is het meisje waar Jan minder verliefd* t *op is.*
this is the girl where John less in-love on is

b. *Dit is het meisje waar Jan minder* t *op verliefd is.*
this is the girl where John less on in-love is

c. *_Dit is het meisje waar Jan_ **t** ***op*** *minder verliefd is.*
this is the girl where John on less in-love is

(64) a. *Dit is het meisje waar Jan te verliefd* t *op is.*
this is the girl where John too in-love on is

b. *_Dit is het meisje waar Jan_ **t** ***op*** *te verliefd is.*
this is the girl where John on too in-love is

Finally, morphological comparatives are derived by movement of the adjective to a comparative morpheme in Q. PP-complements preceding a comparative adjective must hence have been moved themselves (into the QP domain) and consequently do not allow subextraction:

(65) a. *Dit is het meisje waar Jan verliefder* t *op is.*
this is the girl where John more-in-love on is

b. *_Dit is het meisje waar Jan_ **t** ***op*** *verliefder is.*
this is the girl where John on more-in-love is

Although the above constitutes a successful account of a complex pattern of extraction, it does not provide an argument for Corver's analysis of degree expressions. The data follow equally straightforwardly if class-2 degree expressions are modifiers rather than functional heads. On that view, too, any PP-complement preceding a type-A adjective has undergone movement and therefore resist subextraction. Hence, the data in (59)–(61) can be explained in essentially the same way.

Like Corver, we assume that the complement of class-B adjectives can either precede or follow the selecting head. So, the complement of a class B adjective is in a derived position only if it precedes a degree expression: adjoined to AP if it precedes a class-2 degree modifier, in spec-DegP or adjoined to DegP if it precedes a class-1 degree head. The data in (62)–(64) again follow in the same way as in Corver's analysis.

We argued earlier that the comparative morpheme derives type-A adjectives (which select their complement to the right). The data in (65) can therefore be explained on a par with those in (59). As before, the analysis mirrors Corver's, except that it does not rely on a tripartite functional structure in the extended adjectival projection. In other words, we find ourselves in the same situation as in the previous section: like the distributional evidence discussed there, the extraction data do not distinguish Corver's approach from the one defended here.

To conclude, in several domains an analysis of class-2 degree items as modifiers is more successful than an analysis in which they head a QP (see Sections 2.1–2.5). In addition, there are two domains, the distribution of and extraction from PP-complements to adjectives, which are neutral between the two accounts. Our overall conclusion must be that an account which distinguishes degree heads and degree modifiers is to be preferred.

3. The semantics of degree expressions

3.1. The semantics of gradable adjectives

The standard view of the semantics of degree distinguishes a gradable property, an ordered set of degrees (a scale) and a 'selector' that determines the degree to which the property holds of an object, usually by comparing this degree to other degrees on the scale (Seuren 1973, Kamp 1975, Cresswell 1976, Klein 1980, and many others). If this semantics was transparently realized in syntax, then one would expect degree expressions to map onto a tripartite structure: Deg (selector), Q (the scale) and A (the gradable property), as in the indirect degree systems proposed by Bresnan and Corver.

The syntactic evidence of Section 2 suggests that structures involving a degree expression have just two components, a degree expression and a property. As for bare gradable predicates, we have found no reason to assume that they are anything else than truly bare. What kind of semantics would go with this?

Our analysis, which incorporates the insights of previous proposals (especially Klein 1980), is based on the view that adjectives (and other gradable expressions) are themselves scalar and hence need not be combined with an expression that turns them into a scale (compare Kennedy 1997, 2000a). Naturally, they may be combined with selectors that modify the scale in a variety of ways. This approach is often implemented by associating a gradable predicate with a degree argument (Cresswell 1976, Hellan 1981 and von Stechow 1984a, b). Here we explore an alternative formalization, based on second-order quantification, according to which a gradable adjective is a set of properties ordered by strength (but otherwise identical). Degree modification is selection of a property from this scale. In other words, the properties associated with Q and A in tripartite analyses are represented in a single node on the approach taken here.

One advantage of this formalization is that it explains the order in which an AP combines with a degree expression and with a subject. Since the AP denotes a set of properties, it cannot directly combine with a subject. Rather, the AP can only be used as a predicative expression if the set of properties is reduced to a single property: this is precisely what merger of a degree expression achieves. As a result, the structure in (66a) is ungrammatical, while (66b) is allowed.

(66) a. *[Degree expression [subject [AP]]]
b. [Subject [degree expression [AP]]]

Indeed, examples corresponding to (66a) are invariably unacceptable:

(67) a. **There seems to be very* (*much*) *someone happy about passing the exam.*
b. *There seems to be someone very happy about passing the exam.*

This is not necessarily true of other theories in which gradable adjectives are scalar. Kennedy (1997) treats such adjectives as functions that take an individual as input and yield a degree d_1 (a point on the scale introduced by the adjective). A degree expression introduces another degree d_2 and specifies a relation between d_1 and d_2 (for example, $d_1 > d_2$). Thus, *John is taller than Bill* is interpreted as *tall(john)* $>$ *tall(bill)*. As far as we can see, this analysis does not explain why an adjective should combine with a degree expression before it combines with a subject, rather than the other way around. In other words, although the semantic characterizations of adjective and degree expression allow the order of merger in (66b), it is not clear how the apparently universal absence of a structure like (66a) can be accounted for.[10]

A second advantage of the semantic analysis to be developed here goes back to a comparison made by von Stechow (1984a) between the sentences in (68a) and (68b).

(68) a. *John is taller than Mary.*
b. *John's height exceeds Marys's height.*

10. Indeed, Kennedy (2000b) explicitly allows a predicate to combine with an argument before it combines with a degree expression. The underlying structure of *Pug is too stinky to go to the Ritz* is as in (i).

(i) $[_{IP}$ _ *is* $[_{DegP}[_{Deg}$ *too* $[_{AP}$ Pug $[_{A'}$ *stinky*]]] *to go to the Ritz*]]

Intuitively, these sentences are different in that (68b) ascribes a property to an abstract object: John's height has the property that it exceeds Mary's height. (68a), on the other hand, ascribes a property to a concrete object: John has the property that he is taller than Mary. However, theories of degree expressions based on comparison of degrees treat the two sentences as essentially equivalent in interpretation. Our proposal entails that a predicate of a particular strength is directly applied to *John* in (68a) and the comparison, therefore, does not involve the tallness of John and the tallness of Mary, but the strength of the predicates that apply to them.

To be more concrete, our analysis is based on the idea that the semantics of *tall* consists of a set of predicates expressing a variable degree of length, ranging from zero to infinity, as in (69).[11]

(69) $[\![\text{tall}_A]\!] = \text{SCALE}_{\text{length}} = \langle \lambda x.\text{tall}_0(x), \ldots, \lambda x.\text{tall}_i(x), \ldots, \lambda x.\text{tall}_\infty(x) \rangle$, where $\lambda x.\text{tall}_i(x)$ applies to an individual x iff x is i-tall.

We can now form an expression like (70), which states that P is an element of a set of tallness-properties and that John has the property P.

(70) $\exists\, P\, [P \in [\![\text{tall}_A]\!]\ \&\ P(\text{John})]$

As just discussed, a gradable adjective like *tall*, being an ordered set of properties $\mathcal{P}$, cannot be directly applied to an individual. First, a property must be selected from the scale of tallness-properties. In the case of bare APs, we hypothesize that this is the result of existential closure. Alternatively, as we argue extensively in Section 3.2, a degree expression attached to the AP may act as an existential quantifier over the ordered set of properties.

If the degree expressed by a bare adjective is indeed fixed by existential closure, it is remarkable that the vast majority of gradable adjectives express a high degree when used in isolation. As often observed (see among many others Sapir 1944 and Bresnan 1973), in *John is tall*, John is tall to a high degree; the sentence cannot be used if John is in fact short. One might think that this is due to lexical specification, and thus part of the semantics of *tall*. This cannot be the case, however, as *tall* does not mean 'tall to a high degree' in examples like *John is four feet tall* or *John is as tall as Bill; they are both very short*.

The only plausible alternative is to attribute the high-degree interpretation of bare adjectives to the way existential closure operates: it is restricted to the higher regions of the adjective's scale. However, what counts as the higher regions of a scale is notoriously context dependent. If John is seven feet (2.10 m)

11. This view has some similarity with that developed in Klein 1980. However, the problems with Klein's theory, as discussed by von Stechow (1984a), do not carry over to our proposal. It is beyond the scope of the present article to make an explicit comparison.

tall, we might well call him tall. But a tree of the same size cannot normally be called tall. It seems, then, that existential closure is anchored on a contextually determined point on the scale $\mathcal{P}$, which for now we refer to as $P_{average}$ (for discussion, see Bartsch and Venneman 1973, Cresswell 1976, Klein 1980, von Stechow 1984a and many others); its nature is explored more fully in Section 3.3.

We represent the semantics of *tall* in *John is tall* as in (71), where $P_{average}$ is given as a parameter of the definiendum:[12]

(71) $[\![\text{tall}(P_{average})_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{D}istance\ (\text{Considerable},\ \mathcal{U}p\ (P_{average}, [\![\text{tall}_A]\!]))\ \&\ P(x)]$

The functor $\mathcal{U}p$ uses $P_{average}$ to split the ordered set introduced by the adjective into two subsets and selects the one containing stronger properties. Subsequently, the functor $\mathcal{D}istance$ constructs a singleton set whose member is a predicate P that is considerably stronger than the bottom of the scale (which is now $P_{average}$). Finally, the existential quantifier selects the property out of the set so formed. It is this property that is ascribed to the subject of the AP. (We will return to the functor $\mathcal{D}istance$ in the next section.)

This view of bare adjectives is reminiscent of many others in the literature (Cresswell 1976, Hellan 1981, Atlas 1984, Von Stechow 1984a and Bierwisch 1989), but differs from them in two respects. First, most authors do not assume operations that manipulate the scale to which existential quantification applies. Rather, the semantics is based on the direct comparison of points taken from a scale. As we argue in Section 4.2, an appropriate semantics for (stacked) degree expressions must be based on the manipulation of scales (see Seuren 1978, 1984, von Stechow 1984b and Kennedy 2000a for related discussion). Second, most authors do not assume an equivalent of the function $\mathcal{D}istance$ (but see von Stechow 1984b). Some evidence supporting the presence of this function in bare adjectives will be given later on.

How is the semantics in (71) arrived at? The lexical semantics of *tall* in (69) asserts that it consists of a set of predicates ordered by the strength of the property they express. The addition of the existential operator in (71) is due to existential closure. Closure reduces the set to a single predicate, as indicated by the lambda bound variable. However, (71) also contains the information that the selected degree must be stronger than the relevant anchor point. We propose that this restriction is added by the two interpretive rules in (72), which apply in the order given.[13] We show later that each part of the restriction inserted

---

12. A TLR reviewer expresses concern about the use of vague notions like ‘considerable’ in our semantic ontology. We share this concern, but it is not our objective to develop a formalism that allows a precise characterization of vague terms. Rather, we are interested in the functors encoded by degree expressions.

by these rules is independently required to capture the semantics of degree expressions.

(72) *Scale enrichment rule I*
SCALE$_{bare}$ → $\mathcal{U}p$ (P$_{average}$, SCALE)
Scale enrichment rule II
SCALE$_{derived}$ → $\mathcal{D}istance$ (Considerable, SCALE)

These rules can be seen as instructions concerning the interpretation of bare scalar expressions. If the grammar did not contain them, a sentence like *John is tall* would fail to be informative, as it would only express the tautology that John has some degree of length.

It would be possible to formulate a single enrichment rule that combines the two rules given in (72). However, we will see that each of the enrichment rules can be applied individually, even though in the case of base adjectives they must both apply. The reason for this is that scale enrichment rule II cannot apply to a bare scale. In other words, its input must be a scale that has already been restricted in some way. We do not know why this is the case, but we will encounter degree expressions that impose exactly the same requirement on the scale to which they apply. Finally, it seems that scale enrichment rule I cannot ever be the final operation that applies to a scale. Again, it is unclear why this is so, but it is clear that *John is tall* cannot be true if John's height is merely average or just slightly above average. It is the subsequent obligatory application of enrichment rule II that is responsible for the high degree reading of bare adjectives.

There is independent evidence that the interpretation of bare adjectives involves comparison with an anchor point: P$_{average}$ can license the presence of two types of 'satellite', namely PPs headed by *for* and infinitival clauses (see Bierwisch 1989 for a good overview of the types of satellites licensed by scalar expressions). Thus, (73) expresses that John's length exceeds the average height for a submariner.

(73) *John is tall for a submariner/to go on a submarine.*

Not all adjectives are interpreted in the way suggested by (72). For instance, the adjective *short* denotes a low rather than a high degree of length, which would seem to suggest that in this case existential closure selects a degree below P$_{average}$. If so, alternative versions of the scale enrichment rules would have to be assumed. This would raise a number of difficult questions that are best avoided. Instead, we propose, in the spirit of Sapir (1944), Von Stechow (1984b), Bierwisch (1989) and Kennedy (1997, 2000a), that *short* and *tall* both

13. As will become clear, the order stipulated in (72) is intrinsic.

introduce a scale of length, but that *short* differs from *tall* in that its semantics contains the functor *Reverse* which changes the polarity of this scale. As a result, the top of the scale is used as the bottom in any further operations that are applied. On this view, the semantics of *short* can be schematized as in (74a). If we now apply the scale enrichment rules and closure to *short*, the property selected will indeed be weaker than $P_{average}$. In an example like *John is short* the AP is interpreted as in (74b). Additional evidence for the existence of *Reverse* is presented in Note 17; for discussion of the types of adjective that may introduce a reversed scale, see Bierwisch 1989.

(74) a. $[\![short_A]\!] = \mathit{Reverse}\,([tall_A]) = \langle P_\infty, \ldots, P_i, \ldots, P_0 \rangle$,
where P_i is property that applies to an individual x iff x is i-tall.
b. $[\![short(P_{average})_{AP}]\!] =$
$\lambda x\ \exists\ P\ [P \in \mathit{Distance}\ (\text{Considerable},\ \mathit{Up}\ (P_{average},\ \mathit{Reverse}\ ([\![tall_A]\!]))\ \&\ P(x)]$

Like *tall*, *short* can be combined with satellites that help identify $P_{average}$:

(75) *Pino is short for a giraffe/to enter the high jump competition.*

In addition to existential closure and the scale enrichment rules, languages have degree expressions that lexically encode an existential operator and a specific restriction to be applied to a scalar predicate. In the presence of such an expression, the application of scale enrichment rule II is blocked. Intuitively, this is because the scale enrichment rules are last resort operations that only apply when necessary. The semantics of degree expressions involves an operation on a scale and therefore application of enrichment rule II is superfluous. As we will see, certain degree expressions may still trigger application of enrichment rule I (also as a last resort).

In the following sections we discuss the semantics of degree expressions and in particular the restrictions they introduce; the evidence for the claim that degree expressions uniformly introduce an existential operator will be presented in Section 4. We argue that degree expressions come in two types distinguished by the type of anchor point they rely on (compare Kennedy and McNally 1999). One type is evaluative. As with bare adjectives, the anchor point is related to the average degree to which a property holds of the adjective's subject. The other type is non-evaluative, in that the anchor point is an independently specified, objective, value on the adjective's scale. We arrive at the conclusion that the semantic and syntactic bifurcations of degree expressions, although both well-motivated, do not coincide. In other words, the semantic classes we identify do not underlie the syntactic distinction between class-1 and class-2 degree expressions. This provides a strong argument for a certain arbitrariness in the mapping between syntactic and semantic categories.

3.2. Non-evaluative degree expressions

Non-evaluative degree expressions come in three types: (i) measure phrases like *eight miles* in *eight miles high*; (ii) the comparative morphemes *more*, *less* and *-er*; (iii) the equative morpheme *as*. We will discuss these in turn, starting with measure phrases.

Measure phrases reduce the scale provided by the adjective by selecting a singleton subset. The measure phrase may specify the member of this set exactly (as in *eight miles high*) or may be more vague (as in *several miles high*). The strength of the property selected compares with the strength of the property at the bottom of the scale in the way specified by the measure phrase. With *several miles* this results in the formula in (76a), where $\mathcal{P}$ represents a variable over scales (sets of properties); when *several miles* applies to the bare adjective *high* we obtain (76b).[14, 15]

(76) a. $[\![\text{several miles}]\!] = \lambda\mathcal{P}\,\lambda x\ \exists P\ \exists n\ [n \geq 2\ \&\ P \in \mathcal{D}istance\,(n\ \text{miles}, \mathcal{P})\ \&\ P(x)]$

b. $[\![\text{several miles high}_{AP}]\!] = \lambda x\ \exists P\ \exists n\ [n \geq 2\ \&\ P \in \mathcal{D}istance\,(n\ \text{miles}, [\![\text{high}_A]\!])\ \&\ P(x)]$

For expository convenience we will sometimes say that the functor $\mathcal{D}istance$ takes a measure from the bottom of the scale. This should be interpreted as outlined above.

As already observed by von Stechow (1984b), the function $\mathcal{D}istance$, in conjunction with the notion of a reversed scale, predicts the contrast in (77).

(77) a. *Six feet tall.*

b. **Three centimeters short.*

Recall that *tall* and *short* contain the same scale of length but that in the case of *short* this scale is reversed. As there is such a thing as a minimal length, namely zero, but not such a thing as a maximal length, this analysis implies that *tall* but not *short* introduces a scale with a bottom. Hence, *tall* can be combined

---

14. The view of measure phrases given in the main text can be formalized by saying that measure phrases operate on the subscripts of the properties in a representation like (69). Thus, a point P_i on a scale of tallness is a predicate that corresponds to a set of individuals that are i-tall. On this view, a measure phrase like *2 meters* constructs a singleton set whose member is the set of individuals characterized by the property $(0+2\,\text{m})$-tall, where 0 is the value of tallness that characterizes the predicate at the bottom of the scale.
15. It might be possible to construct a theory without existential quantification in the case of functors that return a singleton set (such as $\mathcal{D}istance$), but in the interest of uniformity we will assume throughout that all semantic operations that reduce the scale provided by a gradable expression return a set.

with a measure phrase while *short* cannot, as it does not provide the measure phrase with a predicate that can form the basis of the procedure that constructs a singleton set.

Let us now turn to the semantics of the second class of non-evaluative degree expressions, starting with *more*. The restriction introduced by *more* is almost identical to the one inserted by scale enrichment rule I (see (72)). First, a point is used to split the ordered set expressed by the adjective. Then the subset containing stronger properties is selected. Finally, existential quantification chooses a property from this set. The difference between *more*'s restriction and the one inserted by (72) lies in the type of point used for comparison. In the case of *more* this is not the average value on the scale of the adjective, but rather some independent point of reference. Consider the semantics of (78).

(78) *John is more intelligent than Bill.*

Whereas *John is intelligent* means that John's intelligence is above a certain average value of intelligence, (78) does not imply that John is intelligent to such a degree. The only thing that is expressed is that John's intelligence exceeds that of Bill. Thus, the semantics of *more* is as in (79), in which P_{ref}, the property i-intelligent that applies to Bill, is the point of comparison.[16] The same semantics can be assigned to the comparative suffix *-er*. It is P_{ref} that licenses the satellite *than Bill* in (78). (P_{ref}, a free variable in the definition, is part of the definiendum.)

(79) $[\![\text{more}(P_{ref})]\!] = [\![\text{-er}(P_{ref})]\!] = \lambda\mathcal{P}\,\lambda x\; \exists P\; [P \in \mathcal{U}p\,(P_{ref}, \mathcal{P})\; \&\; P(x)]$

It stands to reason that the semantics of *less* is closely related to that of *more*. Consider (80).

(80) *John is less intelligent than Bill.*

The point used to split the scale (made explicit by *than Bill*) is a reference point in the sense of (79). But this time the subset on which existential quantification

16. By Bill's value on the scale of intelligence we mean his maximum value. We do not believe, however, that the semantics of this type of satellite should make reference to maximality. There is a general pragmatic license to assume that a rational speaker makes optimally relevant statements (cf. Sperber and Wilson 1986). As a result, a specified value on a scale is taken to be maximal (unless there are overriding considerations). For example, logic dictates that if John has four pounds, he also has three pounds. Yet, the sentence *John has three pounds* is normally taken to mean that he has exactly three pounds. Only certain contexts allow a different interpretation. (See below for further discussion.)

operates is the one with the weaker properties, so that the scale to which it applies must have had the function *Reverse* applied to it:[17]

(81) $[\![\text{less}(P_{ref})]\!] = \lambda\mathcal{P}\,\lambda x\ \exists P\ [P \in \mathcal{U}p\ (P_{ref}, \mathcal{R}everse\ (\mathcal{P}))\ \&\ P(x)]$

Just like *more* and *less*, *as* introduces a reference point P_{ref} that does not equal an average value. If John is as tall as Bill, this does not imply that John is tall but only that John has a value of tallness which corresponds to the (maximum) value of tallness that Bill has. Thus, *as* differs from *more* and *less* in that it does not contain the function *Up*, but instead a function expressing that the selected point equals the reference point (Atlas 1984). We will call this function *At*.[18]

---

17. The semantics of *less* could also be captured by introducing a functor *Down* which selects the subset below the reference point:

(i) $[\![\text{less}(P_{ref})]\!] = \lambda\mathcal{P}\,\lambda x\ \exists\ P\ [P \in \mathcal{D}own\ (P_{ref}, \mathcal{P})\ \&\ P(x)]$

However, the interpretation of expressions which combine *less* with a degree item containing *Distance*, such as the measure phrase *several miles*, provides a clear argument for the representation in the main text. In order to see why, consider the example in (ii).

(ii) *The helicopter is flying several miles less high than the aeroplane.*

If the formula in (i) were adopted, the semantics of *several miles less high* would be as in (iii).

(iii) $[\![\text{several miles less high}(P_{ref})_{AP}]\!] = \lambda x\ \exists P\ \exists n\ [n \geq 1\ \&\ P \in \mathcal{D}istance\ (\text{n miles}, \mathcal{D}own\ (P_{ref}, [\![\text{high}_A]\!])\ \&\ P(x)]$

We have assumed that *Distance* measures upwards from the bottom of the scale it is presented with. Hence, in (iii), it would measure from the bottom of the selected set, which is the lowest degree of height. This is obviously not the correct result: if a helicopter is flying several miles less high than an aeroplane, this does not mean that it is flying at an altitude of several miles. If *less* consists of a combination of *Reverse* and *Up*, however, *Distance* will be given a scale whose bottom value is P_{ref} and whose top is zero. Hence, *Distance* will measure from P_{ref} towards zero, as required.

18. Atlas makes use of universal quantification in order to account for the equative. In his view an equative such as *John is as tall as Peter* means that John has all values of tallness that Peter has. Suppose that Peter is 6 feet tall. Then, on this approach, he has a set of values of tallness ranging from 0 to 6 feet. Therefore, *John is as tall as Peter* is true if John shares all these values of tallness. This does not exclude, however, that John may also have values exceeding 6 feet. As a consequence, *John is as tall as Peter* is still true if John is taller than Peter, but not the other way around. Our approach obtains the same result by using existential quantification. According to the formula in (82), *John is as tall as Peter* means that John has Peter's (maximum) value of tallness. After all, having a particular value of tallness implies having all lower values. Hence, John may in fact be taller than Peter. In this sense the two approaches are equivalent.

(82) $[\![\text{as}(P_{ref})]\!] = \lambda\mathcal{P}\lambda x\ \exists P\ [P \in \mathcal{A}t\,(P_{ref}, \mathcal{P}) \ \&\ P(x)]$

Of course it is sometimes appropriate to use expressions like *as famous* in contexts in where the selected property is stronger than the reference point. This reading is highlighted in examples such as *John is certainly as famous as Bill. In fact, I suspect he is a lot more famous.* Some authors have incorporated this into the semantics of *as*, taking it to mean 'equal to or more than' (Cresswell 1976, Klein 1980 and Horn 1989). Note, however, that the phenomenon under discussion can be observed more generally with expressions that specify a quantity or strength. For instance, *John has three pounds* can in some contexts be used if John in fact has four pounds (for example in *John certainly has three pounds – he just received his salary*; cf. Grice 1989 and others). Therefore, the crucial distinction between degree expressions that introduce the function $\mathcal{U}p$ and those that introduce $\mathcal{A}t$ is that only the latter allow a reading in which the strength of the selected property equals that of the reference point: *John is more famous than Bill* cannot be used if John and Bill are in fact equally famous. Note that $\mathcal{D}istance$ is similar to $\mathcal{A}t$ in this respect, as illustrated by the following example: *John is certainly six feet tall. In fact I suspect he is even taller than that*. This is exactly what we expect: $\mathcal{D}istance$ shares with $\mathcal{A}t$ that its application results in a singleton set.

That expressions involving *as* allow a stronger interpretation than expected is the result of the logical implication that if one has a particular property to a certain degree, one also has it to a lesser degree. Hence, *as famous as Bill* means by implication 'at least as famous as Bill', just as *to have three pounds* means 'to have at least three pounds'. This explains an observation brought to our attention by Annabel Cormack (p.c.), namely that *not as famous as Bill* does not mean 'either more or less famous than Bill' but only 'less famous than Bill' (see Atlas 1984, who discusses this issue extensively). Given that *as famous as Bill* by logical implication means 'at least as famous as Bill', *not as famous as Bill* must mean 'less famous than Bill'. The interpretation 'more famous than Bill' cannot be allowed as it gives rise to a contradiction: if one is more famous than Bill, one is also as famous as Bill. Note that the same happens in *John has three pounds*. *John doesn't have three pounds* does not mean that he has either more or less than three pounds, but only that he has less. Again, the interpretation in which he has more than three pounds would lead to a contradiction. As expected, *John isn't six feet tall*, where the function $\mathcal{D}istance$ has applied, patterns alike.

In view of what we have just said, the example in (83) comes as a surprise. If the numeral in *three pounds* is contrastively focussed, the interpretation otherwise excluded becomes available. The reason for this lies in the well-known fact that contrastive focus implies exclusion of alternatives. It hence refers to the sum of exactly three pounds, with the result that (83) is non-contradictory.

(83) *John doesn't have* ***three*** *pounds – he has* ***four*** *pounds.*

In the same vein, the possibility of interpreting *as* as *equal to or more than* is lost under contrastive focus. Given the semantics in (82), its interpretation should then be restricted to the reference point itself. Hence, contrastive focus should make available an interpretation otherwise excluded, namely one in which *not as famous* is taken to mean 'more famous'. This prediction is borne out by (84a); (84b) shows that *not six feet tall* behaves in the same way.[19]

(84) a. *John isn't* ***as*** *famous as Bill – he is much* ***more*** *famous.*
b. *John isn't* ***six*** *feet tall – he is* ***seven*** *feet tall.*

For discussion of scalar implicatures in cases like the above see Carston (1995, 2003).

There is a further test by which we can determine whether a degree expression introduces the function *Up* or *At*. *At* reduces the adjective's scale to a singleton set, as it selects only those points that are identical to P_{ref}. This implies that further modification by means of a degree item involving the function *Distance* is excluded: it is impossible to measure from the bottom of a singleton set. *Up* is different in this respect, as it derives a non-singleton set consisting of all points above P_{ref}. Hence, it is possible to add a modifier which measures the distance between the bottom of the derived scale and the property selected. The contrast in (85) confirms the semantics of *as* versus that of *more*.[20]

(85) a. *John F. was a good deal more famous than Marilyn.*
b. *John F. was a good deal less famous than Marilyn.*
c. **John F. was a good deal as famous as Marilyn.*

---

19. The logic just discussed equally applies to reversed scales. For instance, if John is as small as Bill, then he may well be smaller. His value on the scale of smallness may be higher on the reversed scale than Bill's, which means closer to zero.
20. A TLR reviewer points out that the examples in (i) seem problematic given our claim that *as* selects a singleton set and should therefore not permit modification.

(i) a. *John is exactly as tall as Benny.*
b. *John is three times as tall as Benny.*

We do not think that these data are incompatible with our analysis. With respect to expressions like *exactly*, it should be noted that they clearly do not modify scales. In an example like *John has exactly six pounds* it modifies *six*, which is not scalar. In contrast, a measure phrase does modify scalar expressions and hence something like **John has a good deal six pounds* is ungrammatical. A very similar argument applies to *three times*, which does not modify a scale either, given the grammaticality of *John has three times two pounds*.

There are three further degree expressions that we classify as non-evaluative but which do not license a satellite: *this*, *that* and the question word *how*. The relevant data are given below:

(86) a. **John is this tall as Mary.*
b. **John is tall but he is not that tall as Mary.*
c. **How tall as Mary is John?*

In the case of *this* and *that*, the reason that no satellite is licensed presumably lies in the type of point on which the semantics of these expressions is anchored. The reference points introduced by *more*, *less* and *at* are pronominal in the sense that they can either take a grammatical antecedent (the satellite) or be given a value through context. Deictic and discourse-anaphoric expressions, however, do not have the capacity to take a grammatical antecedent and hence the anchor point of *this* and *that* is always identified contextually. In the formula below we refer to this anchor point as $P_{context}$:

(87) $[\![\text{this/that}(P_{context})]\!] = \lambda\mathcal{P}\lambda x\ \exists P\ [P \in \mathcal{A}t\,(P_{context}, \mathcal{P}) \ \&\ P(x)]$

Although these degree words do not license a satellite, they should be classified as non-evaluative because no average value enters into the identification of $P_{context}$: if one says *John is this tall*, there is no implication that he has an average length.

The question word *how* does not license a satellite for very different reasons: it does not contain $\mathcal{A}t$ or indeed any function, but instead asks the hearer to identify a predicate in the scale $\mathcal{P}$ by applying an appropriate function to that scale:

(88) a. *How tall is John?*
b. *Six feet tall. / Very tall. / As tall as Jack is. / Taller than Jack is.*

We may therefore represent the semantics of *how* as follows:

(89) $[\![\text{how}]\!] = \lambda\mathcal{P}\lambda x\ \text{W}{\scriptstyle\text{H}}(P)\ [P \in \mathcal{P}\ \&\ P(x)]$

3.3. Evaluative degree expressions

The three semantic classes of non-evaluative degree expression are complemented by three classes of evaluative degree expression: (i) evaluative measure phrases like *a little*; (ii) an expression involving comparison, namely *too*; and (iii) a semi-equative expression, namely *enough*.

Evaluative degree expressions differ from non-evaluative ones in that the semantic operations they encode are based on a point of comparison that is derived from a contextually determined variant of $\mathcal{P}$. In order to see what we mean by this, let us go back to the meaning of *tall* in the example *John is tall*. The definition in (71) does not sufficiently express the context dependence of the interpretation of *tall*. When computing $P_{average}$, we do not apply the function $\mathcal{Average}$ to $\mathcal{P}$, but rather to a contextually determined variant of it. For instance, if John is an adult man, we are usually not taking into account the length of babies and children when we want to know whether he is tall or not. If we did, practically every adult man would be tall. Therefore, *John is tall* can only be interpreted if we know what contextual factors have to be taken into account. In the literature this phenomenon is often handled in terms of a comparison class that contains those individuals that are contextually relevant for the interpretation of the adjective (cf. Klein 1980). This idea can be implemented in our model in the following way. On the basis of $\mathcal{P}$ we create $\mathcal{P}'$, which consists of the subset of $\mathcal{P}$ compatible with the comparison class.[21] We can now make the definition of *tall* in (71) more precise, by specifying that $P_{average}$ is in fact derived on the basis of $\mathcal{P}'$:

(90) $[\![\text{tall}(\mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{Distance}\ (\text{Considerable},\ \mathcal{Up}\ (\mathcal{Average}\ (\mathcal{P}'),\ [\![\text{tall}_A]\!]))\ \&\ P(x)]$

The preceding discussion implies that scale enrichment rule I should be reformulated as below:

(91) *Scale enrichment rule I*
$\text{SCALE}_{bare} \rightarrow \mathcal{Up}\ (\mathcal{Average}\ (\mathcal{P}'),\ \text{SCALE})$

There are satellites that make explicit the contextually derived subset $\mathcal{P}'$. Thus, in (92a) $\mathcal{P}'$ consists of lengths typical of boys, while in (92b) it consists of lengths that allow one to go on a submarine.

(92) a. *John is tall for a boy.*
b. *Alexander is tall to go on a submarine.*

As one would expect, there are also degree expressions that introduce a restriction which depends on a reference point which is derived from $\mathcal{P}'$. These expressions license the same type of satellites that bare adjectives do (see Bierwisch 1989 for related discussion):

(93) a. *John is too tall for a boy.*

---

21. Because it is orthogonal to the main line of argumentation, we do not address here the difficult problem of defining $\mathcal{P}'$. See von Stechow 1984a, Moltmann 1992 and Meier 1999 for discussion. In particular, much of what Meier 1999 proposes can be made compatible with the analysis defended here.

a′. *Alexander is too tall to go on a submarine.*
b. *John is competent enough for a promotion.*
b′. *Alexander is competent enough to be promoted.*
c. *John is very old for a professional tennis player.*
c′. *Alexander is very old to be a professional tennis player.*
d. *John is a little heavy for a jockey.*
d′. *Alexander is a little heavy to race a horse.*
e. *John was a good deal over the limit for a policeman.*
e′. *Alexander was a good deal over the limit to be driving a Ferrari.*

We argue below that these expressions all compute a reference point on the basis of $\mathcal{P}'$, but differ in the way this computation is achieved. Next to *Average*, the functions *Maximum* and *Minimum* can be used, depending on the lexical specifications of the degree expression. We will call expressions of this type evaluative, as the contextual selection of $\mathcal{P}'$ usually depends on an evaluation.

None of the non-evaluative degree items discussed earlier take satellites introduced by *for* or *to*:

(94) a. **John is more intelligent for a sailor.*
a′. **Alexander is more intelligent to join the navy.*
b. **John is older for a boy.*
b′. **Alexander is older to be in the third grade.*
c. **John is less competent for a promotion.*
c′. **Alexander is less competent to be promoted.*
d. **John is as muscular for a long-distance runner.*
d′. **Alexander is as muscular to be running the marathon.*

Conversely, none of the evaluative degree expressions take satellites introduced by *as* or *than*; neither do bare adjectives:

(95) a. **John is tall than/as Bill.*
b. **John is too tall than/as Fred.*
c. **John is competent enough than/as Peter.*
d. **John is very old than/as Max.*
e. **John is a little heavy than/as Adrian.*
f. **John was a good deal over the limit than/as Karl.*

In conclusion, there is strong evidence for a distinction between evaluative and non-evaluative degree expressions.

Let us now consider the semantics of evaluative degree expressions in more detail. The motivation for analyzing the degree expressions in (93) as involving a contextually determined subset $\mathcal{P}'$ is that they do not uniformly refer to $P_{average}$, but pick out different points in $\mathcal{P}'$. Consider first (93a′). This sentence does not mean that Alexander's height exceeds the average for submariners, but rather that there is a range of heights which qualify a person as a potential submariner and that Alexander's height exceeds all values in this interval. Hence, we can express the semantics of *too* by referring to the maximum value in the interval of contextually given admissible values. This is expressed by (96), which is identical to the formula that characterizes *more*, except for the point that *Up* takes as its first argument.

(96) $[\![\text{too}(\mathcal{P}')]\!] = \lambda\mathcal{P}\lambda x\ \exists P\ [P \in \mathit{Up}\ (\mathit{Maximum}\ (\mathcal{P}'), \mathcal{P}) \ \&\ P(x)]$

Once we accept that the values compatible with the contextually given comparison set are represented as an interval, the semantics of an example like (93b′) is straightforward. If Alexander is competent enough to be promoted to lieutenant, this does not mean that he has the average competence of a lieutenant but that his competence equals or exceeds the minimum required of lieutenants. This is expressed by (97), given the logical truth that, if a stronger property holds, weaker properties hold as well.

(97) $[\![\text{enough}(\mathcal{P}')]\!] = \lambda\mathcal{P}\lambda x\ \exists P\ [P \in \mathit{At}\ (\mathit{Minimum}\ (\mathcal{P}'), \mathcal{P}) \ \&\ P(x)]$

(97) is identical to the semantics of *as*, except for the point that *At* takes as its first argument.

The average value of $\mathcal{P}'$ (previously $P_{average}$) also plays a role in the semantics of evaluative degree items involving the function *Distance*. Such items restrict the domain of existential closure to those properties that lie at a particular distance from the bottom of the scale. However, if the bottom of the scale is given by an interval, measuring becomes problematic since different points in the interval give different outcomes. Therefore, if there are to be any evaluative degree items involving *Distance*, they must measure from a particular point in the interval, by hypothesis its average.

Consider the semantics of *a little tall* in *Alexander is a little tall to go on a submarine*. This sentence does not mean that Alexander exceeds the maximum length for submariners. Neither does it mean that his length is just above the minimum expected of submariners. Rather Alexander is somewhat taller than the average submariner. Thus, the semantics of *a little tall* can be formalized as below:

(98) $[\![\text{a little tall}(\mathcal{P}')]\!] = \lambda x\ \exists P\ [P \in \mathit{Distance}\ (\text{Small},\ \mathit{Up}\ (\mathit{Average}\ (\mathcal{P}'), [\![\text{tall}_A]\!])) \ \&\ P(x)]$

This formula expresses that the scale of length introduced by *tall* is divided by the average of the admissible lengths for submariners and that the lower part of the scale is discarded. Alexander's length is found by travelling a short distance upwards from the bottom of the resulting scale.

Let us consider how much of the semantics in (98) is contributed by *a little*. This question is pertinent given that the semantics in (98) partly overlaps with that introduced by the scale enrichment rule in (91). Considerations of parsimony suggest that we minimize the semantics of *a little* to (99a). As the formula requires a derived scale as input, on a par with scale enrichment rule II (see (72)), merger of *a little* is impossible if the bare scale introduced by *tall* does not undergo enrichment rule I, yielding (99b).

(99) a. $[\![\text{a little}]\!] = \lambda\mathcal{P}\lambda x\ \exists P\ [P \in \mathcal{D}istance\,(\text{Small}, \mathcal{P}_{\text{derived}})\ \&\ P(x)]$
b. $[\![\text{tall}(\mathcal{P}')]\!] = \mathcal{U}p\,(\mathcal{A}verage\,(\mathcal{P}'), [\![\text{tall}_A]\!])\ \&\ P(x)$

As was the case with *a little tall*, if Alexander is very tall to go on a submarine, he need not exceed the maximum length for submariners. This confirms that *very*, like *a little*, measures from the average over $\mathcal{P}'$, rather than from the interval's top value. Thus, the semantics of *very tall* is comparable with (98), except that the distance from the average of $\mathcal{P}'$ is large:

(100) $[\![\text{very tall}(\mathcal{P}')]\!] = \lambda x\ \exists P\ [P \in \mathcal{D}istance\ (\text{Large}, \mathcal{U}p\ (\mathcal{A}verage\ (\mathcal{P}')), [\![\text{tall}_A]\!]))\ \&\ P(x)]$

Assuming, as before, that part of this formula is due to the application of scale enrichment rule I, we may specify the semantics of *very* as below:

(101) $[\![\text{very}]\!] = \lambda\mathcal{P}\lambda x\ \exists P\ [P \in \mathcal{D}istance\,(\text{Large}, \mathcal{P}_{\text{derived}})\ \&\ P(x)]$

There is some independent evidence that *very* encodes $\mathcal{D}istance$, as it is excluded from a construction that in general does not tolerate elements containing this function. As Bowers (1975, 1987) has noticed, a curious restriction holds of the class of degree expression that can accompany APs fronted to the specifier of DP. (102) shows that a wide variety of degree expressions can modify a fronted AP (see also Bresnan 1973).

(102) a. *This is too big a car to park in this space.*
b. *This is as big a car as John has.*
c. *How big a car is this?*
d. *That big a car nobody can afford.*
e. *John was more intelligent an applicant than I expected.*
f. *John was intelligent enough a student to pass the exam.*

However, those degree expressions we have analyzed as introducing the function $\mathcal{D}istance$ cannot:

(103) a. **John was a little tall a man to join the navy.*
b. **John was 2 metres tall a man.*

Very patterns with the latter group and may therefore be classified as a measuring expression itself:

(104) *John was very tall a man to join the navy.

A final evaluative degree item involving the function $\mathcal{Distance}$ to be discussed here is *a good deal*. There is good reason to suppose that the semantics of *a good deal* lexically encodes the very same meaning that is introduced by scale enrichment rule II:

(105) $[\![\text{a good deal}]\!] = \lambda\mathcal{P}\ \lambda x\ \exists P\ [P \in \mathcal{Distance}\ (\text{Considerable}, \mathcal{P}_{\text{derived}})\ \&\ P(x)]$

As expected if the semantics of *a good deal* is that of a measure phrase, it cannot modify a fronted adjective:

(106) **John was a good deal different a boy.*

That *a good deal* measures the same distance as scale enrichment rule II is confirmed by the fact that it can only be attached to a restricted set of scalar expressions:

(107) a. *John turned out to be a good deal different from his brother.*
b. *Yet, the twins looked a good deal alike.*
c. ?*John seems a good deal dependent on his parents.*

(108) a. **That car is a good deal big.*
b. **It is a good deal dirty, too.*
c. **Not to mention a good deal old.*

As pointed out by Bresnan (1973), many, but not all, bare adjectives have a high-degree reading. Consider *different*. If one says that John is different from his brother, the implication is not that John is different from his brother to a high degree. Rather, there is some degree of difference that John must minimally have, but this may well be weaker than the average degree to which brothers are different from each other. So the lexical semantics of *different* is as in (109b), where $P_{\text{cut-off}}$ represents the point above which one can be called different (see Kennedy and McNally 1999 for related discussion). Existential closure then derives (109c).

(109) a. $[\![\text{different}]\!] = \text{SCALE}_{\text{difference}} = \{P_0, \ldots, P_i, \ldots, P_\infty\}$, where P_i is a property that applies to an individual x iff x is i-different.
b. $[\![\text{different}(P_{\text{cut-off}})_A]\!] = \mathcal{Up}\ (P_{\text{cut-off}}, [\![\text{different}]\!])]$
c. $[\![\text{different}(P_{\text{cut-off}})_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{Up}\ (P_{\text{cut-off}}, [\![\text{different}]\!])\ \&\ P(x)]$

Since *different*, as opposed to most adjectives, has a lexical restriction, it does not undergo scale enrichment. Consequently, it does not have a high degree reading when used in isolation. However, it can acquire such a reading when combined with a modifier like *a good deal*. If (105) is applied to (109b), we obtain the following formula:

(110) $[\![$a good deal different$(P_{\text{cut-off}})_{AP}]\!] = \lambda x\ \exists P\ [P \in$ *Distance* (Considerable, *Up* $(P_{\text{cut-off}}, [\![\text{different}]\!])$ & P(x)]

What (110) expresses is that a value is identified at a considerable distance above $P_{\text{cut-off}}$. As a result, *a good deal different* has the high-degree reading that *different* lacks. By contrast, **a good deal big* has the same semantics as *big* when used in isolation. In both cases, the adjective undergoes enrichment rule I. In the case of the bare adjective, this is followed by application of enrichment rule II, while in the case of **a good deal big* the same high degree reading is contributed by *a good deal*:[22]

(111) a. $[\![$a good deal big$(\mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ [P \in$ *Distance* (Considerable, *Up* (*Average* $(\mathcal{P}')$, $[\![\text{big}_A]\!]$)) & P(x)]

b. $[\![\text{big}(\mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ [P \in$ *Distance* (Considerable, *Up* (*Average* $(\mathcal{P}')$, $[\![\text{big}_A]\!]$)) & P(x)]

The ungrammaticality of **a good deal big* finds a natural explanation in the notion of representational economy (cf. Chomsky 1995). Suppose a well-formed extended projection of a head α blocks an alternative well-formed extended projection of α if it contains less material (see Grimshaw 1991 for a definition of extended projection and Grimshaw 1997 for its relevance in the definition of candidate set). Then, *big* blocks **a good deal big* because it acquires a high-degree reading without extension of its projection. But since *different* and *a good deal different* differ in interpretation, one cannot block the other. Hence, if *a good deal* is analyzed as in (105), its limited distribution follows.[23] (Note

22. Note that this argument supports the claim, implicit in the formulas proposed so far, that *very* has a stronger semantics than *a good deal* (as expressed by the first argument of the function *Distance*). Clearly, *very red* is not blocked by *red*.
23. Interestingly, *a good deal* cannot felicitously be attached to DPs (cf. i). This may be due to semantic reasons. In general DPs are not scalar expressions and if they are used in that way, their gradable interpretation is probably due to contextual coercion (see Jackendoff 1997 for discussion of this phenomenon). Recall that *a good deal* has the same semantics as that contributed by scale enrichment rule II and hence it cannot normally be attached to bare gradable expressions. Instead, its distribution, abstracting away from cases in which it modifies a degree head, is restricted to gradable expressions whose lexical semantics contributes a cut-off point, such as *different*. However, it is unlikely that under coercion DPs could acquire a cut-off point.

that representational economy does not affect the application of the scale enrichment rules, since their application gives rise to a new meaning, while representational economy compares candidates with identical semantics.)

The proposed semantics of measure phrases like *a little* is not inherently evaluative. In the cases we have discussed, the evaluative nature of the extended projection that hosts the measure phrase is due to the application of scale enrichment rule I. However, the semantics of measure phrases of this type only asks for a derived scale; it does not stipulate what type of operation must have applied. We therefore predict that a non-evaluative expression will result if the measure phrases in question modify a degree expression that does not compute a reference point on the basis of $\mathcal{P}'$. (In those circumstances, application of scale enrichment rule II will be blocked, because this rule only applies to bare scales.) This prediction is corroborated by the observation that satellites licensed by the presence of $\mathcal{P}'$ cannot occur in the examples below:

(112) a. *John is a little taller than Bill* (**for a jockey*).
b. *John is very much older than Peter* (**for a professional tennis player*).
c. *John was a good deal more over the limit than Neil* (**for a policeman*).

The division of labour between the semantics of the measure phrases under discussion and scale enrichment rule I receives support from the observation that expressions like *a little* can combine with comparatives without requiring computations involving $\mathcal{P}'$. If we say that John is a little taller than Fred, we do not mean that there is an average value by which people in the relevant comparison class are taller than Fred and that John's height exceeds this value by a little (as expressed by (113a)). Rather, what it means is that John's height exceeds Fred's height by a small amount (as (113b) expresses). This implies that *a little* must not encode the semantics supplied by scale enrichment rule I. In the same vein, *very much older* and *a good deal more over the limit* have the semantics in (113c) and (113d).

(113) a. $[\![\text{a little taller}(P_{ref}, \mathcal{P}')]\!] =$
$\lambda x\, \exists P\, [P \in \mathit{Distance}(\text{Small}, \mathit{Up}(\mathit{Average}(\mathcal{P}'), \mathit{Up}(P_{ref}, [\![\text{tall}_A]\!])))\ \&\ P(x)]$ (incorrect semantics)
b. $[\![\text{a little taller}(P_{ref})]\!] =$
$\lambda x\, \exists P\, [P \in \mathit{Distance}(\text{Small}, \mathit{Up}(P_{ref}, [\![\text{tall}_A]\!]))\ \&\ P(x)]$
c. $[\![\text{very much older}(P_{ref})]\!] =$
$\lambda x\, \exists P\, [P \in \mathit{Distance}(\text{Large}, \mathit{Up}(P_{ref}, [\![\text{old}_A]\!]))\ \&\ P(x)]$

(i) *?A good deal the Hollywood celebrity he always wanted to be, Robin ordered the marble 10-person jacuzzi.*

d. ⟦a good deal more over the limit(P_{ref})⟧ =
λx ∃P [P ∈ *Distance* (Considerable, *Up* (P_{ref}, ⟦over the limit⟧)) & P(x)]

Let us now return to the main line of argumentation. Degree expressions fall in two syntactic categories (functional heads and modifiers) and a morphological one. We have also seen that there is a basic semantic bifurcation, between evaluative degree expressions (whose semantics is based on an interval of admissible values) and non-evaluative ones (whose semantics is based on a reference point). Evidence for this division comes from various interpretative facts, as well as from the type of satellites licensed by the two classes of degree expression. Despite the fact that they employ a different object of comparison, evaluative and non-evaluative degree expressions are similar in that their semantics can be described in terms of the functions *Up*, *At* and *Distance*.

Once we bring these results together, it becomes clear that the two syntactic classes do not map uniformly onto the semantic classes of evaluative and non-evaluative degree expression (see (114)). Similarly, the semantic classes created by *Up*, *At* and *Distance* do not exclusively contain either heads or modifiers. Hence we may draw the conclusion that selection by degree heads must be syntactic rather than semantic in nature. Thus, the data considered so far provide an argument for a certain arbitrariness in the mapping between syntactic and semantic categories.

(114)

	Evaluative		Non-evaluative	
Up	*too*	class 1	*more, less, -er*	class 2 morphology
At	*enough*	class 2	*as*	class 1
Distance	*very*	class 1	*very*	class 1
	a little	class 2	*a little*	class 2
	straal	morphology	*3 meters*	class 2

Note: *very* and *a little* are in principle non-evaluative, but may trigger application of scale enrichment rule I, which inserts evaluative semantics.

(114) shows that there is not only a bifurcation between two syntactic classes of degree expression, but also between those two syntactic classes on the one hand and a morphological class on the other. Irrespective of one's view of morphology, degree expressions that form a word with their host must be formally distinguished from degree expressions that do not. But since morphologically dependent degree expressions do not map to a single semantic class either, our

overall conclusion that formal properties are independent of semantic properties is strengthened still further.[24]

3.4. The semantics of much

The simplest argument against the view that the c-selectional properties of degree expressions reduce to their semantics can be based on examples involving *much*. We argue in this section that the lexical semantics of *much* does not affect the semantics of *too much under the weather*. This PP is interpreted on a par with the DegP *too unwell*, which suggests that the modifier *too much* and the functional head *too* have the same semantics.

Corver (1997a) assumes two variants of *much*; one a dummy, the other semantically charged. The initial motivation for drawing such a distinction comes from the contrast in their distribution. Most non-adjectival phrases do not tolerate the semantically charged variant of *much*, but they can be combined with dummy *much* in the presence of a degree head:

---

24. Note that the functions contained in the restrictions of degree expressions take an ordered set as their second argument but do not necessarily rely on the notion of degree (in the sense used throughout this paper). Therefore, a natural extension suggests itself: the degree items could as well be applied to other types of ordered sets. Of course, this possibility is only available for class-2 expressions, as only these may be attached to other categories than AP. As has been extensively argued in the semantic literature (see Link 1983, Bach 1986 and Krifka 1986 amongst others), both NPs and VPs may introduce a set of individuals (singular and plural individuals or portions of matter) partially ordered with respect to each other by the part-of relation. These partially ordered sets are closely related to the ordered sets provided by the APs, even though their structure is slightly different: a series of individuals or portions with the same cardinality/amount do not form a separate set, whereas, for instance, the individuals which are i-tall are united in a set in the representation of the adjective *tall*. However, as the information necessary to form the new structure is present, it does not come as a surprise that class-2 expressions are able to specify cardinality or amount when attached to NPs and VPs of the type just described (cf. Doetjes 1997 for further discussion):

(i) a. *More students attended the lecture than expected.*
b. *There was less coffee in the can than I had hoped for.*
c. *John goes to the cinema more than any of his friends.*
d. *Mary studied enough to pass for the exam.*

The advantage of this approach is that these data can be captured without assuming that class-2 degree expressions are systematically multiply ambiguous.

(115) a. **John is much under the weather.*
b. *John is too much under the weather to come to work.*

APs, on the other hand, do not tolerate dummy *much*, as we have already shown in the previous section. Nevertheless, some can be combined with its semantically charged counterpart:

(116) *John is much different from his brother.*

Although the above data indicate that we must make some sort of distinction between the two variants of *much*, a unitary account does not seem beyond reach. The only difference between the two variants is that 'dummy' *much* lacks the high degree reading that characterizes 'semantically charged' *much*. However, since there is a rule, namely scale enrichment rule II, that could relate the two, the contrast may be reducible to that between adjectives that have and have not undergone this rule. In the same way that bare adjectives acquire a high degree reading, bare *much* may do so.

Let us assume, more concretely, that *much*'s lexical entry contains the semantic formula in (117a). If *much* is not combined with a degree expression, it will undergo scale enrichment rule II in (72), followed by existential closure. This results in the formula in (117b).[25]

(117) a. $[\![\text{much}_A]\!] = \lambda\mathcal{P}\,\mathcal{P}$
b. $[\![\text{much}_{AP}]\!] = \lambda\mathcal{P}\,\lambda x\ \exists P\ [P \in \mathcal{Distance}\ (\text{Considerable}, \mathcal{P}_{\text{derived}})\ \&\ P(x)]$

(117b) is what Corver refers to as 'semantically charged' *much*. If this variant is applied to *different*, in (118a), we get the formula in (118b) and thus the high degree reading that *different* lacks.

(118) a. $[\![\text{different}(P_{\text{cut-off}})_A]\!] = \mathcal{Up}\ (P_{\text{cut-off}}, [\![\text{different}]\!])$
b. $[\![\text{much different}(P_{\text{cut-off}})_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{Distance}\,(\text{Considerable}, \mathcal{Up}\ (P_{\text{cut-off}}, [\![\text{different}]\!]))\ \&\ P(x)]$

Much behaves like other scalar expressions in that scale enrichment rule II cannot apply to it if it combines with a degree item (since scale enrichment rules are rules of last resort). Consequently, *too much* is not associated with a high degree reading. Its semantics is the result of composing the formula in (117a)

25. The semantics of *little* is identical to that of *much*, except that it also contains the functor $\mathcal{Reverse}$. Hence, scale enrichment rule II gives rise to a low-degree reading in *This is little different from what is normally the case*. This low-degree reading is not necessarily implied by *John is too little into syntax to go to MIT*, as in this context scale enrichment rule II is blocked by the presence of a degree expression.

with that in (96). The resulting expression has exactly the same semantics as *too* itself:

(119) ⟦too much($\mathcal{P}'$)⟧= $\lambda\mathcal{P}\,\lambda x\ \exists P\ [P \in \mathcal{U}p\,(\mathcal{M}aximum\,(\mathcal{P}'), \mathcal{P}\,)\ \&\ P(x)]$

In other words, if *much* fails to undergo scale enrichment rule II, it does not make a semantic contribution. We can now understand why *too much under the weather* and *too unwell* have identical semantics (on the assumption that *under the weather* and *unwell* do). Combining *too* with the adjective gives (120b), while combining *too much* with the PP results in (120c).

(120) a. ⟦unwell⟧ = ⟦under the weather⟧ = $\text{SCALE}_{\text{illness}}$
b. ⟦too unwell($\mathcal{P}'$)$_{AP}$⟧ = $\lambda x\ \exists P\ [P \in \mathcal{U}p\,(\mathcal{M}aximum\,(\mathcal{P}'), \text{SCALE}_{\text{illness}})\ \&\ P(x)]$
c. ⟦too much under the weather($\mathcal{P}'$)$_{PP}$⟧ = $\lambda x\ \exists P\ [P \in \mathcal{U}p\,(\mathcal{M}aximum\,(\mathcal{P}'), \text{SCALE}_{\text{illness}})\ \&\ P(x)]$

Since *much* does not always contribute to the interpretation of the expression in which it is contained, its distribution is restricted by economy. We have assumed that a well-formed extended projection of a head α blocks an alternative well-formed extended projection of α if it contains less material. Consequently, *much* insertion is restricted to situations in which there are no competitors. Potential competitors either give rise to a different semantics or are ruled out on syntactic grounds. In the case of *too much under the weather*, for instance, the potential competitor **too under the weather* is ungrammatical, so *much* insertion is allowed. There are three further cases to consider.

First of all, *much* cannot be combined with gradable expressions that can also acquire a high-degree reading through scale enrichment rule II. As we have seen, there are some adjectives whose lexical semantics contains a restriction. These do not undergo the rule in (91) and may consequently be modified by *much* (see (116)). However, adjectives lacking a lexical restriction, such as *big*, acquire a high-degree reading through scale enrichment rules I and II (see (121b)). In this case, *much* cannot be inserted, as this would give rise to the same semantics. As before, it is input to scale enrichment rule II, but since this rule itself requires a derived scale, application of enrichment rule I is triggered as well. The output of these operations is given in (121b′).

(121) a. The car is (*much) big.
b. ⟦big($\mathcal{P}'$)$_{AP}$⟧ = $\lambda x\ \exists P\ [P \in \mathcal{D}istance\,(\text{Considerable}, \mathcal{U}p\,(\mathcal{A}verage\,(\mathcal{P}'), [\![\text{big}_A]\!]))\ \&\ P(x)]$
b′. ⟦much big($\mathcal{P}'$)$_{AP}$⟧ = $\lambda x\ \exists P\ [P \in \mathcal{D}istance\,(\text{Considerable}, \mathcal{U}p\,(\mathcal{A}verage\,(\mathcal{P}'), [\![\text{big}_A]\!]))\ \&\ P(x)]$

Note that this situation closely resembles the one found with *a good deal* (cf. (107) and (108)). In both **a good deal big* and **much big* an extended projection contains more material than necessary, since neither *a good deal* nor *much* is required to obtain the targeted semantics.

Second, *much* cannot be combined with a class-1 expression when the modifier so formed attaches to an AP. The reason is that class-1 degree expressions can directly attach to APs themselves, yielding identical semantics. Hence, insertion of *much* violates economy:

(122) a. The car is too (*much) big.
b. $[\![\text{too big}(\mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'), [\![\text{big}_A]\!])\ \&\ P(x)]$
b′. $[\![\text{too much big}(\mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'), [\![\text{big}_A]\!])\ \&\ P(x)]$

The third and final case consists of *much* support with a class-2 degree item. Since class-2 expressions can attach to phrases of any category, insertion of *much* again violates economy:

(123) a. *John is more* (**much*) *into syntax.*
b. $[\![\text{more into syntax}(P_{ref})_{PP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{Up}\,(P_{ref}, [\![\text{into syntax}]\!])\ \&\ P(x)]$
b′. $[\![\text{more much into syntax}(P_{ref})_{PP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{Up}\,(P_{ref}, [\![\text{into syntax}]\!])\ \&\ P(x)]$

In summary, there is only one lexical entry for *much*, with a minimal semantic specification. Depending on context, *much* may or may not undergo scale enrichment rule II, something which is true of the vast majority of adjectives. Enrichment creates semantically charged *much*; lack of enrichment results in its dummy variant. The distribution of *much* is governed by economy: it may only be inserted if there is no grammatical competitor. Hence, semantically charged *much* cannot combine with high-degree adjectives, while dummy *much* cannot combine with any adjective at all.

We may conclude that dummy *much* may support any class-1 item that attaches to a non-adjectival expression. The implication is that for every class-1 degree expression there is a class-2 alternate derived by attaching it to *much* (see (124)). This fact alone suffices to eliminate any hope of reducing the c-selectional properties of degree expressions to their semantics.

(124)

	Evaluative		Non-evaluative	
Up	*too*	class 1	*more, less*	class 2
	too much	class 2	*-er*	morphology
At	*enough*	class 2	*as*	class 1
			as much	class 2
Distance	*very*	class 1	*very*	class 1
	very much	class 2	*very much*	class 2
	a little	class 2	*a little*	class 2
	straal	morphology	*3 meters*	class 2

Note: *very* and *a little* are in principle non-evaluative, but may trigger application of scale enrichment rule I, which inserts evaluative semantics.

The at least partly arbitrary relation between the syntactic status of a degree expression and its semantics implies that we may expect a great deal of variation in the realization of degree across languages. Of course, languages that are historically related are likely to map syntax and semantics in a similar way, but we expect the three formal classes to be distributed differently across the various semantic functions in unrelated languages. It would take us too far afield to pursue this matter here.

4. Interface issues: Recursion of degree expressions

4.1. Recursion as a result of specifier-head agreement

Traditional accounts of the syntax of degree expressions, such as those of Bresnan (1973) and Jackendoff (1977), assume that there is a unique position for which all degree expressions compete. This predicts that they are in complementary distribution. Indeed the examples in (125) are ungrammatical.

(125) a. **John F. was very more famous.*
b. **John F. was more very famous.*

The facts are more complicated than these examples suggest. (126) shows that certain combinations of degree expressions *are* grammatical:

(126) a. *John F. was very much more famous than Marilyn.*
b. *John is too big for this sweater, and Bill is too big for it as well. But it seems to me that John is less too big for it than Bill.*

Since class-1 and class-2 expressions do not compete for the same syntactic position in our analysis, we cannot attribute the contrast between (125) and (126) to the syntax alone. On the other hand, it cannot be attributed to just the semantics either. We have argued that degree expressions uniformly introduce an existential quantifier. Although this could potentially explain the ungrammaticality of (125) – combining an adjective with more than one degree expression would give rise to vacuous quantification – (126) shows that this line of reasoning is incorrect.

We argue below that the above patterns of complementarity and recursion follow from the interaction of semantic and syntactic factors. The ungrammatical examples indeed involve vacuous quantification, but at the syntax-semantics interface certain structures allow a reduction in the number of existential quantifiers, as a result of which the grammatical examples are salvaged. More specifically, this reduction is conditioned by specifier-head agreement.

The essence of specifier-head agreement is the sharing of a feature between an X^0 and a maximal projection, as a result of which the shared feature is mapped onto a single semantic entity: the feature on the head is 'deleted' under identity with that in the specifier. This is the intuition behind analyses of verbal concordance in terms of specifier-head agreement, but something along these lines is also necessary in the interpretation of embedded wh questions, where a wh operator moves to the specifier of a subordinating head. In languages that allow phonological realization of the relevant head, such as some variants of Dutch, this head typically has wh properties (if used on its own, it introduces a yes/no question; see (127b)). For most speakers, (127c) is ungrammatical.

(127) a.

Ik	*vraag*	*me*	*af*	$[_{CP}$	*wie*	$[_{C'}$	*of*	[	*Peter*	*op*	*z'n*
I	wonder	me	PRT		who				Peter	on	his
feestje	*uitgenodigd*	*heeft*]]]									
party	invited	has									

b.

Ik	*vraag*	*me*	*af*	*of*	*Peter*	*Marie*	*uitgenodigd*	*heeft.*
I	wonder	me	PRT	if	Peter	Mary	invited	has

c.

**Ik*	*vraag*	*me*	*af*	$[_{CP}$	*wie*	$[_{C'}$	*dat*	[	*Peter*	*op*	*z'n*
I	wonder	me	PRT		who		that		Peter	on	his
feestje	*uitgenodigd*	*heeft*]]]									
party	invited	has									

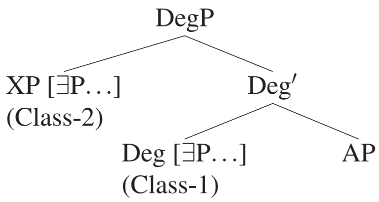
From a lexical perspective, both *wie* and *of* contain a wh feature. From a semantic perspective, however, there is only one wh feature, as (127a) is a simple wh question. Hence, one of the two wh features is ignored by the semantics (the feature in the head is 'deleted' under identity).

If LF treats specifier-head configurations in this way, combining degree expressions in such a configuration need not result in vacuous quantification, as long as the existential quantifier of the head deletes under identity with that of its specifier. This situation obtains in two configurations.

4.2. Specifiers of class-1 degree items

First, a class-1 item can take a class-2 expression as its specifier in a structure like (128). The existential quantifier in the head is deleted under identity with that in the specifier, due to the way in which specifier-head relationships are treated at LF.

(128)



As we saw in (126b), a class-2 degree expression can be attached externally to a class-1 head. The relevant part of the example is repeated here for convenience:

(129) *John is* [$_{\text{DegP}}$ *less* [$_{\text{Deg}'}$ *too* [$_{\text{AP}}$ *big*]]] *for it than Bill.*

The semantics of (129) is derived as follows. The specifier-head relation between *less* and *too* allows the latter to be reduced from (130b) to (130b′), which applied to *big* yields (130c). (130a) applied to (130c) results in (130d).

(130)
a. $[\![\text{less}(P_{ref})]\!] = \lambda\mathcal{P}\,\lambda x\ \exists P\ [P \in \mathcal{Up}\,(P_{ref}, \mathcal{Reverse}\,(\mathcal{P}))\ \&\ P(x)]$
b. $[\![\text{too}(\mathcal{P}')]\!] = \lambda\mathcal{P}\,\lambda x\ \exists P\ [P \in \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'), \mathcal{P})\ \&\ P(x)]$
b′. $[\![\text{too}(\mathcal{P}')]\!] = \lambda\mathcal{P}\,[\mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'), \mathcal{P})]$
c. $[\![\text{too big}(\mathcal{P}')_{AP}]\!] = \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'), [\![\text{big}_A]\!])]$
d. $[\![\text{less too big}(P_{ref}, \mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ [P \in \text{Up}(P_{ref}, \mathcal{Reverse}\,(\mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'), [\![\text{big}_A]\!])))\ \&\ P(x)]$

The semantics of *less too big* can be described as follows. *Big* provides an ordered set, which is split on the basis of the $\mathcal{P}'$ introduced by *too*; hence, the option of adding *for this sweater* as a satellite to *too big*. The resulting scale undergoes reversal, after which a reference point and the function $\mathcal{Up}$ are used to further reduce the domain of existential quantification. The reference point licenses the satellite *than Bill*. The existential quantifier therefore selects a point that lies between two contextually determined points, namely the degree of bigness above which one is too big for the sweater in question and Bill's degree of bigness.

Note that the interpretation of *less too big* shows that the semantics of degree expressions must be based on the manipulation of ordered sets, as proposed here and in much earlier work (see for instance Kennedy 2000a and references cited there). More specifically, it cannot be based on the direct comparison of points taken from a scale, as proposed by Cresswell (1976), amongst others. In the latter approach the semantics of *John is less too big for this sweater than Bill* incorrectly allows an interpretation in which John is in fact not too big for the sweater in question. It would be asserted of Bill's size, $size_{Bill}$, that it exceeds $size_{sweater}$, the point at which the sweater no longer fits ($size_{Bill} > size_{sweater}$). Of John's size, $size_{John}$, it would be asserted that $size_{John} < size_{Bill}$. These assertions are compatible with a situation in which $size_{John} < size_{sweater}$, so that he would in fact fit the sweater.

Of course, other degree expressions involving $\mathcal{Up}$ can appear in the specifier of *too* as well. For example, given an appropriate context, *more too big* is grammatical (131a). Its semantics, in (131b), is identical to that of *less too big*, except that the function $\mathcal{Reverse}$ is absent.

(131) a. *John is too big for this sweater, and Bill is too big for it as well. But it seems to me that John is* [$_{DegP}$ *more* [$_{Deg'}$ *too* [$_{AP}$ *big for it*]]] *than Bill.*
b. $[\![\text{more too big}(P_{ref}, \mathcal{P}')_{AP}]\!] = \lambda x\, \exists P\, [P \in \mathcal{Up}\, (P_{ref}, \mathcal{Up}\, (\mathcal{Maximum}\, (\mathcal{P}'), [\![\text{big}_A]\!])) \,\&\, P(x)]$

The specifiers discussed so far all involve class-2 degree expressions. However, nothing excludes merger of a class-1 degree item as a specifier, although this triggers *much* support. If a degree head appears in a specifier position on its own, its c-selectional requirements are not met, as selection is restricted to the head-complement relation. Hence, merger of a degree head requires prior merger with dummy *much*. Limiting ourselves for the time being to degree expressions involving the function $\mathcal{Up}$, the following example bears out our expectations:[26]

(132) *Superman is usually too clumsy to save a cat in trouble; yet its owner almost always forgives him. But today he was* [$_{DegP}$ *too much* [$_{Deg'}$ *too clumsy*]] *to be forgiven*.

As predicted, **too too clumsy* is ungrammatical and remains so no matter the context.[27] The semantics of *too much too clumsy*, in (133c), is derived as

26. A TLR reviewer wonders why in *too much too clumsy* the restriction of the innermost degree head is not deleted under identity with the restriction of *too much*. The answer to this question lies in the notion of representational economy. Suppose that the restriction were deleted. Then *too much too clumsy* would be synonymous with *too clumsy*. But since the latter contains less material, it blocks the former. Hence, *too much too clumsy* will only be grammatical if deletion is restricted to the existential operator.

27. A TLR reviewer points out that *too too A* and *very very A* are acceptable (at least for some speakers). Note that the semantics of such expressions is not compositional in that it involves intensification rather than application of semantics of the second degree expression to the

follows. *Too much* has the semantics in (133a) and *too* that in (133b). The latter formula is reduced to (133b') under specifier-head agreement, after which it is applied to *clumsy*, resulting in (133c). Finally, (133a) applies to (133c) yielding the semantics of *too much too clumsy* in (133d). In this formula the contextually determined interval associated with *too much* concerns "too clumsy-ness" rather than clumsiness. The interval associated with clumsiness relates to the innermost *too*.

(133) a. $[\![\text{too much}(\mathcal{P}'_1)]\!] = \lambda\mathcal{P}\lambda x\ \exists P\ [P \in \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'_1), \mathcal{P}) \ \&\ P(x)]$
b. $[\![\text{too}(\mathcal{P}'_2)]\!] = \lambda\mathcal{P}\lambda x\ \exists P\ [P \in \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'_2), \mathcal{P}) \ \&\ P(x)]$
b′. $[\![\text{too}(\mathcal{P}'_2)]\!] = \lambda\mathcal{P}\ \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'_2), \mathcal{P})$
c. $[\![\text{too clumsy}(\mathcal{P}'_2)]\!] = \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'_2), [\![\text{clumsy}_A]\!])$
d. $[\![\text{too much too clumsy}(\mathcal{P}'_1, \mathcal{P}'_2)_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'_1), \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'_2), [\![\text{clumsy}_A]\!])) \ \&\ P(x)]$

The semantics in (133d) should license two satellites, as it mentions two contextually determined intervals. Although it seems that multiple satellites of the same type often give rise to reduced acceptability, (134a) seems grammatical. In this example, *to save the cat* is licensed by $\mathcal{P}'_2$ and *to be forgiven* by $\mathcal{P}'_1$. That it is indeed the outermost *too* which licenses the outermost satellite is corroborated by the ungrammaticality of (134b) (where contrastive focus is used to force a reading of the first infinitival as the satellite of the innermost degree head).

(134) a. *Superman was too clumsy to save the rabbit; yet he was forgiven by its owner. ?But he was* $[_{\text{DegP}}$ *too much* $[_{\text{Deg}'}$ *too clumsy*$]]$ *to save the **cat** to be forgiven.*
b. *Superman was too clumsy to save the rabbit; yet he was forgiven by its owner. *But he was* $[_{\text{DegP}}$ *too clumsy*$]$ *to save the **cat** to be forgiven.*

Not only degree expressions involving the function $\mathcal{Up}$ but also those involving $\mathcal{At}$ can occupy the specifier of the degree head *too*, as shown by the following example:

(135) *Superman is usually too clumsy to save a cat in trouble; yet his reputation remains unblemished. But today he was* $[_{\text{DegP}}$ *enough* $[_{\text{Deg}'}$ *too clumsy*$]]$ *for his reputation to be destroyed.*

---

output of the first. Thus, *too too A* means something like 'much too A', not 'too much too A'. We abstract away from these cases, assuming that they do not involve recursion of degree heads.

The semantics of *enough too clumsy* is given in (136). As before, specifier-head agreement is responsible for the elimination of the existential operator in *too*.

(136) $[\![$enough too clumsy$(\mathcal{P}'_1, \mathcal{P}'_2)_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathcal{At}\ (\mathcal{Minimum}\ (\mathcal{P}'_1),\ \mathcal{Up}\ (\mathcal{Maximum}\ (\mathcal{P}'_2), [\![clumsy_A]\!])) \ \&\ P(x)]$

The two contextually determined intervals in (136) should license two satellites. Indeed, in (137a) *to save the cat* is licensed by $\mathcal{P}'_2$ and *for his reputation to be destroyed* by $\mathcal{P}'_1$. That *enough* indeed licenses the outermost satellite is corroborated by the ungrammaticality of (137b).

(137) a. *Superman is usually too clumsy to save an animal in trouble; yet his reputation remains unblemished. ?But today he was* [$_{DegP}$ *enough* [$_{Deg'}$ *too clumsy*]] *to save his neighbour's cat for his reputation to be destroyed.*

b. *Superman is usually too clumsy to save an animal in trouble; yet his reputation remains unblemished. *But today he was* [$_{DegP}$ *too clumsy*] *to save his neighbour's cat for his reputation to be destroyed.*

As we have seen, *too* also tolerates a class-1 item in its specifier, provided *much* support takes place:

(138) a. *The incredible Hulk was* [$_{DegP}$ *as much* [$_{Deg'}$ *too clumsy*]] *to save the cat as Superman.*

b. **The incredible Hulk was* [$_{DegP}$ *as* [$_{Deg'}$ *too clumsy*]] *to save the cat as Superman.*

The derivation is similar to that outlined in (133) and again relies on the elimination of the existential operator of *too* under specifier-head agreement: (139a) combines with (133c) to yield (139b).

(139) a. $[\![$as much$(P_{ref})]\!] = \lambda\mathcal{P}\,\lambda x\ \exists P\ [P \in \mathcal{At}\,(P_{ref}, \mathcal{P})\ \&\ P(x)]$

b. $[\![$as much too clumsy$(P_{ref}, \mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ [p \in \mathcal{At}\,(P_{ref}, \mathcal{Up}\,(\mathcal{Maximum}\,(\mathcal{P}'), [\![clumsy_A]\!]))\ \&\ P(x)]$

The contextually determined interval in (139b) licenses the satellite *to save the cat* in (138b), whereas P_{ref} licenses *as Superman.* That this is so is confirmed by the ungrammaticality of (140).

(140) **The incredible Hulk was* [$_{DegP}$ *too clumsy*]] *to save the cat as Superman.*

The degree head *too* also allows further modification by degree items involving $\mathcal{Distance}$:

(141) a. *The projected flight path is* [$_{DegP}$ *several miles* [$_{Deg'}$ *too* [$_{AP}$ *high*]]].
b. *The projected flight path is* [$_{DegP}$ *a little* [$_{Deg'}$ *too* [$_{AP}$ *high*]]].

The derivation again relies on elimination of an existential operator in a specifier-head configuration. Function application then results in a semantics in which *Distance* applies to the output of *Up*. In other words, *several miles* selects a property whose strength is increased by several miles in comparison to the strength of the property at the bottom of the scale (which was derived by *too*; see (142)). The semantics of (141b) differs from (142) only in the first argument of *Distance*.

(142) $[\![\text{several miles too high}(\mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ \exists n\ [n \geq 1\ \&\ P \in \mathit{Distance}\,(n \text{ miles}, (\mathit{Up}\,(\mathit{Maximum}\,(\mathcal{P}'), [\![\text{high}_A]\!]))\ \&\ P(x)]$

The measure phrases in (141) are maximal projections. However, as expected, the degree head *very*, which also employs the function *Distance*, may appear in the specifier of *too*, provided *much* support takes place. (The resulting semantics is similar to (142), except for the first argument of *Distance*.)

(143) a. *John is* [$_{DegP}$ *very much* [$_{Deg'}$ *too big*]] *for this sweater.*
b. **John is* [$_{DegP}$ *very* [$_{Deg'}$ *too big*]] *for this sweater.*

Perhaps surprisingly, *much* can also appear in the specifier *too* on its own:

(144) *Superman was, in my opinion,* [$_{DegP}$ *much* [$_{Deg'}$ *too clumsy*]] *to save the cat.*

It is clear that this is not a case of *much* support, as *much* is not the complement of a degree head. Therefore, *much* will undergo scale enrichment rule II and existential closure. It will in other words acquire the semantics in (145a). The specifier-head configuration results again in the semantics (133c) for *too clumsy*; when (145a) applies to it, we obtain (145b).

(145) a. $[\![\text{much}_{AP}]\!] = \lambda\mathcal{P}\ \lambda x\ \exists P\ [P \in \mathit{Distance}\,(\text{Considerable}, \mathcal{P}_{\text{derived}})\ \&\ P(x)]$
b. $[\![\text{much too clumsy}(\mathcal{P}')_{AP}]\!] = \lambda x\ \exists P\ [P \in \mathit{Distance}\,(\text{Considerable}, \mathit{Up}\,(\mathit{Maximum}\,(\mathcal{P}'), [\![\text{clumsy}_A]\!]))\ \&\ P(x)]$

This formula is associated with a single contextually determined interval. As a consequence, *much too clumsy* can only be associated with a single satellite. Thus, semantically charged *much* is predicted to be like *a little*, *very* and *a good deal* in that it does not license a satellite (compare (112)). A TLR reviewer points out the ungrammaticality of the example in (146), which confirms this prediction.

(146) *Superman was too clumsy to save the rabbit; yet he was forgiven by its owner. *But he was much too clumsy to save the cat to be forgiven.*

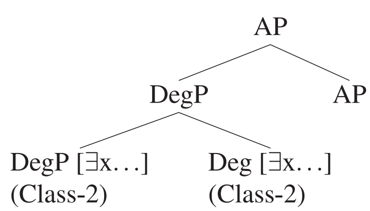
Although specifier-head agreement makes it possible to combine two degree expressions, not every such structure will be well-formed, since the restrictions of the degree expressions may not be compatible. Degree items involving the functions *At* or *Distance* reduce the scale provided by the adjective to a singleton set. They thereby block further attachment of degree modifiers, as all such modifiers take an ordered set of properties as their input. Thus, the degree head *as* (involving *At*) resists insertion of degree expressions in its specifier (see (147a)). Similarly, the degree head *very* (involving *Distance*) does not accept the modifiers that *are* accepted by *too* (see (147b)).

(147) a. **John was* [$_{\mathrm{DegP}}$ *less/more/too much/a little/enough/as much/very much/much* [$_{\mathrm{Deg'}}$ *as clumsy]] as Bill.*
b. **John was* [$_{\mathrm{DegP}}$ *less/more/too much/a little/enough/as much/very much/much* [$_{\mathrm{Deg'}}$ *very clumsy*]] *for this type of work.*

4.3. *Specifiers of class-2 degree items*

There is a second configuration that allows the combination of two degree expressions, namely one in which a class-2 modifier occupies the specifier position of another class-2 modifier, as in the left-branching structure below (for concreteness we represent class-2 degree items as intransitive Degs):

(148)



Structures of this type are easily found, although they may require some pragmatic priming. We discuss some examples below involving the modifier *more* (the data can be replicated for *less*).

First of all, *more* may take another class-2 item introducing *Up* in its specifier:

(149) *The blue and the yellow sweaters are both more expensive than the red one. But the blue sweater is* [$_{\mathrm{AP}}$ [$_{\mathrm{DegP}}$ *less* [$_{\mathrm{Deg'}}$ *more*]] [$_{\mathrm{A'}}$ *expensive*]].

For reasons of space, we will not discuss the semantics of (149) and subsequent examples. However, the reader may verify that the appropriate interpretation for each example can be derived once we allow omission of an existential quantifier under specifier-head agreement.

More may also combine with a class-1 degree item involving the function $\mathcal{Up}$, provided that *much* support takes place. Indeed, *too much more* is an allowed combination:

(150) a. *Susan is more intelligent than her supervisor; yet they have a good working relation. Mary, however, is* $[_{\text{AP}}\ [_{\text{DegP}}$ *too much* $[_{\text{Deg}'}$ *more*$]]\ [_{\text{A}'}$ *intelligent*$]]$ *than her supervisor for them to get along.*

b. *Susan is more intelligent than her supervisor; yet they have a good working relation. *Mary, however, is* $[_{\text{AP}}\ [_{\text{DegP}}$ *too* $[_{\text{Deg}'}$ *more*$]]\ [_{\text{A}'}$ *intelligent*$]]$ *than her supervisor for them to get along.*

A second group of modifiers that can be combined with *more* introduce the function $\mathcal{At}$. For instance, it is possible to use *enough more*, given an appropriate context.

(151) *Susan is more intelligent than her supervisor. Indeed she is* $[_{\text{AP}}\ [_{\text{DegP}}$ *enough* $[_{\text{Deg}'}$ *more*$]]\ [_{\text{A}'}$ *intelligent*$]]$ *to irritate him thoroughly.*

Again it is also possible for a class-1 degree item introducing $\mathcal{At}$ to appear in the specifier of *more*, but only if *much* support takes place:

(152) a. *Susan and Mary are more intelligent than their respective supervisors. But only Susan is* $[_{\text{AP}}\ [_{\text{DegP}}$ *as much* $[_{\text{Deg}'}$ *more*$]]\ [_{\text{A}'}$ *intelligent*$]]$ *than her supervisor as she claims.*

b. *Susan and Mary are more intelligent than their respective supervisors. *But only Susan is* $[_{\text{AP}}\ [_{\text{DegP}}$ *as* $[_{\text{Deg}'}$ *more*$]]\ [_{\text{A}'}$ *intelligent*$]]$ *than her supervisor as she claims.*

Measure phrases form the third group of specifiers *more* can take:

(153) *Susan is more intelligent than her supervisor. Indeed she is* $[_{\text{AP}}\ [_{\text{DegP}}$ *30 IQ points* $[_{\text{Deg}'}$ *more*$]]\ [_{\text{A}'}$ *intelligent*$]]$ *than him.*

If *much* support takes place, the measuring expression may be a head:

(154) a. *Susan is more intelligent than her supervisor. Indeed she is* $[_{\text{AP}}\ [_{\text{DegP}}$ *very much* $[_{\text{Deg}'}$ *more*$]]\ [_{\text{A}'}$ *intelligent*$]]$ *than him.*

b. *Susan is more intelligent than her supervisor. *Indeed she is* $[_{\text{AP}}\ [_{\text{DegP}}$ *very* $[_{\text{Deg}'}$ *more*$]]\ [_{\text{A}'}$ *intelligent*$]]$ *than him.*

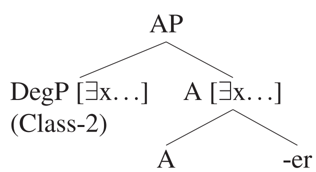
Finally, as was the case with *too*, semantically charged *much* may appear in the specifier of *more*. (155b) confirms our earlier conclusion that *much*, like other measure phrases, does not license a satellite.

(155) a. *Susan is more intelligent than her supervisor. Indeed she is* [$_{AP}$ [$_{DegP}$ *much* [$_{Deg'}$ *more*]] [$_{A'}$ *intelligent*]] *than him.*
b. **Susan is* [$_{AP}$ [$_{DegP}$ *much* [$_{Deg'}$ *more*]] [$_{A'}$ *intelligent*]] *than her supervisor to not know the answer.*

The proposed interpretation of specifier-head relationships also allows a straightforward account of degree modifiers that accompany morphological comparatives, such as *higher* in (156).

(156) *The flight path is eight miles higher than originally planned.*

Since the relevant part of this example has the structure in (157), deletion of an existential quantifier under specifier-head agreement is allowed.

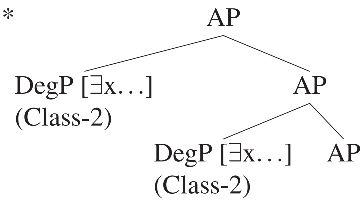
(157)


As we saw in the previous section, specifier-head agreement makes it possible to combine two degree expressions, but only if their restrictions are compatible. Hence, class-2 degree expressions involving the functions *At* or *Distance* do not tolerate another degree expression as their specifier, since they reduce the ordered set provided by the adjective to a point. Thus, the following are all ungrammatical:

(158) a. **John was* [$_{PP}$ [$_{DegP}$ *less/more/too much/a little/as much/very much/much* [$_{Deg'}$ *enough*]] [$_{P'}$ *into syntax*]] *to go to MIT.*
b. **John felt* [$_{PP}$ [$_{DegP}$ *less/more* [$_{Deg'}$ *a little*]] [$_{P'}$ *insecure about his hair*]] *than Bill.*
c. **John felt* [$_{PP}$ [$_{DegP}$ *too much/enough* [$_{Deg'}$ *a little*]] *insecure about his hair*]] *to go out.*
d. **John felt* [$_{PP}$ [$_{DegP}$ *as much/very much* [$_{Deg'}$ *a little*]] *insecure about his hair*]] *as Bill.*
e. **John felt* [$_{PP}$ [$_{DegP}$ *much* [$_{Deg'}$ *a little*]] *insecure about his hair*]].

So far we have assumed without argumentation that the stacking of class-2 degree items takes place in a left-branching structure like (148). But this structure must be contrasted with the right-branching alternative in (159). In the latter, the two degree expressions are not in a specifier-head configuration, and therefore the two existential quantifiers must be interpreted separately, resulting in vacuous quantification.

(159) *



The prediction, then, is that all the grammatical examples discussed above have a left-branching structure. There are three tests which show that this prediction is correct. First, the specifier of a class-2 item must accompany it when it appears in extraposition. Thus, there is a very sharp contrast between (160b) and (160c).

(160) a. *John is very much more into syntax than Gillian.*
b. *John is into syntax very much more than Gillian.*
c. **John is very much into syntax more than Gillian.*

Second, the left-branching structure is supported by coordination. As shown by the contrast below, the left-branch constituents we argue for can be coordinated, but constituents implied by a right-branching structure cannot:

(161) a. *Did you say that Gillian was* [*too much more*] *or* [*too much less*] *intelligent than Bill to get the job?*
b. **Did you say that Gillian was too much* [*more intelligent*] *or* [*less manipulative*] *than Bill to get the job.*

Third, although marginal in many cases, there are structures that allow movement of the proposed left-branch constituents in both English and Dutch:

(162) a. [*How much more*] *is John* [$_{PP}$ *t* [$_{PP}$ *into syntax*]] *than Gillian?*
b. [*Hoeveel meer*] *is Jan* [$_{AP}$ *t* [$_{AP}$ *gesteld op Marie*]] *dan op Susanne?*
how much more is John fond of Mary than of Susanne

In contrast, movement of constituents implied by a right-branching structure consistently leads to ungrammaticality:

(163) a. *Into syntax, John is certainly very much more.*
b. **More into syntax, John is certainly very much.*
c. *Very much more into syntax, John certainly is.*

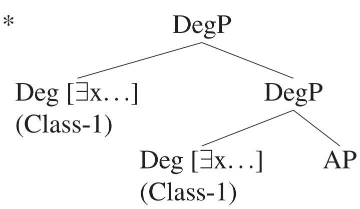
In sum, more than one degree expression can occur in the extended adjectival projection as long as each occupies the specifier position of the next. The examples given so far involved two degree expressions, but nothing excludes further recursion:

(164) *The projected flight path was too high, but it was eight miles less too high than originally calculated.*

4.4. Impossible combinations

Since the syntax-semantics mapping limits recursion to specifier-head relationships, degree expressions are otherwise in complementary distribution. As already demonstrated, two class-2 modifiers cannot be combined in a right-branching structure (see (159)). A similar point holds of class-1 degree expressions. Such expressions are heads selecting an adjectival complement. Consequently, if we abstract away from cases involving *much* support, they can only be combined in a structure like (165). Although in this structure the c-selectional requirements of the two degree heads are satisfied (in line with Grimshaw (1991) we assume that DegPs are adjectival in nature), reduction of the number of existential quantifiers under specifier-head agreement is impossible. The result is vacuous quantification and hence ungrammaticality.

(165) *

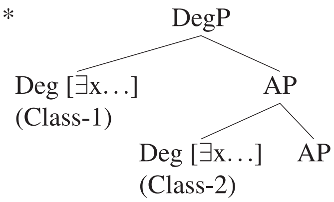


Indeed, a sequence of two class-1 degree items is always excluded. As we have seen (cf. (132), (138b) and (143b)), it is impossible to insert a class-1 degree item in the specifier of another class-1 degree item. One example, repeated from (143b), is given in (166a). The same example shows that the structure in (165) is ungrammatical. Otherwise, it would be ruled in under the structure in (166b).

(166) a. **John is* [$_{\text{DegP}}$ *very* [$_{\text{Deg}'}$ *too big*]] *for this sweater.*
b. **John is* [$_{\text{DegP(very)}}$ *very* [$_{\text{DegP(too)}}$ *too big*]] *for this sweater.*

Another structure that is categorically excluded is one in which a class-1 degree expression is attached after attachment of a class-2 modifier. Clearly, the degree head and the modifier are not in a specifier-head relationship, so that it is impossible to delete one of them under identity with the other. As before, this results in vacuous quantification:

(167) *

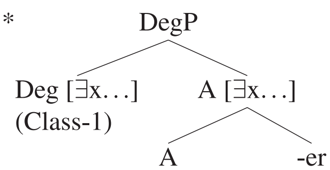


Structures of this type are indeed not attested. The relevant examples were discussed earlier in connection with the impossibility of inserting a class-1 degree item in the specifier of a class-2 modifier. We repeat (154b), which illustrates this, as (168a). The same example shows that (167) is ungrammatical under the parse in (168b).

(168) a. **She is* [$_{AP}$ [$_{DegP}$ *very* [$_{Deg'}$ *more*]] [$_{A'}$ *intelligent*]] *than him.*
b. **She is* [$_{DegP}$ *very* [$_{AP}$ [$_{DegP}$ *more*] [$_{A'}$ *intelligent*]]] *than him.*

A final structure ruled out across the board is one in which a class-1 degree expression combines with a morphological comparative. There are two candidate structures for such expressions. In the first the AP containing the comparative morpheme is a complement of the degree head:

(169) *



This structure is ruled out by the ban on vacuous quantification, since the adjective and the degree head are not in a specifier-head relation. Indeed examples like (170) are ungrammatical.

(170) **The flight path is* [$_{DegP}$ *very* [$_{AP}$ *higher*]] *than originally planned.*

A second potential structure would have *very* in the specifier of *higher*, where it would fail to satisfy its c-selectional requirements (cf. 171a). Hence, the only way to combine a morphological comparative with a class-1 degree expression is by inserting dummy *much*, as in (171b):

(171) a. **The flight path is* [$_{AP}$ *very* [$_{A'}$ *higher*]] *than originally planned.*
b. *The flight path is* [$_{AP}$ [$_{DegP}$ *very much*] [$_{A'}$ *higher*]] *than originally planned.*

To summarize, we have argued that the mapping principles that relate syntactic and semantic representations are sensitive to the essentially syntactic relation between a specifier and its head. Only if two degree expressions entertain such a syntactic relation can one of the existential quantifiers be deleted under identity with the other. Hence, it is only possible to combine two degree expressions in the extended adjectival expressions if this syntactic relation obtains.

5. Summary and conclusion

The literature on degree expressions contains proposals according to which degree expressions are modifiers (cf. Bresnan 1973 and Jackendoff 1977) and proposals according to which they are heads (cf. Abney 1987, Zwarts 1992 and Corver 1997a, b, Kennedy (1997, 2000a). We have argued that both positions are correct, but for different classes of degree expression. Class-2 degree expressions are modifiers, but class-1 degree expressions are heads. In addition, a class of morphological degree expressions was identified. We have shown that an attempt to reduce the formal classes to semantic classes must fail. Instead, the sole defining property of class-1 expressions is that they are functional heads which c-select a complement of a specific category. Class-2 expressions, being modifiers, lack such selectional requirements. Morphological degree expressions, finally, are dependent morphemes: they select for a morphological host.

In the final part of the paper we have argued that the stacking of degree expressions is constrained by an LF mapping rule that allows two existential quantifiers in a specifier-head configuration to be reduced to one. Whenever this mapping rule fails to apply, combining two degree expressions gives rise to vacuous quantification.

University College London
University College London
Utrecht University and Leiden University

References

Abney, Stephen (1987). The English noun phrase in its sentential aspect. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Atlas, Jay David (1984). Comparative adjectives and adverbials of degree: an introduction to radically radical pragmatics. *Linguistics and Philosophy* 7: 347–377.

Bach, Emmon (1986). The algebra of events. *Linguistics and Philosophy* 9: 5–16.

Baker, Mark (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago: The University of Chicago Press.

Bartsch, Renate and Theo Venneman (1973). *Semantic Structures: A Study in the Relation between Syntax and Semantics*. Frankfurt: Athenaeum Verlag.

Bierwisch, Manfred (1989). The semantics of gradation. In *Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation*, Manfred Bierwisch and Ewald Lang (eds.), 71–261. Berlin: Springer-Verlag.

Borer, Hagit (2000). Exoskeletal vs. endoskeletal explanations: syntactic projections and the lexicon. Paper presented at the UC San Diego conference Explanation in Linguistics, December 1999.

Bowers, John (1975). Adjectives and adverbs in English. *Foundations of Language* 13: 529–562.

— (1987). Extended Xbar theory, the ECP and the Left Branch Condition. *Proceedings of the Seventh West Coast Conference on Formal Linguistics 6*, Stanford Linguistics Association, Stanford University, Stanford, California.

Bresnan, Joan (1973). Syntax of the comparative clause construction in English. *Linguistic Inquiry* 4: 275–343.

Carston, Robyn (1995). Quantity maxims and generalised implicature. *Lingua* 69: 213–244.

— (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell Publishing.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Corver, Norbert (1990). The syntax of left branch constructions. Unpublished Ph.D. dissertation, Tilburg University.

— (1997a). Much-support as a last resort. *Linguistic Inquiry* 28: 119–164.

— (1997b). The internal syntax of the Dutch extended adjectival projection. *Natural Language and Linguistic Theory* 15: 289–368.

Doetjes, Jenny (1997). Quantifiers and selection: On the distribution of quantifying expressions in French, Dutch and English. Unpublished Ph.D. dissertation, Leiden University.

Cresswell, Maxwell (1976). The semantics of degree. In *Montague Grammar*, Barbara Partee (ed.), 261–292. New York: Academic Press.

Don, Jan (1993). Morphological conversion. Unpublished Ph.D. dissertation, Utrecht University.

Emonds, Joseph (1976). *A Transformational Approach to English Syntax*. New York: Academic Press.

— (1985). *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.

Grice, Paul (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

Grimshaw, Jane (1979). Complement selection and the lexicon. *Linguistic Inquiry* 10: 279–325.

— (1991). Extended projection. Ms., Brandeis University.

— (1997). Projection, heads and optimality. *Linguistic Inquiry* 28: 373–422.

Hellan, Lars (1981) *Towards an Integrated Analysis of Comparatives*. Tübingen: Narr.

Horn, Laurence (1989). *A Natural History of Negation*. Chicago: The University of Chicago Press.

Jackendoff, Ray (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

— (1977). *X-bar Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

— (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Kamp, Hans (1975). Two theories of adjectives. In *Formal Semantics of Natural Language*, Edward Keenan (ed.), 123–155. Cambridge: CUP.

Kennedy, Christopher (1997). Projecting the adjective: the syntax and semantics of gradability and comparison. Ph.D. dissertation, University of California, Santa Cruz (published 1999 by Garland).

— (2000a). Polar opposition and the ontology of 'degrees'. *Linguistics and Philosophy* 24: 3370.

— (2000b). Lectures on the syntax and semantics of gradability and comparison. Presented at Kyushi University.

Kennedy, Christopher and Louise McNally (1999). From event structure to scale structure: degree modification in deverbal adjectives. In *Proceedings of SALT 9*, Cornell Linguistics Circle Publications, Cornell University, Ithaca, NY.

Klein, Ewan (1980). Semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4: 1–45.

Krifka, Manfred (1986). Nominal-Referenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen. Unpublished Ph.D. dissertation, Ludwig-Maximilians-Universität, München.

Link, Godehart (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, Use and Interpretation of Language*, Rainer Bäuerle, Christophe Schwarze and Arnim von Stechow (eds.), 302–323. Berlin: Walter de Gruyter.

Marantz, Alec (1997). No escape from syntax. In *University of Pennsylvania Working Papers in Linguistics 4.2: Proceedings of the 21st Annual Penn Linguistics Colloquium*, Alexis Dimitriadis, Laura Siegel, Clarissa SurekClark, and Alexander Williams (eds.), 201225. Department of Linguistics, University of Pennsylvania.

Meier, Cécile (1999). Konsekutive Konstruktionen und Relative Modalit"at. Unpublished Ph.D. dissertation, University of Tübingen.

Moltmann, Friederike (1992). Coordination and comparatives. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.

Ross, Jonathan (1967). Constraints on variables in syntax. Unpublished Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge.

Sapir, Edward (1944). Grading: A study in semantics. *Philosophy of Science* 11: 93–116.

Sperber, Dan and Deirdre Wilson (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Seuren, Pieter (1973). The comparative. In *Generative Grammar in Europe*, Ferenc Kiefer and Nicolas Ruwet (eds.), 528–564. Dordrecht: Reidel.

— (1978). The structure and selection of positive and negative gradable adjectives. In *Papers from the Parasession on the Lexicon*, *CLS 14*, Donka Farkas, Wesley Jacobsen and Karol Todrys (eds.), 336–346. Chicago: Chicago Linguistic Society.

— (1984). The comparative revisited. *Journal of Semantics* 3: 109–141.

Trommelen, Mieke and Wim Zonneveld (1986). Dutch morphology: Evidence for the right-hand head rule. *Linguistic Inquiry* 17: 147–170.

von Stechow, Arnim (1984a). Comparing semantic theories of comparison. *Journal of Semantics* 3: 1–77.

— (1984b). My reaction to Cresswell's, Hellan's and Seuren's comments. *Journal of Semantics* 3: 183–199.

Zwarts, Joost (1992). X′-syntax – X′-semantics. Unpublished Ph.D. dissertation, Utrecht University.