Decomposing Antonyms?[*]

Irene Heim
Dept. of Linguistics and Philosophy
Massachusetts Institute of Technology

heim@mit.edu

Abstract

Are the marked members of antonym-pairs such as *long – short* decomposed in the syntax? Büring has recently argued that they are, on the basis of evidence about the distribution of Rullmann-ambiguities and crosspolar anomalies. But the readings of marked antonyms in the complements of matrix modals seem to argue for the opposite conclusion. The dilemma that results defies a simple solution. Perhaps it tells us something about the workings of Comparative Deletion.

1 Introduction

This paper considers the question of whether the marked members of antonym pairs like *long – short* are decomposed in the syntax. Daniel Büring has recently updated and substantially expanded an argument first envisaged by Rullmann (1995), supporting the conclusion that the surface form *short* sometimes spells out a collocation of two abstract items **little** and **long** which do not form a semantic constituent (Büring 2007a, b). I will quickly review Büring's arguments and the analysis that he takes them to support. I then show that this analysis overgenerates unattested readings in a class of examples where *shorter* is systematically not equivalent to *less long*. These data were previously discussed in Heim (2006), where they led me to conclude that antonyms are <u>not</u> decomposed.

Pairs of antonyms such as *long – short*, *old – young*, *fast – slow* are characterized by the truth-conditional equivalence of (1a - c), where α^+ stands for the first member in one of these pairs (e.g., *long*, *old*, *fast*) and α^- for its second member (*short*, *young*, *slow*).

(1)   a.   x is α^+er than y is.

       b.   y is α^-er than x is.

       c.   y is less α^+ than x is.

One way to capture these equivalences (and other patterns of reasoning with antonyms) involves interpreting α^+ and α^- as negations of each other, as in the following set of lexical entries.

(2)   a. $[\![\textbf{long}]\!] = \lambda x.\ (0, L(x)]$

       b. $[\![\textbf{short}]\!] = \lambda x.\ (L(x), \infty)$

       c. $[\![\textbf{er}]\!] = \lambda A.\ \lambda B.\ B \supset A$

       d. $[\![\textbf{less}]\!] = \lambda A.\ \lambda B.\ B \subset A$

("L" stands for the measure function length, and sets of degrees are given in interval notation.) In the syntactic structure at Logical Form (LF), the comparative morpheme **er** combines first with the *than*-complement (a possibly elliptical wh-clause) and then with the matrix-clause it has scope over. Both of these clauses denote sets of degrees, and the comparative is true iff the matrix-set is a proper superset of the *than*-clause set. Given that $[\![\alpha^-]\!]$ (e.g., $[\![\textbf{short}]\!]$) maps an individual to the complement of the set to which $[\![\alpha^+]\!]$ (e.g., $[\![\textbf{long}]\!]$) maps it, the equivalences between (1a) and (1b) and between (1b) and (1c) boil down to the set-theoretic fact that A is a proper subset of B iff the complement of B is a proper subset of the complement of A.

The entries in (2) exemplify what I call a "<u>lexical</u> negation theory of antonymy". In such a theory, the denotations of *long* and *short* are related by the operation of (predicate) negation, but there is no meaningful part of the syntactic representation of *short* that expresses this operation. I contrast this with a "<u>syntactic</u> negation theory of antonymy". This type of theory has no listing for a minimal meaningful item **short** in its lexicon; instead it always generates the surface form *short* by spelling out a collocation of two meaningful units, one of which is the same as what spells out *long* and the other is some kind of negation operator. The proposal in Büring (2007a, b) is an instance of such a theory. It posits an abstract item (called **little**) which expresses predicate negation and which figures in the pre-spell-out representations of both *short* and *less*. *shorter* and *less long* in fact are alternative spell-outs of the same underlying representations. I present Büring's analysis in section 2, review its motivation in section 3, and show that it overgenerates in section 4. In section 5, I introduce a modification of the proposal from Heim (2006) and suggest that it might account for the data under suitable assumptions about Comparative Deletion. This conclusion, however, will be very tentative, and the main purpose of the paper is to draw attention to a difficult dilemma.

2   A syntactic negation theory of antonymy

Büring's proposal can be summed up in the following lexical entries (3) and spell-out rules (4).[1]

(3)   a. $[\![\textbf{long}]\!] = \lambda x.\ (0,\ L(x)]$
       b. $[\![\textbf{er}]\!] = \lambda f.\ \lambda A.\ \lambda B.\ f(B) \supset f(A)$
       c. $[\![\textbf{little}]\!] = \lambda A.\ -A$
       d. $[\![\textbf{much}]\!] = \lambda A.\ A$

(4)   a.   **long** > *long*
       b.   **er** > *er*
       c.   **er little** > *less*
       d.   **little long** > *short*

A simple comparative not involving **little** is analyzed in (5). Here and henceforth, grey material feeds spell-out but not interpretation, struck-out words (and lambdas and indices) are interpreted but not spelled out, struck-out grey material is neither interpreted nor spelled out, and the rest is both interpreted and spelled out. The representations are generated by a copy-theory of movement, with rightward covert movement of DegP and late merger of the *than*-clause (cf. Bhatt & Pancheva 2004). DPs (*the rope*, *the wire*) are abbreviated by single letters and assumed to reconstruct for interpretation. For ease of reference, I use the labels "PF" and "LF" for the truncated versions of the syntactic representation which are missing respectively the material that doesn't affect spell-out and the material that isn't interpreted. (But strictly speaking, there is only one representation that is interpreted at both interfaces.)

(5)   a.   The rope is longer than the wire is.
       b. $[\lambda_1.$ r is [er ~~much~~]$_1$ ~~r~~ long] ~~er much~~ than ~~wh~~ $[\lambda_2.$ w is ~~wh$_2$~~ ~~w long~~]
       c.   PF: r is er long than w is
             er long > *er long* (> *longer*)
       d.   LF: $[\lambda_1.\ t_1$ r long] [er much] $[\lambda_2.\ t_2$ w long]
             $[\![\textbf{long}]\!](r) \supset [\![\textbf{long}]\!](w)$

The second line (5b) shows the full syntactic representation with all copies of movement chains and all elided material. The third line (5c) ("PF") shows how to get to the surface form, by deleting from (5b) all struck-out material as well as all lambdas, indices and brackets, and then spelling out what is left. The fourth line (5d) ("LF") shows how to compute the meaning from (5b), by first deleting all the grey material and

---

[1]There are minor differences from Büring's own implementation. Throughout this paper, I generate the adjective's subject as its innermost argument. Basically, I just want to abstract away from the question of how exactly the adjective's arguments are introduced and ordered.

then plugging in denotations form the lexicon and using semantic composition rules and logical inferences as usual. Note that the combination of Büring's meaning for **er** with his **much** amounts to our previous simpler meaning for **er** in (2c). The reason why **much** is needed is to have a uniform semantic type for **er** which allows it to also combine directly with **little** (see below).[2]

Now let us turn to derivations with **little**. One syntactic representation we can generate is just like (5b) except with **little** replacing **much**. This is in (6a) and looks like (6b) to the spell-out rules and like (6c) to the semantics.

(6)   a.   [λ_1. r is [er little]$_1$ ~~r~~ long] ~~er little~~ than ~~wh~~ [λ_2. w is ~~wh$_2$ w long~~]
      b.   PF: r is er little long than w is
      c.   LF: [λ_1. t$_1$ r long] [er little] [λ_2. t$_2$ w long]

Since the meaning of the complex **er little** is that of **less** in (2d), (6c) says that ⟦**long**⟧(r) ⊂ ⟦**long**⟧(w), i.e., the rope is shorter than the wire, or equivalently, the rope is less long than the wire. But how is (6b) spelled out? Büring's proposal says that it can actually be spelled out in two different ways. We can either use the spell-out rules in (4b) and (4d), so **er** > *er* and **little long** > *short*, which gives us *er short = shorter*. Or we can use the rules in (4c) and (4a), so **er little** > *less* and **long** > *long*, which yields *less long*. (6a) and its meaning ⟦**long**⟧(r) ⊂ ⟦**long**⟧(w) are therefore paired with both the surface strings *the rope is shorter than the wire is* and *the rope is less long than the wire is*. In general, spell-out operates as a series of replacement operations that rewrite substrings of the original string. Each such operation must be sanctioned by a spell-out rule, and everything in the original string must eventually be rewritten. Other than that, the rules apply optionally and freely, and in particular, those rules like (4c, d), which amalgamate two input items into one output item, can (but need not) apply whenever a string contains these input items next to each other.

There is another syntactic representation that also spells out in the same two ways and also has the same semantic interpretation. Here **little** is generated not as the argument of **er** but adjoined to the adjective phrase.

(7)   a.   [λ_1. r is [er ~~much~~]$_1$ little ~~r~~ long]

                  ~~er much~~ than ~~wh~~ [λ_2. w is ~~wh$_2$ little w long~~]
      b.   PF: r is er little long than w is
      c.   LF: [λ_1. t$_1$ little r long] [er much] [λ_2. t$_2$ little w long]

Since **little long** effectively means **short** as interpreted in (2b) and **er much** means **er** as in (2c), (7c) says again that ⟦**long**⟧(r) ⊂ ⟦**long**⟧(w). And (7b) is indistinguishable from

---

[2]This is how Büring overcomes the compositionality problem which plagued an earlier version of this analysis in Rullmann (1995). See Heim (2006).

(6b) to the spell-out component, so both pronunciations (*the rope is shorter than the wire is* and *the rope is less long than the wire is*) can be derived as before.

A key feature of this theory is that two hierarchically different arrangements of the items **er**, **little**, **long** in the syntax can end up with the same pronunciation. Of particular interest to us is the fact that the surface word *short* can spell out a **little** and **long** which never form a constituent with each other in (6a) and whose interpreted copies are quite far apart from each other in (6c). We have yet to see the motivation for this theory, but when we do so below, we particularly want to scrutinize the justification for this type of semantics-morphology mismatch.

3   Crosspolar anomalies and nomalies

Büring's analysis, as introduced above, can be seen as an updated and compositional version of a theory entertained by Rullmann (1995, attributed there to a suggestion by Barbara Partee). Part of the motivation that Büring gives for it (see Büring 2007a) also goes back to Rullmann and turns on the ambiguity of sentences like (8a, b).

(8)   a.   He is less tall than he is allowed to be.
        b.   He is shorter than he is allowed to be.

I will concentrate here on another argument, however, which was newly contributed by Büring (2007b) and in some ways is more straightforward, since it is based on simpler structures that don't involve ellipsis or modal operators. The backdrop for this argument is Kennedy's (2001) work on what he dubbed "crosspolar anomaly", the deviance of comparatives like (9b) compared to (9a, c).

(9)   a.   The rope is longer than the gap is wide.
        b.   *The rope is longer than the gap is narrow.
        c.   The rope is shorter than the gap is narrow.

Kennedy, working with a non-decompositional analysis of antonyms essentially like that in (2) above, proposes that the comparative operators **er** and **less** cannot relate two sets of degrees that are on opposite ends of a scale. I.e., they can compare two initial segments of the spatial distance scale[3], as in (9a), or two final segments as in (9c), but not an initial with a final segment as in (9b). Perhaps this is because such comparisons will be necessarily false, or because the comparative morphemes actually carry a presupposition that its two relata must be both initial or both final segments. We need not decide here on the exact nature of the constraint, just on this descriptive

---

[3]By the "spatial distance scale", I mean the shared ordered set of degrees into which the measure functions associated with **long**, **wide**, **high**, etc. map their arguments.

generalization. Büring accepts Kennedy's constraint, but points out that it is *prima facie* falsified by the missing fourth member of the paradigm in (9), namely (10).

(10)   The rope is shorter than the gap is wide.

By Kennedy's constraint, (10) should be just as deviant as (9b), but it is actually grammatical. Why is that?

Büring's answer is that (10) is good because it spells out two distinct syntactic representations, only one of which violates the constraint. The bad derivation is (11) and the good one is (12) (compare (7) and (6) above).

(11)   a. $[\lambda_1.$ r is [er ~~much~~]$_1$ little ~~r~~ long] ~~er much~~ than ~~wh~~ $[\lambda_2.$ g is ~~wh$_2$~~ ~~g~~ wide]
         b.   r is er little long than g is wide
         c. $[\lambda_1.$ t$_1$ little r long] [er much] $[\lambda_2.$ t$_2$ g wide]

(12)   a. $[\lambda_1.$ r is [er little]$_1$ ~~r~~ long] ~~er little~~ than ~~wh~~ $[\lambda_2.$ g is ~~wh$_2$~~ ~~g~~ wide]
         b.   r is er little long than g is wide
         c. $[\lambda_1.$ t$_1$ r long] [er little] $[\lambda_2.$ t$_2$ g wide]

(11) violates Kennedy's constraint because there is an interpreted **little** in the matrix clause but none in the *than*-clause. As a result (seen in (11c)), the matrix clause denotes a final segment of the distance-scale and the *than*-clause an initial segment. (12), on the other hand, is just fine: neither the matrix nor the *than*-clause contain an interpreted **little** (see (12c)), and so we are comparing two initial segments. Both of the representations in (11) and (12) spell out either as (10) or as *The rope is less long than the gap is wide*. Indeed, both of these sentences are fine, as predicted by the fact that they both have at least one derivation that respects the semantic constraint. The crucial point of Büring's argument is that (10) could not be grammatical unless it allowed the derivation in (12). This in turn shows that **little long** must be able to surface as *short* even when **little** and **long** are not sisters in the syntactic structure and not composed with each other at LF.

4   When *shorter* cannot paraphrase *less long*

A general prediction of Büring's analysis is that *shorter* and *less long* can be substituted for each other in all environments without any change in meaning or in available readings. This must be qualified a bit. Büring (2007a) actually says that while both [**er little**] **long** and [**er much**] [**little long**] may surface as either *less long* or as *shorter*, there is a preference for the former to surface as *less long* and the latter to surface as *shorter*. In other words, the spell-out rules in (4c) and (4d) <u>can</u> ignore the syntactic bracketing, but this is a somewhat marked option, and the preferred or unmarked way to

spell these structures out is to amalgamate only syntactic sisters. To implement this preference formally, we would presumably have to move away from the simplistic assumption that spell-out rules don't see bracketing at all. Instead maybe spell-out is generally cyclic and only the intervention of special structure-flattening operations can make it seem otherwise. But be that as it may, Büring still is committed to the view that both spell-outs are always possible, even if one is more marked. This is enough to make the data I now will present a problem for his theory. The relevant readings in these examples will be strictly impossible, not just dispreferred.

Consider the following scenario[4]: Polly and Larry both are supposed to be in Boston by 8pm at the latest. They are both driving there. It is 5:30pm, and Polly is just setting out from Providence RI, whereas Larry is leaving from New Haven CT, about twice as far from Boston as Providence. Sentence (13) is a true description of this state affairs.

(13)   Larry needs to drive faster than Polly needs to drive.

The reason why Larry needs to drive faster is that he needs to cover more distance within the same two and a half hours. Another true statement, in fact a possible paraphrase of (13), is (14).

(14)   Polly needs to drive less fast than Larry needs to drive.

Polly needs to drive less fast because she needs to cover less distance. But (15) is not a paraphrase of (13) and (14) and in fact claims something quite different.

(15)   Polly needs to drive more slowly than Larry needs to drive.

For (15) to be true, there would have to be some kind of penalty or disadvantage to arriving early. If this is not a feature of our scenario, then (15) is actually false: Polly may drive more slowly, but she doesn't need to. So (15) and (14) differ in truth-value here. Upon reflection, we may judge that (14) could also have been understood to express the same falsehood that (15) does. But the fact remains that (15) does not share the true reading of (14), and this is unexpected if Büring's analysis is right.

Before we look more closely at how this type of case counterexemplifies Büring's predictions, let's have a second example that makes the same point. This is the tale of the sexist boarding school. In the dormitory of this school, rooms are allocated according to seniority. The top-floor single rooms, which are the most attractive and private, are reserved for the oldest pupils. Specifically, they are restricted to girls who are 16 or older and to boys of 14 or older. Commenting on the sexism of this policy, I might truthfully utter (16).

---

[4]This and the following example were introduced in Heim (2006).

(16)   The older pupils get the single rooms on the top floor,
         but girls have to be older than boys (to qualify for this privilege).

I might also utter (17) to express the same thought.

(17)   ..., but boys have to be less old than girls (to qualify ...).

But if I were to say (18), that would sound quite wrong.

(18)   #..., but boys have to be younger than girls (to qualify ...).

Boys don't <u>have</u> to be younger, they just <u>can</u> be! Again, (17) has a true reading which (18) lacks.

A further piece of data concerns the licensing of the German modal *brauchen* ('need'), which is an NPI when taking infinitive complements. The following German translations of (14) and (15) differ in grammaticality.[5]

(19)   a.   Polly braucht weniger schnell zu fahren als Larry.
               Polly needs   less   fast   to drive than Larry
         b.   *Polly braucht langsamer zu fahren als Larry.
                Polly needs   slower   to drive than Larry

This indicates that some kind of negative element can scope over and license *brauchen* in (19a), but not in (19b).

Büring's analysis does not block the unattested readings of (15), (18), and (19b). Since *slower* and *younger* can spell out the syntactic configuration [**er little**] **fast/old**, it should be possible to have the following derivation for (15).

(20)   a.   [λ_1. p needs to ~~p~~ drive [er little]$_1$ fast]
               ~~er little~~ than ~~wh~~ [λ_2. l needs to ~~l~~ drive ~~wh$_2$ fast~~]
         b.   PF: p needs to drive er little fast than l needs to drive
               er little fast > *er slow* (> *slower* or *more slowly*)
         c.   LF: [λ_1. need p drive t$_1$ fast] er little [λ_2. need l drive t$_2$ fast]
                   $\{d: \forall w \in Acc(@): Speed_w(p) \geq d\} \subset$
                       $\{d: \forall w \in Acc(@): Speed_w(l) \geq d\}$

---

[5](14) and (15) can also be translated into German with the modal *müssen* ('must'), in which case the judgments parallel the ones for English: both sentences are grammatical, but only one has the reading true in the scenario.

The meaning of the LF in (20c) is that Polly's required minimum speed is below Larry's required minimum speed. This is the meaning of (13) and is true in the scenario. But intuitively this true reading is unavailable for the *more slowly* sentence in (15). The analysis overgenerates.

What is the difference between the "cross-polar nomaly" case exemplified in (10) above and the current example? Descriptively, it seems that negative antonyms (*short*, *slow*) can spell out a **little** that is semantically composed with the comparative operator, but only as long as no modal operator scopally intervenes between the comparative operator and the adjective. I do not see a natural way to make Büring's analysis sensitive to this distinction. The movement of comparative operators, after all, is <u>covert</u> movement: *-er* is always spelled out in the AP even when its logical scope includes higher operators, and *less* also evidently remains within the modal's surface complement even when the negation in *less* outscopes the modal. (We see this in the true readings of (14), (17).) In covert movement, the copy that is relevant to spell-out is (only) the one at the bottom of the chain. The surface string is effectively determined by the structure that obtains <u>before</u> DegP movement, and this looks the same in (20) as it does in (12): **er**, **little**, Adj start out in exactly the same configuration. PF-operations do not "see" the top of the DegP-movement chain and therefore cannot be sensitive to whether it eventually crosses a modal or not.

5   *little* and its distribution

The juxtaposition of the data in section 4 with the data in section 3 creates a tough dilemma, and I still don't know what is the best way to approach it. The suggestion I will make here is quite tentative and, even if on the right track, would take a lot more work to implement properly. I will start by reformulating Büring's insight about the distribution of crosspolar anomaly as a generalization about the distribution of silent instances of *little*. I will then speculate that this distribution might be understood within a theory of Comparative Deletion, but I do not yet have a concrete idea of what such a theory would be.

Clearly we cannot simply go back to the lexical theory in (2) or fix up Büring's theory by turning the mere preference for one spell-out over the other into an inviolable requirement. Either of these moves would get us the correct predictions about the sentences in (15), (18), and (19), but it would fail to capture the grammaticality of the crosspolar nomaly (10), and it would also fail to predict that *less*-comparatives such as (14) and (17) are actually ambiguous. We need something in between, a theory that permits *less*-α^+-comparatives to express a proper superset of the readings of the corresponding *more*-α^--comparatives. In Heim (2006), I accomplished this by decomposing *less* but not α^-. *less* always spelled out **er** and a scopally mobile **little**, but *short* was just the lexical item **short**. For reasons that will become clearer soon,

however, this proposal is not suited to make sense of Büring's crosspolar nomaly data. So here I will pursue a variation which posits **little** in the sources of both *less* and *short*, though a mobile **little** in one and an immobile one in the other.

Here is my lexicon.

(21) a. $[\![$**long**$]\!] = \lambda x.\ (0, L(x)]$
      b. $[\![$**er**$]\!] = \lambda A.\ \lambda B.\ B \supset A$
      c. $[\![$**little**$]\!] = \lambda A.\ -A$
      d. $[\![$**little***$]\!] = \lambda d.\ \lambda A.\ d \notin A$

The first two entries are the same as in the lexical theory (2), the third is Büring's **little** from (3c), and the fourth is a differently Schönfinkeled variant from Heim (2006). There are no entries for **less** or **short**. These surface forms come about by non-trivial spell-out rules as in (22).[6]

(22) a.   **er little*** > *less*
      b.   **little long** > *short*

little has type <dt,dt> and thus can modify the adjective in the structure [**er** [**little long**]], as in Büring-style derivations like (7) above. **little** cannot move out of this structure (the result would be uninterpretable for any type trace). Nor can **little** enter into the structure [[**er little**] **long**] (because my **er** is not typed to take it as an argument), so it can't hitch a ride that way either. This **little** is therefore frozen in place, and so we capture the fact that the negation in α^- is bound to scope below any modal. We derive the right prediction about (15), (18) and (19b). E.g., (15) has only one derivation, (23), which pairs it with the meaning that is false in our scenario.

(23) a. $[\lambda_1.$ p needs to ~~p~~ drive er$_1$ little fast]
           ~~er~~ than ~~wh~~ $[\lambda_2.$ l needs to ~~l~~ drive ~~wh$_2$~~ ~~little fast~~]
      b.   PF: p needs to drive er little fast than l needs to drive
           er little fast > *er slow* (> *slower* or *more slowly*)
      c.   LF: $[\lambda_1.$ need p drive t$_1$ little fast] er $[\lambda_2.$ need l drive t$_2$ little fast]
           $\{d: \forall w \in Acc(@): Speed_w(p) < d\} \supset$
                $\{d: \forall w \in Acc(@): Speed_w(l) < d\}$

---

[6]The spell-out rules in (22) make it look as if we need to distinguish **little** from **little*** at PF. However, this is not really necessary, since we can also rely on syntactic bracketing to make the right distinction. If spell-out is cyclic, it will automatically group **little** with the adjective and **little*** with **er**. (In the alternative of the next footnote, where there only is a single **little**, we have no choice but to rely on bracketing.)

(23c) says that the permissible maximum speed for Polly is below the permissible maximum speed for Larry. In other words, she is not allowed to drive as fast as he is allowed to – not true unless there is a penalty for arriving early.

A different behavior is found with **little***, the item that underlies *less*. This has a type to take the QR-trace of **er** as an argument and form with it another QRable phrase of type <dt,t>.[7] We can therefore generate two different readings for a *less*-comparative like (14), of which one paraphrases (13) and the other (15). The two derivations start with identical pre-movement structures, and the choice point is in the scoping of the **little***-phrase, above or below the modal.

(24)   a.   [λ_1. p needs to ~~[er₁ little*]~~ [λ_2. ~~p~~ drive [er₁ little*]₂ fast]]
            ~~er~~ than ~~wh~~ [λ_3. l needs to ~~[wh₃ little*]~~ [λ_4. ~~l~~ drive ~~[wh₃ little*]₄ fast~~]]

       b.   PF: p needs to drive er little* fast than l needs to drive
            er little* fast > *less fast*

       c.   LF: [λ_1. need [t_1 little*] [λ_2. p drive t_2 fast]]
            er [λ_3. need [t_3 little*] [λ_4. l drive t_4 fast]]
            $\{d: \forall w \in Acc(@): Speed_w(p) < d\} \supset$
                    $\{d: \forall w \in Acc(@): Speed_w(l) < d\}$

(25)   a.   [λ_1. ~~[er₁ little*]~~ [λ_2. p needs to ~~p~~ drive [er₁ little*]₂ fast]]
            ~~er~~ than ~~wh~~ [λ_3. ~~[wh₃ little*]~~ [λ_4. l needs to ~~l~~ drive ~~[wh₃ little*]₄ fast~~]]

       b.   PF: p needs to drive er little* fast than l needs to drive
            er little* fast > *less fast*

       c.   LF: [λ_1. [t_1 little*] [λ_2. need p drive t_2 fast]]
            er [λ_3. [t_3 little*] [λ_4. need l drive t_4 fast]]
            $\{d: \neg\forall w \in Acc(@): Speed_w(p) \geq d\} \supset$
                    $\{d: \neg\forall w \in Acc(@): Speed_w(l) \geq d\}$

---

[7]Having to posit two distinct **little**'s is clumsy. I think it could be avoided as follows. Suppose only the **little** of type <dt,dt> exists and we always base-generate the configuration [er [**little** [subj Adj]]]. We then can proceed with derivations as in (23), or else we can begin with a short movement of the adjective (including its subject), leaving a trace of type <d,t>. The resulting configuration will be [subj Adj] λP.[**er** [**little** P]]. From this point on, the newly created unit λP.[**er** [**little** P]] can move in the same way as the **little***-phrase in the text (e.g. in (24), (25). It is in fact semantically equivalent to it: after QRing out **er** and leaving a trace of type d, both [d **little***] and λP.[d [**little** P]] express the generalized quantifier λP. d $\notin$ P. To combine this approach with a natural story about spell-out, it makes sense to regard the movement of the adjective as an overt movement. We can then say that **little** always spells out with the structurally closest other item that is subject to spell-out. If the adjective has stayed in place, this will be the adjective and we obtain the antonym; if the adjective has moved out, the closest other item is **er** and we get **less**.

In general, when *less* surfaces, it spells out a **little***, and the negation in this interacts scopally with clause-mate modals. See Heim (2006) for a broader exposition (as well as Rullmann 1995 and Büring 2007a).

We have captured my data in section 4, but what about Büring's paradigm of cross-polar anomalies and nomalies? Let us warm up with (26), an example that has not presented any difficulty to previous analyses.

(26)   The rope is less long than the gap is wide.

(26) worked out fine for Kennedy, when *less* was treated as the primitive comparison operator in (2d), and it worked for Büring, since one of his parses for *less* is effectively an internally complex but otherwise identical counterpart of primitive **less**. What about the present system? Since (26) is not a crosspolar anomaly, it better contain a covert **little** or **little*** inside its *than*-clause. In this theory (unlike Büring's), any occurrences of **little**(*) in comparatives are always within one or the other argument of the comparison operator **er**; there is no option of **little**(*) forming part of the comparison operator itself. Even when we start out with the constituent [**er little***], **er** must eventually strand [**t little***] somewhere in the matrix clause before it reaches a place where it can be interpreted. Therefore, since one **little*** (spelling out in *less*) is manifestly present in the matrix, another matching one in the *than*-clause is needed to meet Kennedy's constraint. The derivation for (26), on its grammatical construal, must be something like (27).

(27)   a.   [λ_1. ~~[er₁ little*]~~] [λ_2. r is [er₁ little*]₂ ~~r~~ long]]
              ~~er~~ than ~~wh~~ [λ_3. ~~[wh₃ little*]~~] [λ_4. g is ~~[wh₃ little*]~~₄ ~~g~~ wide]]

         b.   PF: r is er little* long than g is wide
              er little* long > *less long*

         c.   LF: [λ_1. [t₁ little*] [λ_2. t₂ r long]] er [λ_3. [t₃ little*] [λ_4. t₄ g wide]]
              (L(r), ∞) ⊃ (W(g), ∞)

So far, so good, but how exactly does there manage to be a silent **little*** in the *than*-clause? I could have asked the same question about other *less*-comparatives to which I applied the analysis in Heim (2006). We never saw an overt reflex of the putative second **little*** in any of those examples. But perhaps this was not so mysterious, given that the examples all had some degree of uncontroversial ellipsis in the *than*-clause to begin with, e.g. VP-ellipsis, Comparative ellipsis, or at least Comparative Deletion (an AP-sized or AdvP-sized gap as in (14)). In those cases, then, it was not unreasonable to blame the silence of the second **little*** on whatever ellipsis or deletion process affected the surrounding material. (26) is a bit different in this regard. This looks like a sub-deletion comparative, normally thought to involve no further unpronounced material but the null wh-operator. So why is **little*** deleted there too? I must say that (26) isn't really "sub-deletion" after all (at least not if "sub-deletion" is by definition a structure with

nothing covert but a bare **wh**), but a kind of Comparative Deletion. Even though the adjective, *wide*, has stayed behind, there still is deletion of a larger phrase surrounding the **wh**, namely the *little*-phrase **wh little***. Presumably, this deletes under identity with the matching *little*-phrase in the main clause, in whatever way it is that Comparative Deletion operates.

Now we are all set to turn to the surprising crosspolar nomaly in (10) *The rope is shorter than the gap is wide*. As in (26), the present proposal leaves me only one way to avoid violating Kennedy's semantic constraint: I must posit a covert **little** or **little*** inside the *than*-clause, in order to convert the initial scale-segment denoted by **the gap wide** into a final segment comparable with the one denoted by the main clause **the rope short**. This is where I exploit the assumption that *short* does contain a **little** – not a **little** that scopes around, but still one that can antecede a matching one under Comparative Deletion. The derivation I have in mind is (28).

(28)   a. $[\lambda_1.$ r is er$_1$ ~~r~~ little long]

          ~~er~~ than [~~[wh little]~~ $[\lambda_2.$ g is ~~[wh little]$_2$~~ ~~g~~ wide]]

       b.   PF: r is er little long than g is wide

            er little long > *er short* (*shorter*)

       c.   LF: $[\lambda_1.$ t$_1$ r little long]] er [little $[\lambda_2.$ t$_2$ g wide]]

            $(L(r), \infty) \supset (W(g), \infty)$

Notice that the semantic vacuity of **wh** makes it possible to bracket it with **little** (not just with **little***) and still get an interpretable structure. This derivation yields the correct and non-anomalous meaning.

For the time being, this is the best story I have been able to come up with that reconciles the frozen scope of the negation in α^- (data in section 4) with the crosspolar nomaly of α^-*er than* α^+ (data in section 3). But my confidence in it is limited by my failure to understand the mechanics of Comparative Deletion. There are sophisticated syntactic proposals in the literature, such as Lechner (1999) and Kennedy (2002), but I don't know how to integrate them with the semantic and syntactic assumptions in this paper. This will have to await further research.

Two remarks before closing: It would have been nice not to have to invoke Comparative Deletion or any other specific ellipsis process at all and instead just assume there is a phonetically null **little** which can be generated freely and doesn't need an antecedent. (In that case, we also needn't have bothered to decompose **short** at all.) In its fully general version, this idea wreaks instant havoc. If covert **little**'s could be hiding below the surface wherever they fit into the structure and meaning, every *more*-comparative could be read as a *less*-comparative, and we would predict unattested ambiguity all over the place. Even restricting the silent-**little** option in some commonsensical way, say as a last resort to save a structure from ungrammaticality,

would not have worked. If a silent **little** in the *than*-clause can rescue (10) from anomaly, why not a silent **little** in either the matrix or the *than*-clause to do the same for (9b) **The rope is longer than the gap is narrow*? Clearly, the silent-**little** option must be tightly regulated by the grammar. Büring did away with this problem by not allowing any silent **little**'s at all (except when part of a larger elision including surrounding material); his comparative operator [**er little**] manages to negate two clauses with just a single (overt) **little**. But I have to face the problem in this form, and so I need to put my money on Comparative Deletion or some other appropriately constrained mechanism of recoverable deletion.

My second remark concerns the Rullmann-ambiguity in sentences of the type (8b) *He is shorter than he is allowed to be*. If the present proposal can be made to work for crosspolar nomaly, it will cover these cases too. A high **wh little** in the *than*-clause, if sanctioned by Comparative Deletion under identity with the **little** in *short*, will give us the less-than-maximum-reading. In Heim (2006), I suggested that this reading might have an entirely different source, connected to the same mechanism by which some speakers get an ambiguity even in the *more*-comparative *He is taller than he is allowed to be*. But what I have since learned from Büring's work is that this way out is not general enough. Whether or not it is plausible for the Rullmann examples, it throws no light at all on the crosspolar nomalies. There are no modals in those.

References

Bhatt, Rajesh and Roumyana Pancheva (2004) "Late Merge of Degree Clauses", *Linguistic Inquiry* **35**, 1–45.

Büring, Daniel (2007a) "*More* or *Less*", to appear in *Proceedings of CLS 43*.

Büring, Daniel (2007b) "Cross-Polar Nomalies", to appear in *Proceedings of SALT 17*.

Heim, Irene (2006) "*Little*", in C. Tancredi et al. (eds) *Proceedings of SALT16*, CLC: Ithaca, Cornell University.

Kennedy, Christopher (2001) "Polar Opposition and the Ontology of 'Degrees'", *Linguistics and Philosophy* **24**, 33–70.

Kennedy, Christopher (2002) "Comparative Deletion and Optimality in Syntax", *Natural Language and Linguistic Theory* **20**, 553–621.

Lechner, Winfried (1999) *Comparatives and DP-Structure*, PhD thesis, University of Massachusetts Amherst.

Rullmann, Hotze (1995) *Maximailty in the Semantics of Wh-Constructions*, PhD thesis, University of Massachusetts Amherst.