

Domains for Adverbs

Øystein Nilsen

Utrecht institute of Linguistics OTS, Trans 10, 3512 JK Utrecht, the Netherlands

Abstract

Speaker oriented adverbs are argued to be excluded from downwards entailing environments, i.e. they are argued to be positive polarity items. This behavior is derived by analyzing them as *domain shrinking* possible world quantifiers (cf. Kadmon and Landman (1993, L&P16) on negative polarity and domain *expansion*). It is argued that this analysis derives the major distributional properties of these adverbs. Particular attention is given to the adverb *possibly*.

Key words: adverb order, positive polarity items, functional heads, veridicality, monotonicity, syntax-semantics interface, possible worlds

1 Introduction

Why are sentential adverbs¹ ordered? This question has become relatively central to linguistic theory since the work of Cinque (1999) and Alexiadou (1997). The present paper represents an attempt to give it a partial answer. I will mainly concern myself with high (speaker oriented) adverbs, but some discussion of lower (temporal) adverbs and manner adverbs is included.

* I gratefully acknowledge valuable discussions with Raffaella Bernardi, Patrick Brandt, Paul Dekker, Fred Landman, Marika Lekakou, Eric Reuland, Robert van Rooy, Henriëtte de Swart, Kriszta Szendrői and Henk Zeevat about the contents of this paper. Usual disclaimers apply.

¹ Throughout, I use the notions ‘adverb’ and ‘adverbial’ interchangeably and in a purely descriptive sense. Alexiadou (2001) argues that the word class ‘adverb’ is questionable, Dowty (2000) questions the argument–adjunct distinction, and Julien (2000) argues that the notion of grammatical word is irrelevant for grammatical theory. Thus it is not obvious that any class of adverbs or adverbials can be adequately defined and distinguished from, say, auxiliaries and verbal affixes. On this, see also Nilsen and Vinokurova (2000).

It is well known that some adverbs are polarity items. For instance, adverbs like *yet* and *any longer* are negative polarity items (NPI). Perhaps less widely acknowledged is the fact that some adverbs are positive polarity items (PPI). In [van der Wouden \(1997\)](#) it is argued that the Dutch adverbs *al* ‘already’ and *niet* ‘not’ are PPIs. In this paper I will argue that surprisingly many sentential adverbs are PPIs. In particular, I will argue that ‘speaker oriented’ adverbs, like *allegedly*, *fortunately*, *possibly*, *evidently* should be treated as PPIs. I will use this fact to derive the relative ordering in which adverbs can cooccur. The observation is that some adverbs, in addition to being PPIs, themselves induce the environments that PPIs are excluded from. This results in a classification of adverbs according to which environments they can/cannot/must appear in, and which expressions can/cannot occur within their scope. It relates adverb ordering to Bellert’s observation ([Bellert, 1977](#)) that certain adverbs are degraded in e.g. questions, imperatives and antecedents of conditionals.

Classifying adverbs in this fashion already gives us predictions for possible adverb sequences which differ from existing theories ([Cinque, 1999](#); [Ernst, 2002](#)). However, I have not answered the initial question, i.e. why sentential adverbs are ordered, unless I can explain why a given adverb should be sensitive to a given property of its environment. In order to approach this question, I will adopt the framework developed in [Chierchia \(2001\)](#) for (negative) polarity items² and attempt to extend it to positive polarity phenomena. I will focus primarily on the adverb *possibly* here, but I think the approach can be extended to other speaker oriented adverbs as well.

Quantification is quite generally subject to contextual restriction ([Westerståhl, 1988](#)). For example, if I say that *Stanley invited everybody*, I do not generally mean that he invited everybody in the whole world. The idea is essentially that negative polarity *any* is lexicalized with a (universally closed) variable ranging over *domain expanding functions*. The application of such domain expansion is governed by a *strengthening condition*: the result of widening the quantificational domain for a quantifier in a proposition must be stronger than (entail) the corresponding proposition without widening.

This derives the major distributional patterns for NPIs in a rather plausible way. I suggest that this approach can be straightforwardly extended to positive polarity items if we assume that PPIs are associated with *domain shrinkage* rather than domain expansion. This approach is then argued to explain the distribution of some (though not all) sentential adverbs.

Before proceeding, I would like to point out one complication, also noted in ([Cinque, 1999](#), this volume). It is often stated that Cinque’s hierarchy ([Cinque,](#)

² Chierchia bases his theory on that of [Kadmon and Landman \(1993\)](#), and it is closely related to proposals by [Krifka \(1995\)](#) and [Lahiri \(1997\)](#). These proposals are discussed in section 3.2.

1999) should be given a semantic explanation.³ The complication is this: there are adverb–adjective pairs which are apparently synonymous, but, nevertheless, differ in distribution. For instance, the adverb *probably* cannot occur under *never*, whereas the (proposition embedding) adjective *probable* can. Compare (1a) to (1b):

- (1) a. *Stanley never probably ate his wheaties.
- b. It was never probable that Stanley ate his wheaties.

In order to make plausible a *semantic* explanation for the oddness of (1a), it seems one has to be able to establish that *probably* is not synonymous with *probable* in certain respects, otherwise (1b) should be odd as well. What is more, one has to show that whatever semantic difference there is between the two, that difference must be the culprit of the contrast in (1).⁴ I try to establish such a difference between the pair *possibly*, *possible* below.

The article is organized as follows. In section 2, I present the main data, that speaker oriented adverbs are excluded from the same type of environment that license NPIs. In section 3, I present the theories of polarity items that I will employ to derive this behavior and discuss some of their prediction with respect to adverb ordering. Section 4 develops a semantics for the adverb *possibly*, deriving its semantic differences with the adjective possible and, at the same time, deriving its distribution. In section 5, the final section, I discuss how the present account bears on current theories of adverb ordering phenomena.

2 Adverbs and Polarity

2.1 NPIs and speaker oriented adverbs

In Bellert (1977) it was observed that certain sentential adverbs have a narrower distribution than one might expect. For example, she observes that *speaker oriented adverbs* (henceforth SOA), such as evaluatives (*fortunately*), evidentials (*evidently*) and some modals (*possibly*) are degraded in questions (3a). As it turns out, these types of adverbs are also degraded in antecedents of

³ Such semantic derivation would not automatically remove the motivation for the hierarchy. Its motivation is that certain *syntactic* phenomena, such as verb placement in Romance, interact with adverb ordering in intriguing ways. I return briefly to this point in section 5.

⁴ I take this to be a problem for Ernst’s (2002) semantic approach to adverb distribution. This author takes adverb–adjective pairs like the ones under discussion to be, in fact, systematically synonymous.

conditionals (3b), imperatives (3c) under negation (3d), under clause-embedding predicates like *hope* (3e), as well as within the scope of monotone decreasing subject quantifiers like *no N* (3f). I use Norwegian examples to demonstrate this. As far as I have been able to determine, English works the same way, i.e the English translations are also degraded. In (3), ‘ADV’ represents any of the adverbs in (2):⁵

- (2) **heldigvis, tydeligvis, paradoksalt nok, ærlig talt,**
 fortunately evidently paradoxically honestly
 muligens, kanskje, sannsynligvis, angivelig, neppe
 possibly maybe probably allegedly hardly

- (3) [Norwegian]
 a. Spiste Ståle (*ADV) hvetekakene?
 ate S (*ADV) the-wheaties
 “Did Stanley (*ADV) eat the wheaties?”
 b. Hvis Ståle (*ADV) spiste hvetekakene, ...
 if S (*ADV) ate the-wheaties
 c. (*ADV) Spis (*ADV) hvetekakene!
 (*ADV) eat (*ADV) the-wheaties
 d. Ståle spiste (ADV) ikke (*ADV) hvetekakene.
 S ate (ADV) not (*ADV) the-wheaties
 “Stanley (ADV) didn’t (*ADV) eat the wheaties.”
 e. Jeg håper Ståle (*ADV) spiste hvetekakene.
 I hope S (*ADV) ate the-wheaties
 f. Ingen studenter (*ADV) spiste hvetekakene.
 no students (*ADV) ate the-wheaties

The same adverbs can appear in degree clauses (4a) and under clause embedders like *think* (4b) (compare (3e) with the verb *hope*), as well as, of course, ordinary declaratives (4c).

⁵ *kanskje* ‘maybe’ is better in questions than the other ones. However, when *maybe* occurs in questions, as in (i), the question is not whether or not it is *possible* that S ate his wheaties, rather, it seems to be equivalent to the same question with *maybe* removed.

(i) Did Stanley (maybe) eat his wheaties (, maybe)?

I have no suggestions as to why this should hold here (but see [van Rooy \(2002\)](#) on polarity items and questions). One can also find occurrences of *probably* in antecedents of conditionals which are not that bad.

(ii) If Le Pen will probably win, Jospin must be disappointed.

I take the slipperiness of some these intuitions to be comparable to that found with relative adverb ordering. Consequently, I will try to stick to phenomena for which the intuitions are sharper.

- (4)
- a. Ståle var så sulten at han (ADV) spiste hvetekaker.
S was so hungry that he (ADV) ate wheaties
 - b. Jeg tror Ståle (ADV) spiste hvetekakene.
I think S (ADV) ate the-wheaties
 - c. Ståle spiste (ADV) hvetekakene.
S ate (ADV) wheaties

It seems that SOA are excluded from the environments that license negative polarity items (NPI) like English *any*, Greek *kanena* or Dutch *ook maar iets*. Thus, SOA appear to be positive polarity items (PPI). I give Dutch examples in (5) and Greek ones in (6).⁶

- (5) [Dutch]
- a. Heeft Jan **ook maar iets** gegeten?
has J anything eaten
 - b. Als je **ook maar iets** hebt gegeten, . . .
if you anything has eaten
 - c. Eet **ook maar iets**!
eat anything
“Eat (something)!”
 - d. Er heeft geen enkele student **ook maar iets** gegeten.
there has no single student anything eaten
 - e. Ik hoop dat Jan **ook maar iets** heeft gegeten.
I hope that J anything has eaten
 - f. *Ik denk dat Jan **ook maar iets** heeft gegeten.
I think that J anything has eaten

- (6) [Greek]
- a. Efage o Yiannis **kanena**?
ate J anything
 - b. An o Yiannis efage **kanena**, . . .
if J ate anything
 - c. Fae **kanena**!
eat anything
 - d. Elpizo na efage o Yiannis **kanena**.
hope-1sg that ate J anything
 - e. *Nomizo pos o Yiannis efage **kanena**.
think-1sg that J ate anything

⁶ Some Dutch speakers find (5-c) and (5-e) degraded. What is crucial here is that some speakers accept them, so these environments are relevant for NPI licensing.

The emerging generalization is the following: whenever an NPI like *ook maar iets* or *kanena* is licensed, speaker oriented adverbs are degraded. In other words, SOA are PPIs.

3 Approaches to polarity items

3.1 Monotonicity and veridicality

NPIs like *ook maar iets* and *kanena* are usually classified as ‘weak’ NPIs. These are licensed in all downwards entailing (DE) contexts, as well as certain other contexts, including epistemic *might*. NPIs that are only licensed in a subset of DE environments (viz. antiadditive environments) are called ‘strong’ NPIs. A function f is antiadditive iff

$$f(a \vee b) = f(a) \wedge f(b).$$

It is DE iff it is order reversing, i.e. whenever $a \sqsubseteq b$, it holds that $f(b) \sqsubseteq f(a)$, where “ \sqsubseteq ” denotes semantic strength. It is *Antimorphic* (AMo) iff it is both AA, and *antimultiplicative*. A function f is antimultiplicative iff

$$f(a \wedge b) = f(a) \vee f(b).$$

Examples of AA operators include *nobody*, *not*, *never*. Classical negation is antimorphic, i.e. it is both AA and antimultiplicative.⁷ DE operators that are not AA include *less than half of the N* and *rarely*.

Giannakidou (1997); Zwarts (1995) argue that there is a well defined class of semantic environments, *non-veridical* (NV), in which weak NPIs are licensed. Informally, a (propositional) operator Op is *veridical* if $\text{Op}(\varphi)$ entails φ for any φ ; *non-veridical* if $\text{Op}(\varphi)$ does not entail φ , and *anti-veridical* if $\text{Op}(\varphi)$ entails $\neg\varphi$. I give a more formal definition below. In general, we have that all antiadditive environments are DE, and that all DE environments are NV, although the converse is not generally true (cf. van der Wouden (1997); Zwarts (1998); Bernardi (2002) for discussion and formal proofs):

$$AMo \subseteq AA \subseteq DE \subseteq NV$$

This is important, because it shows us that, whenever some expression is (anti-)licensed in NV environments, it is (anti-)licensed in DE, AA and AMo

⁷ Antiadditivity and antimultiplicativity combine to yield the de Morgan Laws for classical negation. See van der Wouden (1997) for discussion of these properties and their relevance for different classes of polarity items.

environments as well. Thus, an analysis in terms of NV-ness can incorporate results from the analyses in terms of DE-ness.

Formally, veridicality can be defined as follows (Bernardi, 2002)⁸:

Definition 3.1 *Let f be a boolean function with a boolean argument,*

- (1) $f \in D_t^{D_t}$. f is said to be
 - (a) veridical iff $\llbracket f(x) \rrbracket = 1 \models \llbracket x \rrbracket = 1$
 - (b) non-veridical iff $\llbracket f(x) \rrbracket = 1 \not\models \llbracket x \rrbracket = 1$
 - (c) anti-veridical iff $\llbracket f(x) \rrbracket = 1 \models \llbracket x \rrbracket = 0$
- (2) $f \in D_t^{D_t^{(a,t)}}$ (i.e. f is a generalized quantifier). f is said to be
 - (a) veridical iff $\llbracket f(x) \rrbracket = 1 \models \exists y \in D_a(\llbracket x(y) \rrbracket = 1)$
 - (b) non-veridical iff $\llbracket f(x) \rrbracket = 1 \not\models \exists y \in D_a(\llbracket x(y) \rrbracket = 1)$
 - (c) anti-veridical iff $\llbracket f(x) \rrbracket = 1 \models \nexists y \in D_a(\llbracket x(y) \rrbracket = 1)$

Some examples are in order. *Possibly* is non-veridical because *possibly*(φ) does not entail (“ \models ”) φ . *Obviously* is veridical, because *obviously*(φ) $\models \varphi$. *I hope that* $_$ is non-veridical, because *I hope that* $\varphi \not\models \varphi$. The same appears to hold for *I think that* $_$, but Giannakidou argues that whenever it is true of John that he thinks that φ , then φ must be true in his epistemic model. This does not hold for *hope*. In this sense, then, *think* can be said to be *weakly veridical*. Weak veridicality is thus veridicality relativized to epistemic states.

Giannakidou’s motivation for using the notion is, of course, that some NPIs are apparently licensed outside of DE contexts. One example is adversative predicates (7). These are argued by Linebarger (1987) not to be DE.

- (7) a. I’m surprised he has any potatoes.
- b. I’m glad he has any potatoes.

However, Kadmon and Landman (1993) argue rather convincingly that these environments are, in fact, DE, once a certain contextual perspective is fixed and we limit ourselves to cases where their factive presupposition is satisfied. These authors argue along similar lines for antecedents of conditionals, which have also been argued not to be DE (Linebarger (1987)).

Another case for NVness are questions which cannot straightforwardly be said to be DE. In van Rooy (2002) it is shown that questions are, in fact, DE in their subject position under a Groenendijk and Stokhof (1984) style semantics. He gives an analysis of polarity items in questions extending and refining the results of Kadmon and Landman (1993); Krifka (1995), without appealing

⁸ Bernardi (2002) also generalizes the definition to n -ary boolean functions. I omit this here, as it will not be relevant to our discussion.

directly to NVness or DEness, however. It appears that the remaining cases for NVness are existential modals (e.g. *might*, cf. (8a)) imperatives (8b) and generics (8c). Generics was actually given a treatment in Kadmon and Landman (1993), so one might suspect that the other two can be dealt with as well.

- (8)
- a. John might have bought anything.
 - b. Bring anything!
 - c. John buys anything he finds.

There are also some problems for the NV-based approach. One obvious one is that languages differ with respect to which NV operators they allow to license weak NPIs. For instance, in Greek, *isos* ‘possibly’ licenses *kanena*, but the corresponding adverbs do not license Dutch *ook maar iets* or English *any*.

- (9)
- | | | | | | |
|-------|----|-----------|------|--------------|-----------------|
| Greek | a. | O Yiannis | isos | efage | kanena . |
| | | J | | possibly ate | anything |
- b. * John possibly ate anything.
 - c. * Perhaps John ate anything.

Similarly, we have seen that the Dutch and Greek verbs for *hope* license *ook maar iets* and *kanena*. According to our informants, English *hope* does not license *any*, as seen in (10).

- (10) *I hope John has any potatoes.

Another problem is that English *any* and Dutch *ook maar iets*, although they are licensed in other NV environments, they are actually reported to be degraded under a purely DE operator, i.e. one which is not AA. This result is slightly alarming, because the NPIs in question are licensed in other (upwards entailing) NV environments, as well as in AA environments, and, as we have seen, $AA \subseteq DE \subseteq NV$.

- (11)
- a. ? Less than three students have eaten anything.
 - b. ? Minder dan drie studenten hebben ook maar iets
less than three students have anything
gegeten.
eaten

This does not show that the class of licensors for these items cannot be characterized as (a subset of) NV, but it does seem to show that something more needs to be said.

3.1.1 Speaker orientation, NV and DE

Before moving on I would like to point out that the NV-based approach gives us some predictions for adverb ordering as it stands. The idea is this. Suppose that we rephrase the generalization made in the previous section in the following way:

- (12) SOA are excluded from X environments.
where X ranges over AMo, AA, DE and NV.

Let us now look at Cinque’s universal hierarchy with adverbs in their respective specifiers Cinque (1999)

- [illegible]

Out of the adverbs in (13), the ones in (14a) are NV. Hence, if this is the relevant property that SOA are “allergic to”, we expect them not to be able to precede and outscope SOA. To the NV class, I would like to add the ones in (14b).

- (14) a. allegedly, probably, perhaps, possibly, usually, no longer
b. hardly, never, rarely, not

Let us begin with some relatively clear contrasts. We expect, for example, that there should be a contrast between *often* and *rarely* with respect to their ability to outscope SOA. (15a) is an example found on the internet. It seems to contrast minimally with (15b) which should be a near semantic equivalent.⁹ We expect to find a similar contrast between *always* and *never*. (16a), also

⁹ I did find some examples of the string *rarely possibly* but these seem to involve misspellings. Here is one example:

(i) The receiver requires a line of sight to the satellites that is relatively unobstructed by foliage and buildings, and this is rarely possible on such a compact campus.

from the internet, is taken from an advertizement for an internet game. Again there is a fairly sharp contrast with (16b), which is what we expect.¹⁰

- (15) a. His retaliations killed or endangered innocents and **often possibly** had little effect in locating terrorists.
 b. ?? His retaliations killed or endangered innocents and **rarely possibly** had an effect in locating terrorists.
- (16) a. This is a fun, free game where you're **always possibly** a click away from winning \$1000!
 b. ?? This is a fun, free game where you're **never possibly** further than a click away from winning \$1000!

In (16a), one could in principle argue either that *possibly* directly modifies the noun phrase *one click away...*, or that *always* directly modifies *possibly* or both. This is rendered implausible by the following facts. The translation of (16a) into Norwegian is equally grammatical (17a), and in (17a), *muligens* 'possibly' is separated from the noun phrase by the copula. Furthermore, the subject of the clause can intervene between the two adverbs (17b). For (17b), in turn, one could object that *alltid* 'always' could directly modify the subject, but this is refuted by the ungrammaticality of (17c), where this supposed constituent is moved to the V2-initial position. One could be tempted to say that the position of the subject in (17b) is due to some PF-reordering mechanism. But this cannot be the case (at least on available ideas of what PF-reordering could do), since the subject takes surface scope with respect to the adverbs: the sentence means that it is always the case that, for at least one player, it is possible that that player is one click away from the prize. Finally, there is a contrast, also in Norwegian between (17a) and (17d), where *aldri* 'never' is substituted for *alltid* 'always', along with some other changes to make the example more pragmatically plausible.

- (17) a. Dette er et morsomt, gratis spill hvor spillerne **alltid**
 this is a fun free game where the-players always
muligens er et klikk fra å vinne \$1000!
 possibly are one click from to win \$1000
- b. Dette er et morsomt, gratis spill hvor **alltid** [en av
 this is a fun free game where always one of
 spillerne] **muligens** er et klikk fra å vinne \$1000!
 the-players possibly is one click from to win \$1000

¹⁰ I have deliberately tried to make the (b)-examples roughly synonymous with the (a) examples. This is to prevent a contrastive interpretation of the (b) examples: Contrastivity is known to suppress (positive) polarity effects (on this, see a.o. Szabolcsi (2002)).

- c. * **Alltid en av spillerne** er muligens et klikk fra å
always one of the-players is possibly one click from to
vinne \$1000!
win \$1000
- d. ?? Dette er et morsomt, gratis spill hvor spillerne **aldri**
this is a fun free game where the-players never
muligens er lenger enn et klikk fra å vinne \$1000!
possibly are further than one click from to win \$1000

Another contrast of the same kind holds between the pair *already* and *not yet*¹¹. Consider (18a) which contrasts with (19a).

- (18) a. In Wall Street, Enron was **already allegedly** going bankrupt.
- b. In Wall Street, Enron was **allegedly already** going bankrupt.
- (19) a. ?? (In Wall Street,) Enron was **not yet allegedly** going bankrupt.
- b. In Wall Street, Enron was **allegedly not yet** going bankrupt.

(20a), from the internet, does not contrast sharply with (20b); my informants find the latter example quite acceptable. This requires an explanation.

- (20) a. Although Beckham's absence will be felt England are **still probably** the best national team in Europe.
- b. (Beckham's absence will be felt and) England are **no longer probably** the best national team in Europe.

Consider what is the semantic contribution of an adverb like *no longer*. In general, *no-longer*(*p*) is truth-conditionally equivalent to $\neg p$, with the added presupposition that *p* was true at some previous time.¹² This presupposition makes it virtually impossible to exclude a contrastive reading of the modifiee. Contrastivity, as we have seen, quite generally interferes with polarity effects (Szabolcsi, 2002). The English word *somewhat* is a PPI, so (21a) is odd unless it is read with heavy stress on *does't*, thus contrasting it with the corresponding positive assertion. Finally, also *somewhat* can appear under *no longer*, again

¹¹ See Löbner (1999) for arguments that *not yet* is the negation of *already*, and that *still* is the dual of *already*, whereas *no longer* is the negation of *still*. *Still* and *already* are both PPIs, but they are allowed in a superset of the environments allowing for SOA. I return to this point shortly.

¹² *no longer*, *still*, *already*, *not yet* also come with aspectual requirements pertaining to the modified predicate (Löbner, 1999).

probably due to the contrastivity inherent to the presuppositional content of the adverb. In this way, the fact that (20b) is acceptable is actually expected on the view that *probably* is a PPI. (19a) also improves considerably if interpreted as an emphatic denial of (18a).

- (21) a. ?? Stanley doesn't like it somewhat.
- b. Stanley no longer likes it somewhat.

If examples like (15a, 16a, 18a, 20a) are not very frequent, there are probably pragmatic reasons for this. In any case, we are not trying to account for the statistical distribution of expressions, and I think it is fair to say that these examples are grammatical.

Two SOA in the same sentence rarely give good results. Consider, for instance (22).

- (22) a. ?? Maybe Stanley probably ate his wheaties.
- b. ?? Probably, Stanley possibly ate his wheaties.

We might take this to indicate that these adverbs are excluded from NV environments, rather than, say, merely DE environments. However, I do not think the oddness of these sentences is due to the status of the adverbs as PPIs. Rather, I think it is due to the fact that they are all epistemic adverbs of the same sort, and one cannot epistemically modify the same sentence twice, if the modifiers are of the same epistemic kind. One reason to think that this is true is the following. *Necessarily* is also epistemic, but it certainly isn't a PPI as it occurs felicitously under negation (23a). Nevertheless, it is bad under the scope of *possibly* (23b).

- (23) a. Stanley didn't necessarily eat his wheaties.
- b. ?? Possibly, Stanley necessarily ate his wheaties.

Another reason not to say that SOA are excluded from NV environments is that *allegedly* which is NV, seems to be able to outscope *probably* and *possibly* (24).

- (24) **Allegedly**, Enron was **probably/possibly** going bankrupt.

Thus it does not seem to be the case that *probably/possibly* is excluded from NV environments. Speech-act adverbs like (*briefly*) might seem to be excluded from all NV environments, since these are degraded under NV epistemic adverbs. For example, (25b) does not have a reading according to which "it is possible that I am brief in saying that S ate his wheaties."

- (25) a. **Briefly**, Stanley **possibly** ate the wheaties.
 b. ?? **Possibly**, Stanley **briefly** ate the wheaties.

This might also be treated as a syntactic binding effect. Thus, *briefly* is essentially a manner adverb containing a variable which can be syntactically bound. If it occurs “too high” in the clause for the subject to bind it, and thus to get a “subject” oriented or standard “manner” reading, it becomes speaker oriented by default.¹³ This is supported by the fact that even *frankly* becomes subject-oriented in certain cases. (26a) can only mean that Stanley is frank, not the speaker. (26b) is apparently ambiguous between the two readings, while (26c) seems to prefer the speaker oriented reading.

- (26) a. Maybe Stanley frankly doesn’t like fish cakes.
 b. Stanley frankly doesn’t like fish cakes.
 c. Frankly, Stanley doesn’t like fish cakes.

Cinque (this volume) points out that (27a) is degraded, while if one of the adverbs is realized as an adjective (27b), the example becomes good. Cinque does not pose this as a problem for the present account, but, since *surely* is not NV (or DE), it might be thought to be one. However, the two adverbs *surely* and *probably* do not seem to be able to cooccur in any order, as seen by comparing (27a) to (27b), so the relative order of the adverbs cannot be the source of the oddness of these examples.

- (27) a. ?? Stanley **surely probably** ate his wheaties.
 b. It is **surely probable** that Stanley ate his wheaties.
 c. ?? Stanley **probably surely** ate his wheaties.

This behavior might be related to the observation made in Cinque (1999) that cooccurrence of two adverbs ending in *-ly* is quite regularly degraded.¹⁴ I do not have an account for this phenomenon, but I would like to point out that it is equally problematic for all existing accounts of adverb ordering I am aware of. Examples like (27a) seem to improve if the two adverbs are not adjacent. In this case the relative ordering *does* seem to matter, since (28b) is significantly worse than (28a). But in this case, the adjectival version is also degraded, i.e. (28c-28d) are also odd, thus contrasting with (27b).

- (28) a. ? **Surely**, Stanley **probably** ate his wheaties.

¹³ For a very similar point of view concerning the manner/subject-oriented distinction, see (Ernst, 2000, (this volume)).

¹⁴ Similar remarks hold of German adverbs ending in *-weise*, Mainland Scandinavian ones ending in *-vis*, or Italian ones ending in *-mente*.

- b. ?? **Probably**, Stanley **surely** ate his wheaties.
- c. ?? It is **probably sure** that Stanley ate his wheaties.
- d. ?? I am **probably sure** that Stanley ate his wheaties.
- e. John is **probably sure** that Stanley ate his wheaties.

What seems to go wrong with the examples where *probably* outscopes *surely/sure* is that it makes little or no sense for the speaker to assert that s/he finds it likely that s/he is sure that Stanley ate his wheaties. In other words, (28e) is good because the adverb *probably* refers to the speakers epistemic state, whereas *sure* refers to *John's* epistemic state.

In sum, it seems that SOA are excluded from DE environments, but generally allowed in NV environments.

3.1.2 A note on ‘phase quantifiers’ and frequency adverbs

We have noted that *yet* (in its temporal use) and *any longer* are NPIs. This obviously limits their distribution. It has also been argued that *already* and *still* are PPIs. But they do not appear to have the same distribution of adverbs like *possibly*, so I will outline my answer to why this is so here. The answer is that the (positive) phase quantifiers¹⁵ are *not* excluded from DE environments, but merely from antiadditive (AA) environments. The latter, as we have seen, forms a subset of the former. *Few students* and *no students* are DE, but only the latter quantifier is AA. The definition of AA is repeated here.

$$f(a \vee b) = f(a) \wedge f(b).$$

Applying this to the quantifiers in question, we see that (29a) is equivalent to (29b), while (30a) is not equivalent to (30b).¹⁶

- (29) a. No students jumped or danced. \iff
- b. No students jumped and no students danced.

This equivalence goes through, because, if the set of students who danced is empty, and the set of students who jumped is empty, then the set of students who did one or the other must also be empty and vice versa.

- (30) a. Few students jumped or danced \niff

¹⁵ This term was introduced (as far as I know) by Löbner.

¹⁶ In the a-examples, only the readings where the quantifier outscopes the disjunction are relevant. For some languages this is not always possible, because in these languages, the disjunction is itself a PPI (Szabolcsi, 2001). As Szabolcsi points out, this does not appear to hold for English *or*, however.

- b. Few students jumped and few students danced.

This equivalence does not go through, because the fact that the set of students who jumped has low cardinality, and the set of students who danced has low cardinality does *not* entail that the union (disjunction) of these two sets has low cardinality. Now consider how *already/still* behave with respect to these quantifiers.

- (31) a. Few students are still/already here.
- b. ?? No students are still/already here.
- c. ?? The students aren't still/already here.

The adverbs are apparently degraded under AA (hence under AMo) operators, but good under merely DE ones. This parallels the fact that *still/already* are substantially better within the scope of *rarely* (which is DE) than they are within the scope of *never* (which is AA).

- (32) a. At 9 AM/PM, Stanley is rarely still/already tired.
- b. At 9 AM/PM, Stanley is never still/already tired.

Thus, phase quantifiers like *already/still* appear to be excluded from AA environments. This allows them more freedom than SOA, because, as we have seen $AA \subseteq DE$, and SOA are excluded from all DE environments. In particular, (17-18) demonstrates that *already/still* enjoy considerable freedom.

We have seen that upwards entailing (UE) frequency adverbs, like *always*¹⁷ and *often* have a very free distribution. *Always* is degraded when it outscopes negation. As discussed in [Beghelli and Stowell \(1997\)](#), this is a quite general property of universal quantifiers. Thus, (33a) is an odd sentence unless it is read with heavy stress on *everybody*, in which case the universal takes scope *below* the negation, not the other way around. No such oddness arises with quantifiers like *many people* (33b). An entirely parallel contrast obtains between (33c-33d). (33c) and

- (33) a. ?? Everybody didn't snore.
- b. Many people didn't snore.
- c. ?? John always didn't snore

¹⁷ *always*, being a universal quantifier, is DE in its restriction and UE in its scope. See [Beaver and Clark \(2002\)](#) for arguments that the restriction of *always* is given by the context, and not by the background, or unfocussed part of the clause. In other words, *always* is not focus-sensitive. Thus a clause (constituent) modified by *always* will be in a UE environment.

- d. John often didn't snore.

Apart from this, it seems that frequency adverbs have a very free distribution.

If we have 40 adverbs, there are $40 \times 40 = 1600$ ordered pairs that we would have to consider in order to exhaust their ordering possibilities, that is, if we limit ourselves to pairs of adverbs. If we do not limit ourselves to pairs, but exclude repetitions of the same adverb, the number is $40!$ ($= 1 \times 2 \times \dots \times 40 \approx 10^{48}$), a truly astronomical number. Hence, I am not going to test all possible orderings here. I hope to have made plausible the idea that limitations on adverb distribution can be treated, to a large extent, as a polarity phenomenon.

3.2 Why DE?

Summing up the findings in this section, we can refine our generalization about SOA as in (34). Why should this generalization hold? This question becomes rather acute when we combine it with the observation made in the introduction, that the adjectival counterparts of these adverbs are *not* excluded from DE environments. Compare (35a) to (35b):

- (34) SOA are excluded from DE environments.

- (35) a. *Jospin didn't possibly win.
b. It is not possible that Jospin won.

If we want to explain adverb distribution from the lexical semantics of the different adverbs, we now have to look for some independently motivated semantic difference between, say, *possible* and *possibly* which we can take to be the source of the contrast in (35). We will now see that there is a difference between these two expressions. In the rest of the paper, I will try to establish that this difference is indeed the culprit. Consider (36).

- (36) a. It's possible that Le Pen will win...
b. # Le Pen will possibly win...
c. # Perhaps Le Pen will win...
d. ... even though he certainly won't.

(36a) followed by (36d) (uttered by the same speaker) appears to make up a consistent statement. (36b)-(36c), on the other hand, cannot consistently be continued with (36d). Restricting our attention to *possible* and *possibly*,

we see that there must be a truth-conditional difference between the two. Impressionistically, the difference is this: (36a) simply states that there is some possibility, however remote and implausible, that Le Pen will win. (36b) does not permit such remoteness: It has it that there is a realistic chance for Le Pen to win. In other words, (36b) constitutes a *stronger statement* than (36a).

3.3 Widen up and strengthen

According to Kadmon and Landman (1993); Krifka (1995); Lahiri (1997); Chierchia (2001), the ungrammaticality of (37a) is a *pragmasemantic* phenomenon. I follow the implementation in Chierchia (2001) here. The idea is essentially that *any* is synonymous with the indefinite article *a* in the sense that they are both existential quantifiers. Thus the meaning of both (37a,37b) could be represented as (37c), where D subscripted to \exists is a contextual restriction, a quantificational domain (Westerståhl, 1988) and \exists represents an existential generalized quantifier, i.e. $\lambda X \lambda Y. X \cap Y \neq \emptyset$.

- (37) a. * John has any potato.
 b. John has a potato.
 c. $\exists_D(\text{potato})(\lambda x. \text{has}(j, x))$

The difference between the two is argued to be that (37a) involves expansion of D to a different quantificational domain $D' \supseteq D$. In other words, *any* invites us to consider more potatoes; it signals reduced acceptance of exceptions. Thus a more accurate representation of the ungrammatical sentence (37a) would be (38) where g is a function from sets to sets, such that $\forall X (X \subseteq g(X))$.

- (38) $\exists_{g(D)}(\text{potato})(\lambda x. \text{has}(j, x))$

So far, this does not explain why (37a) is ungrammatical. The reason for this is that domain expansion is subject to a strengthening condition: The result of domain expansion must entail the same proposition without domain expansion. The fact that there is a potato in some large set such that John has that potato does not entail that he has a potato in a subset of that large set. In order to implement strengthening compositionally, Chierchia defines an operator *Op* that universally closes the function variable g as follows:

Definition 3.2 *Let Δ be a contextually determined set of domain expansions, let φ be sentential constituent containing a free occurrence of g , and φ' be φ with all free occurrences of g removed. Then*

$$\begin{aligned} \mathbf{Op}(\varphi) &= \forall g \in \Delta[\varphi], \text{ if } \forall g \in \Delta[\varphi] \text{ entails } \varphi', \\ \text{else} &= \text{undefined.} \end{aligned}$$

Op does the following: It takes a formula φ containing a free occurrence of a domain expansion variable g and universally quantifies g just in case the result of this whole operation entails φ without domain expansion. Otherwise it is undefined. Consider the following examples.

(39) John doesn't have any potatoes.

Derivation 1

$$\begin{aligned} &\neg \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x)) \\ &\hspace{20em} \text{apply } \mathbf{Op} \blacktriangleright \\ &\mathbf{Op}(\neg \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))) \\ &\hspace{10em} \text{check strengthening} \blacktriangleright \\ &\forall g \in \Delta[\neg \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))] \models \\ &\quad \neg \exists_D(\text{potato})(\lambda x.\text{has}(j, x)) \\ &\hspace{20em} \blacktriangleright \\ &\forall g \in \Delta[\neg \exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))] \end{aligned}$$

(40) * John has any potatoes.

Derivation 2

$$\begin{aligned} &\exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x)) \\ &\hspace{20em} \text{apply } \mathbf{Op} \blacktriangleright \\ &\mathbf{Op}(\exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))) \\ &\hspace{10em} \text{check strengthening:} \blacktriangleright \\ &\forall g \in \Delta[\exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))] \not\models \\ &\quad \exists_D(\text{potato})(\lambda x.\text{has}(j, x)) \\ &\hspace{20em} \blacktriangleright \\ &\text{undefined.} \end{aligned}$$

In words, the bottom line of Derivation (1) says that, for any domain expansion that we are willing to entertain, there is no potato in the expanded domain that John has. This clearly entails that there is no potato in the (unexpanded) quantificational domain that John has, so strengthening is satisfied. In Derivation (2), strengthening is not satisfied, hence application of **Op** is undefined. In Derivation (2), we actually have that the “unexpanded” alternative entails the result of application of **Op**, i.e. we have that

$$\exists_D(\text{potato})(\lambda x.\text{has}(j, x)) \models \forall g \in \Delta[\exists_{g(D)}(\text{potato})(\lambda x.\text{has}(j, x))]$$

This will always happen if domain expansion applies in an upwards entailing (UE) environment. This is because of the fact that if a (small) set A has a non-empty intersection with another set B , then every superset (expansion) A' of A will also have a non-empty intersection with B . Under DE operators, this entailment relation by definition reverses, i.e. the unifying property of DE operators is precisely reversal of entailment relations. Hence, domain expansion is allowed exactly when it occurs in a DE environment. In [van Rooy \(2002\)](#) it is argued that one can derive the stipulative part of this proposal (i.e. that domain expansion requires *strengthening*) from very general and plausible assumptions about the pragmatics of statements, ultimately related to the Gricean notion of *relevance*. The latter notion, he derives from a decision theoretic notion of *utility*. For reasons of space, I cannot go into the details of this here.

This account explains why *any* only occurs in monotone decreasing (entailment reversing) contexts from its lexical semantics. See e.g. [Krifka \(1995\)](#) for a congenial analysis of several other polarity items. In the next section, I will show how the same kind of analysis can be extended to account for the behavior of *maybe* and *possibly*. The idea is essentially that these adverbs are associated with domain *shrinkage*. This explains their contrast with the adjective *possible* noted in the previous section. It also derives their status as weak PPIs.

4 *possibly*: Shrink and Strengthen

4.1 Modal bases

We have seen that there is an intuitive sense in which statements of the form *possibly*(φ) are stronger than statements of the form *it is possible that*(φ). By this we know that, if Γ is a DE operator, $\Gamma(\textit{possibly}(\varphi))$ should be a *weaker* statement than $\Gamma(\textit{it is possible that}(\varphi))$. If the use of *possibly* involves some operation that is subject to a strengthening condition, we can explain the distribution of this adverb in a manner entirely parallel to the explanation of the distribution of *any*. In order to achieve this, we need to look at what *possibly* means.

As a start, I take an information state (belief state) to be a set of possible worlds K , the set of worlds compatible with what we take to be true. If $K = W$, the set of all possible worlds, we are in a state of total ignorance: everything is possible. Upon learning a new proposition p , the new information state K' is given by $K \cap p$. Suppose p is a contradiction. Then $K \cap p = \emptyset$, i.e. the absurd information state. If $|K| = 1$, i.e. it only contains one world, we are

in a state of total information. All this is entirely standard. What is the result, given K , of learning that something is possible i.e. *possibly*(φ)? According to a standard view (Groenendijk et al., 1996), the result should be K if $K \cap \varphi \neq \emptyset$, otherwise it should be \emptyset . Groenendijk et al. (1996) themselves point out that this leads to a situation where one can never really learn anything new from possibility statements. Like Groenendijk et al. (1996), I choose to live with this problem for the purposes of this paper. I add that epistemic possibilities work on *modal bases* Kratzer (1977, 1991). In the default case, we take the modal base to be W . Thus, the meaning of *possible* becomes as in shown in (41):

$$(41) \quad \llbracket possible \rrbracket^K = \lambda p[p \cap W \cap K \neq \emptyset]$$

Given that K is always a subset of W , this is equivalent to $\lambda p[p \cap K \neq \emptyset]$. In order to derive the strengthening effect observed with *possibly*, I take this expression to come with domain shrinkage of the modal base. In (42), g is a variable over domain shrinks, i.e. for all $X, g(X) \subseteq X$.

$$(42) \quad \llbracket possibly \rrbracket^K = \lambda p \lambda g[p \cap g(W) \cap K \neq \emptyset]$$

Just as with the domain expansion associated with *any*, I assume that our function variable must be universally closed. To this effect, we can use Chierchia’s **Op**, but now with respect to domain shrinking functions, and with respect to domains consisting of possible worlds. Thus, *possibly* applied to *Le Pen will win* returns Derivation (3).

Derivation 3

$$\lambda g[\text{win} \cap g(W) \cap K \neq \emptyset]$$

apply **Op** ►:

$$\mathbf{Op}(\lambda g[\text{win} \cap g(W) \cap K \neq \emptyset])$$

check strengthening ►

$$\forall g \in \Delta[\text{win} \cap g(W) \cap K \neq \emptyset] \models [\text{win} \cap W \cap K \neq \emptyset]$$

►

$$\forall g \in \Delta[\text{win} \cap g(W) \cap K \neq \emptyset]$$

Δ now contains various ways in which we can constrain our modal base in various contexts. For instance, If we are discussing the French election, we can shrink W by intersecting it with plausible assumptions about the behavior of the French electorate. The bottom line of Derivation (3) says that the intersection of “win” with such a modal base has a non-empty intersection with K , our information state. Clearly, this can only happen if “win” has a non-empty intersection with K to begin with, so strengthening is satisfied.

This already derives the fact that *possibly* cannot occur in DE contexts. Before I move on to demonstrate this, however, I need to make sure that this setup actually derives the observed difference between *possible* and *possibly*, i.e. *possible*(win) is consistent with *certainly*(¬win) while *possibly*(win) is not. The standard semantics for *certainly* would be as in (43).

$$(43) \quad \llbracket \text{certainly} \rrbracket^K = \lambda p[K \subseteq p]$$

By this, *it is possible that Le Pen will win, even though he certainly won't* comes out as in (44), which is unfortunately inconsistent.

$$(44) \quad [\text{win} \cap K \neq \emptyset] \wedge [K \subseteq \neg \text{win}]$$

A solution that suggests itself is to make use of modal bases with this adverb as well. Thus, we could try to assume that *certainly* also comes equipped with a *W*-shrinking variable. There are two options, according to whether the modal base intersects with *K* or *p*. I consider them in turn.

$$(45) \quad \llbracket \text{certainly} \rrbracket^K = \lambda p \lambda g[g(W) \cap K \subseteq p]$$

Applying this to ¬win, we get the following.

Derivation 4

$$\lambda g[g(W) \cap K \subseteq \neg \text{win}]$$

apply **Op** ►

$$\mathbf{Op}(\lambda g[g(W) \cap \subseteq \neg \text{win}])$$

check strengthening ►

$$\forall g \in \Delta[g(W) \cap K \subseteq \neg \text{win}] \not\models [K \subseteq \neg \text{win}]$$

►

undefined.

The problem is that if we have that $A \subseteq B$, it does not follow from $A \subseteq C$ that $B \subseteq C$. In other words, our function variable *g* is already in a DE environment in Derivation (4). In fact, given our meaning for *certainly*, it should behave as a negative polarity item. But this is plainly wrong, given that (46) is quite impeccable.

$$(46) \quad \text{Chirac will certainly win.}$$

The other option would be to intersect our modal base with *p* rather than with *K*. This yields (47).

$$(47) \quad \llbracket \text{certainly} \rrbracket^K = \lambda p \lambda g [K \subseteq p \cap g(W)]$$

This would make (48a) a stronger statement than (48b). This does not correspond to our intuitions, I think. If anything, asserting (48b) is a stronger statement than the assertion of (48a). (48a) seems to indicate that there is some (however small) room for doubt.

- (48) a. Chirac has certainly won.
 b. Chirac has won.

Be that as it may, the meaning for *certainly* in (47) does not help us. (49), which is the meaning our example with *it is possible that...* would get now, is still inconsistent.

$$(49) \quad [\text{win} \cap K \neq \emptyset] \wedge \forall g \in \Delta [K \subseteq \neg \text{win} \cap g(W)]$$

4.2 Entrenched beliefs

The discussion in the previous subsection seems to indicate that we need something different. I follow [van Rooy \(2001\)](#) in assuming that an information state should be thought of as an ordering relation on the set of propositions $\varphi \subseteq W$. Intuitively, the ordering relation tells us how *plausible* a proposition is, given what we know. To do this, we need a way to compare the plausibility of worlds, given what we know. The definitions below are taken from [van Rooy \(2001\)](#), and ultimately from [Harper \(1976\)](#). They implement what is known as “Harper’s Principle” for determining a similarity relation among worlds:

- (50) Harper’s Principle (HP)
 Only propositions decided by K should count in determining comparative similarity relative to K .

Definition 4.1 Let $S = \wp W$, let $x, y \in W$ and K be a belief state. Then

$$S_y^x K = \{p \in S \mid (K \subseteq p \text{ or } K \subseteq \neg p) \text{ and} \\ ((x \in p \text{ and } y \notin p) \text{ or } (x \notin p \text{ and } y \in p))\}$$

$S_y^x K$ gives us the set of propositions which are decided by K and on whose truth value the two worlds x and y disagree. We can now say that a world x is closer to what we take to be true, or more plausible, than another world y if x disagrees less with what we know than y .

Definition 4.2 Let x, y be worlds. $x \preceq_K y$ iff for all $w \in K$:

$$|S_x^w K| \leq |S_y^w K|$$

The relation \preceq_K gives us a system of spheres of worlds of the kind proposed by Lewis and Stalnaker in the early seventies to account for conditionals and counterfactuals. We now need to know what set of worlds to associate with propositions. This is done by the following *selection* function. It tells us to only consider the most plausible worlds in p .

Definition 4.3 $C_K(p) = \{v \in p \mid \forall u \in p : v \preceq_K u\}$

Given that we can count the number of propositions in $S_y^x K$, we can devise a *quantitative* measure of plausibility.

Definition 4.4 Let $k(w/p)$ be the implausibility of a world w , after updating K with p . Let $k(p/q)$ be the implausibility of a proposition p after updating K with q .

$$\begin{aligned} \text{worlds:} \quad k(w/p) &= |S_u^w C_K(p)|, \text{ for any world in } C_K(p). \\ \text{propositions:} \quad k(q/p) &= \min\{k(w/p) \mid w \in q\} \end{aligned}$$

A proposition p is *accepted* in K iff $k(\neg p/\top) > 0$. p is more implausible than q w.r.t. K if $k(p/\top) > k(q/\top)$. Let $f(p) = k(\neg p/\top)$. Then f measures the level of *plausibility*, or *epistemic entrenchment* of p , given K . We follow [van Rooy \(2001\)](#) in defining an information state (belief state) K as just such an entrenchment relation on a set of possible worlds.

I will now use the entrenchment relation f to give a semantics for epistemically modal statements. In [van Rooy \(2001\)](#), he proposes that what he calls ‘evidential’ attitude reports, like *be certain that*, *be sure that*, *be convinced that* can be treated within the following schema, where f_w^a is the entrenchment function associated with a in w :

$$(51) \quad \llbracket a \alpha \text{ that } p \rrbracket^w = 1 \text{ only if } f_w^a(p) = \text{HIGH}$$

“HIGH” is a contextually determined number. I take it to be determined in the following way: One picks some strongly believed proposition p . $\text{HIGH} = f_w^a(p)$ and $\text{LOW} = f_w^a(\neg p)$. Given that $f_w^a(p) = k_w^a(\neg p/\top)$ by definition, this setup ensures that HIGH and LOW are inversely proportional. In order to implement our treatment of positive polarity, I make use of the following epistemic accessibility relation (\mathcal{E}):

$$(52) \quad \mathcal{E}(a, w) = \{w' \in W : k_w^a(w') \leq n\}$$

where n is either HIGH or LOW.

I write $\mathcal{E}^\uparrow(a, w)$ when n is to be understood as HIGH and $\mathcal{E}^\downarrow(a, w)$ when it is to be understood as LOW. Furthermore, I normally skip specification of a and w . For speaker oriented adverbs, a is always the speaker¹⁸ and w is always the world of evaluation. (52) either gives us the set of worlds whose implausibility is at most LOW (i.e. \mathcal{E}^\downarrow) or the set of worlds whose implausibility is at most HIGH (i.e. \mathcal{E}^\uparrow) Our meaning for *certainly* can now be represented as follows:

$$(53) \quad \llbracket \text{certainly} \rrbracket^K = \lambda p[p \cap \mathcal{E}^\downarrow \neq \emptyset]$$

Given that $C_K(p) = \{v \in p \mid \forall u \in p : v \preceq_K u\}$, i.e. the most plausible worlds in p , (53) actually ensures that if *certainly*(p) is true, then $f(p) \geq \text{HIGH}$. This seems to me the right result. It derives the fact that *certainly*(p) leaves some little room for doubt; it says that p is very plausible, given K . The adjective *possible* can now be given the following representation:

$$(54) \quad \llbracket \text{possible} \rrbracket^K = \lambda p[p \cap \mathcal{E}^\uparrow \neq \emptyset]$$

Again, this ensures that the plausibility of p is at least LOW. Our problematic example, *it is possible that Le Pen will win, even though he certainly won't*, now comes out as follows:

$$(55) \quad [\text{win} \cap \mathcal{E}^\uparrow \neq \emptyset] \wedge [\neg \text{win} \cap \mathcal{E}^\downarrow \neq \emptyset]$$

This is consistent, given the way we determine HIGH and LOW. More in particular, (55) is true just in case $f(\text{win}) = \text{LOW}$ and $f(\neg \text{win}) = \text{HIGH}$. As the reader may have anticipated, I will now derive the meaning of *possibly* from that of *possible* by applying domain shrinkage to \mathcal{E}^\uparrow . Such domain shrinkage is still thought to be contextually restricted, but we define a constraint on possible domain shrinks:.

Definition 4.5 *Let ∇ be the set of domain shrinks and let n be any natural number, s.t. $n < \text{HIGH}$, then in context c ,*

$$\nabla \subseteq \{g_n \mid g_n(\mathcal{E}^\uparrow) = \{w' \mid k_w^a(w') \leq n\}\}.$$

Applying a domain shrink takes us to a domain of worlds whose plausibilities are strictly greater than LOW. The meaning of *possibly* is the following, where, as before, g has to be universally closed by application of **Op**.

$$(56) \quad \llbracket \text{possibly} \rrbracket^K = \lambda p \lambda g[p \cap g(\mathcal{E}^\uparrow) \neq \emptyset]$$

¹⁸ Except, potentially when they occur embedded under an attitude verb like *believe*. Similarly for w . I ignore this complication here.

Op is defined as before, but now operating on $g \in \nabla$. The derivation of (57) is given in Derivation (5).

(57) Le Pen will possibly win.

Derivation 5

$$\begin{array}{rcl}
 & \lambda g[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] & \\
 & & \text{apply Op} \blacktriangleright \\
 & \mathbf{Op}(\lambda g[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset]) & \\
 & & \text{check strengthening} \blacktriangleright \\
 & \forall g \in \nabla[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] \models [\text{win} \cap \mathcal{E}^\uparrow \neq \emptyset] & \\
 & & \blacktriangleright \\
 & \forall g \in \nabla[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] &
 \end{array}$$

Strengthening is satisfied. The bottom line of Derivation (5) states (indirectly) that the plausibility of “win” must be strictly greater than LOW. This clearly entails that it must be greater than or equal to LOW.

We can now test whether (58a) is inconsistent, which is what we want. (58b) is the meaning assigned to this proposition.

- (58) a. $\#$ Le Pen will possibly win, even though he certainly won't.
b. $\forall g \in \nabla[\text{win} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] \wedge [\neg \text{win} \cap \mathcal{E}^\downarrow \neq \emptyset]$

This is clearly inconsistent. The first conjunct states that the plausibility of “win” is strictly greater than LOW whereas the second conjunct has it that the plausibility of “¬win” is greater than or equal to HIGH. But for any p we have that if $f(p) > \text{LOW}$, then $f(\neg p)$ must be $< \text{HIGH}$. Hence (58b) is inconsistent.

4.3 The status of possibly as a PPI

I will now show that our independently motivated semantics for *possibly* derives its distribution. We noted above that whenever strengthening is satisfied in some environment, application of a DE operator Γ to that environment reverses its entailment relations, so strengthening is no longer satisfied after application of Γ . I will go through some examples to assure you that this really works. Consider (59), where *possibly* occurs under negation with its derivation in (6).

(59) ?? Stanley didn't possibly eat his wheaties.

Derivation 6

$$\lambda g \neg [\text{eat} \cap g(\mathcal{E}^\uparrow) \neq \emptyset]$$

apply **Op** ►

$$\mathbf{Op}(\lambda g \neg [\text{eat} \cap g(\mathcal{E}^\uparrow) \neq \emptyset])$$

check strengthening ►

$$\forall g \in \nabla \neg [\text{eat} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] \not\models \neg [\text{eat} \cap \mathcal{E}^\uparrow \neq \emptyset]$$

►

undefined.

Recall that $g(\mathcal{E}^\uparrow) \subseteq \mathcal{E}^\uparrow$. The fact that “eat” has no intersection with the former therefore does not entail that it also has no intersection with the latter. Consider now the alternative Derivation 7 for (59), where the negation enters the derivation *after* application of **Op**.

Derivation 7

$$\lambda g [\text{eat} \cap g(\mathcal{E}^\uparrow) \neq \emptyset]$$

apply **Op** ►

$$\mathbf{Op}(\lambda g [\text{eat} \cap g(\mathcal{E}^\uparrow) \neq \emptyset])$$

MERGE negation ►

$$\neg \mathbf{Op}(\lambda g [\text{eat} \cap g(\mathcal{E}^\uparrow) \neq \emptyset])$$

check strengthening ►

$$\neg \forall g \in \nabla [\text{eat} \cap g(\mathcal{E}^\uparrow) \neq \emptyset] \not\models \neg [\text{eat} \cap \mathcal{E}^\uparrow \neq \emptyset]$$

►

undefined.

Strengthening fails here, too. The fact that not every way of constraining our domain gives us a domain that has an intersection with “eat” does not entail that the unconstrained domain has no intersection with “eat”. The reader might wonder why strengthening applies after the negation has been merged and not before. I clearly need this for the system to work. I follow [Chierchia \(2001\)](#) in assuming that strengthening is checked at the PHASE level of the syntactic derivation [Chomsky \(1999, 2001\)](#).¹⁹

Similar results obtain when *possibly* occurs within the scope of DE subjects like *few students* or DE adverbs like *rarely*. We give the Derivation (8) for *few students* here. We are only interested in the wide scope construal for *few students*, i.e. the reading where the subject outscopes *possibly*.

¹⁹ We might also say that application of **Op** is restricted to the PHASE level. In that case, Derivation (7) would not arise at all.

(60) ?? Few students possibly ate their wheaties.

Derivation 8

$\text{Op}(\lambda g[\text{few}(\text{student})(\lambda x[\text{eat}(x) \cap g(\mathcal{E}^\dagger) \neq \emptyset])])$

check strengthening ►

$\forall g \in \nabla[\text{few}(\text{student})(\lambda x[\text{eat}(x) \cap g(\mathcal{E}^\dagger) \neq \emptyset])] \not\models$
 $[\text{few}(\text{student})(\lambda x[\text{eat}(x) \cap \mathcal{E}^\dagger \neq \emptyset])]$

►

undefined.

If we take *perhaps* and *maybe* to be synonymous with *possibly*, these results extend to these adverbs as well.

4.3.1 Intervention and double licensing

In a recent paper, Szabolcsi (2002) discusses the behavior of the English PPI *some*. She analyzes some intriguing properties of this expression there. I will discuss two of them. In Linebarger (1987), she notes that (anti-)licensing of polarity items can be obstructed by certain kinds of interveners. Chierchia (2001) argues that this cannot be reduced to a syntactic Relativized Minimality effect, essentially, because one cannot identify the class of interveners in a natural way. Below is an example with the PPI *somewhat*.²⁰

- (61) a. * Stanley didn't like this somewhat.
 b. Stanley didn't always like this somewhat.

When *always* intervenes between the negation and *somewhat*, the sentence becomes good. Recall our internet game examples. We saw that *possibly* cannot follow *never* in such cases. Of course, it also cannot follow sentential negation (62a). The interesting thing is that, also in this case, an intervening *always* has a meliorating effect (62b), i.e. (62b) seems to be much better than (62a).

- (62) a. * This is a fun, free game where you're not possibly further than a click away from winning \$1000!
 b. ...where you're not always possibly a click away from winning \$1000!

In order to see how our setup can derive this, I will give an informal presentation of how Chierchia (2001) derives similar intervention effects for the

²⁰ This example is good on a so-called 'metalinguistic' or emphatic denial reading. See Szabolcsi (2002) for discussion.

NPI *any*. First, note that *always* seems to block the licensing of this element (Linebarger, 1987).

- (63) a. Stanley didn't like anything.
b. ?? Stanley didn't always like anything.

Chierchia notes that the class of interveners seem to share the following characteristic: While they do not (necessarily) introduce a scalar implicature in a UE environment, they *do* introduce one in a DE environment. *Always* participates in a (lexicalized) scale with *sometimes*. Thus, (64a) carries the scalar implicature that Stanley didn't always eat his wheaties. Since *always* is the strongest element in the relevant scale, (64b) does not introduce such a scalar implicature. However, under negation, the scale is reversed: (64c) *does* introduce the implicature that Stanley sometimes *did* eat his wheaties. I refer the reader to Krifka (1995); Chierchia (2001) for discussion of an algorithmic way to derive this.

- (64) a. Stanley sometimes ate his wheaties.
 ► Stanley didn't always eat his wheaties.
b. Stanley always ate his wheaties.
c. Stanley didn't always eat his wheaties.
 ► Stanley sometimes ate his wheaties.

Recall that Chierchia assumes that *any* comes with a domain expansion variable g which is universally closed by his **OP**, and that application of **Op** is subject to a strengthening condition. He now argues that if strengthening takes scalar implicatures into account²¹ we can derive the intervention effects as failure of strengthening. In other words, application of **Op** must yield a proposition which is stronger than the strongest meaning of the proposition without domain expansion. Suppose that we have two sets X, Y of potatoes, such that $X \subseteq Y$. Suppose that John did not always have a potato in Y , though he sometimes did. This does not entail that he did not always (though sometimes did) have a potato in X . Chierchia goes through a number of examples, including ones with numeral interveners (e.g. “*John didn't give two people anything”). For our setup, we would have to show that, while the negation reverses the strengthening relation, as we have shown, interveners would have the effect of blocking such reversal. This is not, in general, possible. In fact, most of the interveners are UE in their relevant argument. Suppose that

²¹ Chierchia argues that there are strong reasons to assume that calculation of scalar implicatures are calculated on subconstituents of the clause, and not on the end result of the derivation as commonly assumed. I refer the reader to Chierchia's work for discussion.

Γ is DE and that Θ is UE. Then, p is in a DE environment in both of the following examples:

- (65) a. $\Gamma(p)$
 b. $\Gamma(\Theta(p))$

Hence, addition of an implicature triggered by a (UE) intervener, should not have any rescuing effect on strengthening, as long as the implicature is added on both sides. If we only add it on one side, it could still not rescue a failure of strengthening which is what we would need. Suppose that, instead of being subject to strengthening, PPIs are subject to an *anti-weakening* constraint. In other words, application of a domain shrink must never lead to a *weaker* statement than the same statement without shrinkage. In order to implement this compositionally, we could define a universal closure operator Ω as follows:

Definition 4.6 *Let ∇ be the set of domain shrinks as defined above. φ' is obtained from φ by removing all free occurrences of g and adding scalar implicatures.*

$$\begin{aligned} \Omega(\varphi) &= \forall g \in \nabla[\varphi], \text{ if } \varphi' \not\models \forall g \in \nabla[\varphi] \\ \text{else} &= \text{undefined.} \end{aligned}$$

Though this will derive the intervention effects we have seen for PPIs while retaining the result that they cannot be (directly) outscoped by a DE operator, it does introduce an asymmetry between NPIs (subject to strengthening) and PPIs (subject to anti-weakening) which should be further motivated. One piece of support comes from the fact that that PPIs can be outscoped by non-monotone quantifiers like *exactly three N*. These couldn't satisfy strengthening for the simple reason that they are non-monotone. In other words, (66a) has a reading according to which there are exactly three people such that John gave them something. Similarly, (66b) allows the subject to outscope *possibly*, i.e. it has a reading according to which there are exactly three students such that they possibly ate their wheaties. This already shows that there is a problem with assuming strengthening for PPIs. Furthermore, the same quantifier gives rise to an intervention effect with *any* (66c).

- (66) a. John gave exactly three people something.
 b. Exactly three students possibly ate their wheaties.
 c. ?? John gave exactly three people anything.

We could not generalize the anti-weakening approach to NPIs, because, then, these should be licensed by non-monotone quantifiers, contrary to fact:

(67) * Exactly three students ate anything.

In a sense, the asymmetry between strengthening and anti-weakening reflects the fact that, where NPIs are licensed, PPIs are anti-licensed (Giannakidou, 1997). I go through the derivation of our example (62b), but before doing so, I must device a meaning for *always*. I follow, among others, Beaver and Clark (2002) in assuming that *always* is a universal quantifier over events. In other words, it denotes the subset relation over sets of these, and, crucially, the restriction is always contextually given. In our example, the contextual set could be all events of playing the relevant game. Then, (68a) comes out as (68b) and (68c) comes out as (68d).

- (68) a. You're always a click away from winning.
 b. $\{e \mid \text{play}(e)\} \subseteq \{e \mid \text{1click}(e)\}$
 c. You're always possibly a click away from winning.
 d. $\forall g \in \nabla[\{e \mid \text{play}(e)\} \subseteq \{e \mid \text{1click}(e) \cap g(\mathcal{E}^\uparrow) \neq \emptyset\}]$

$\{e \mid \text{1click}(e) \cap g(\mathcal{E}^\uparrow) \neq \emptyset\}$, where g is an arbitrary domain shrink, is a subset of $\{e \mid \text{1click}(e) \cap \mathcal{E}^\uparrow \neq \emptyset\}$. For ease of exposition, I will refer to the former set of events as A , and the latter as B . Furthermore, I will refer to the restrictor of *always* (in our case $\{e \mid \text{play}(e)\}$), as C . The proposition “You're not always possibly a click away from winning” thus comes out as $\lambda g[C \not\subseteq A]$ before application of Ω . Adding the scalar implicature to this, we get $\lambda g[C \not\subseteq A \wedge C \cap A \neq \emptyset]$. Applying Ω to this, we need to check that anti-weakening is satisfied. This is done by checking that the following non-entailment holds, recalling that $A \subseteq B$:

$$[C \not\subseteq B \wedge C \cap B \neq \emptyset] \not\models \forall g[C \not\subseteq A \wedge C \cap A \neq \emptyset]$$

Indeed it does. One can also see that it is the implicature introduced by the negated *always* which is responsible for the failure of the entailment. In this sense, then, *always* intervenes precisely because it introduces an implicature which blocks the weakening effect which the negation and other DE operators have on domain shrinkage. If *always* does not intervene, anti-weakening is violated. This is because, as we have seen, the following is true for any proposition p .

$$\forall g[p \cap g(\mathcal{E}^\uparrow) \neq \emptyset] \models [p \cap \mathcal{E}^\uparrow \neq \emptyset]$$

Since the negation is DE (even AMo), we therefore have the following, which is a violation of anti-weakening.

$$\neg[p \cap \mathcal{E}^\uparrow \neq \emptyset] \models \forall g\neg[p \cap g(\mathcal{E}^\uparrow) \neq \emptyset]$$

In other words, we have derived that certain interveners can rescue a PPI, just in case they are of the sort that introduce scalar implicatures under DE operators. Without such intervention, the PPI is illicit under DE operators.

Szabolcsi (2002) discusses another intriguing phenomenon about PPIs, namely what she calls ‘double licensing’. She attributes the observation to Jespersen. Consider (69).

- (69) a. * I doubt that Stanley liked it somewhat.
- b. * I think Stanley didn’t like it somewhat.
- c. I doubt that Stanley didn’t like it somewhat.

In this case, the adverbs behave differently from *somewhat*. (70c) does not seem to improve very much, compared to (70a, 70b). In (71), however, there is a detectable improvement.

- (70) a. ?? I doubt that Stanley possibly liked it.
- b. ?? I think Stanley didn’t possibly like it.
- c. ?? I doubt that Stanley didn’t possibly like it.

- (71) I don’t doubt that Stanley possibly liked it.

This bears on the assumption that strengthening/anti-weakening is checked at the PHASE level. For example, for (70a) to come out bad, it is essential that strengthening cannot be checked in the embedded clause. On the other hand, for (72a) to come out good, it seems to be crucial that it *can* be checked in the embedded clause. (72b), in turn seems to require the opposite. It might be that this could be made to follow from properties of the verbs in question. For instance, *think* is known to be a “neg-raising” verb, in the sense that the matrix negation in (72b) seems to be equivalent to negating the embedded clause. For now, I will leave this as an open problem.

- (72) a. Nobody thinks that Le Pen will possibly win.
- b. ?? John doesn’t think that Le Pen will possibly win.

In this section, we have seen that the behavior of *possibly* can be derived from its semantic difference with the proposition-embedding adjective *possible*. More in particular, we have seen that *possibly* yields stronger statements than *possible* and that this can be implemented in van Rooy’s (2001) framework, by assuming that the adverb comes with a variable over domain shrinking functions, mapping the epistemic accessibility relation \mathcal{E}^\uparrow to one of its subsets. Intuitively, what such domain shrinkage does is increasing the level of plausibility ascribed to the modified proposition. We have furthermore seen that if such domain shrinkage is subject to an anti-weakening constraint, and that such anti-weakening takes scalar implicatures into account, the limited distribution of the adverb follows. Hopefully, the anti-weakening constraint

on domain shrinkage can be derived from general pragmatic constraints in the spirit of [van Rooy \(2002\)](#) and the way this author derives the strengthening constraint for NPIs. If so we will have a general explanation of the distribution of SOA and PPIs more generally.

5 Prospects and consequences

I have concentrated on deriving the behavior of *possibly* here. The reason for this is mainly that a similar derivation for all the speaker oriented adverbs we discussed in section 2 would be too ambitious a task for the present paper. Nevertheless, I think the data discussed in section 2 and 3 allows us to be optimistic about the prospects of explaining the distribution of SOA, phase quantifiers and frequency adverbs in terms of their (independently motivated) scopal requirements. For manner adverbs, and aspectual adverbs like *completely* which generally cannot outscope any other adverbs, such an analysis might seem less adequate. They are not polarity items, for example. However, we noted that some manner adverbs *can* occur high, but then, they receive a subject-oriented or speaker-oriented reading. This can be treated as a scopal effect ([Ernst, 2000](#)), if we assume that they come with an implicit variable that must be c-commanded by its binder. For adverbs like *completely*, which do not behave like this, we would need to motivate an analysis according to which such adverbs do *not* come with an implicit variable, or that a subject/speaker-oriented reading would always lead to semantic or pragmatic degradedness. This would explain why they don't receive these readings, but in order to explain why they can't *occur* in high positions, we would need something more. It has been noted quite frequently that these adverbs interact with the argument structure and temporal constitution of the main VP ([Chomsky, 1965](#); [Alexiadou, 1997](#); [Cinque, 1999](#)). Thus, I speculate that, given a proper explanation of that, it might be made to follow that these adverbs cannot apply to a predicate which has already been modified by frequency adverbs, phase quantifiers, or SOA.

There are indications that the reasoning pursued in the present paper extends to some auxiliaries. Consider, for example the pair *might, can*. The former, but not the latter appears to be a PPI. Furthermore, while the former is restricted to an epistemic interpretation, the latter has a wider range of (root) modal uses. Even when *can* is interpreted epistemically it isn't a PPI.²² This fact seems to show that one can't account for the PPI status of *might* by assuming that there is a epistemic modality head M which c-commands a functional

²² [Giannakidou \(1997\)](#) also argues that Greek subjunctive morphology is an NPI. Certain kinds of Germanic (embedded) verb-second clauses appear to be PPIs, i.e. the ("bridge") predicates which allow embedded V2 are never DE.

head Neg hosting the negation. Under standard assumptions, epistemic interpretation of *can* should force it to move to M (at LF), if there is such a head, so we would be lead to expect that epistemic *can* should also be PPI, contrary to fact. Otherwise, one would be forced to say that only the epistemic modals which are, in fact, PPIs occupy M at LF, thus rendering the account vacuous.

- (73)
- a. The university might be in that direction.
 - b. The university might not be in that direction.
might > not. cf. *mightn't
 - c. ?? Nothing interesting might be in that direction.
 - d. ? Few interesting sites might be in that direction.
 - e. ?? I doubt that the university might be in that direction.
- (74)
- a. The university can be in that direction.
 - b. The university can't be in that direction
not > can
 - c. Nothing interesting can be in that direction.
 - d. Few interesting sites can be in that direction.
 - e. I doubt that the university can be in that direction.

5.1 Short Verb Movement

What are the consequences for a syntactic approach like that of Cinque (1999), if a similar account can be developed for all adverbs? I would like to separate this into two questions. The first one is whether or not adverbs are to be analyzed as specifiers of unique (functional) heads. I do not think the present paper bears on that question at all. Ultimately, it depends on how one should analyze purely syntactic phenomena like verb placement among adverbs in the Romance languages (Pollock, 1989; Cinque, 1999). I wish to remain agnostic about this question here.

The second question is this. Assuming that adverbs are specifiers of unique functional heads, how should we derive the ordering of these heads? The standard answer is that functional heads are ordered by syntactic selection. It seems to me that the present paper does bear on this question. In particular, if the present account is on the right track, it seems preferable to assume that functional heads are *not* ordered by selection. To make this point clear, I would like to point out some Norwegian facts that plainly cannot be accounted for by assigning positions to adverbs in a linear sequence of functional heads

(Nilsen, 2001).²³ There are triplets of adverbs that enter into non-linear ordering patterns. We have seen that the Norwegian adverb *muligens* ‘possibly’ has to precede sentential negation. *alltid* ‘always’, on the other hand has to follow it.

- (75) a. Jens hadde ikke alltid pusset tennene sine.
J had not always brushed the-teeth his
- b. *Jens hadde alltid ikke pusset tennene sine.
J had always not brushed the-teeth his

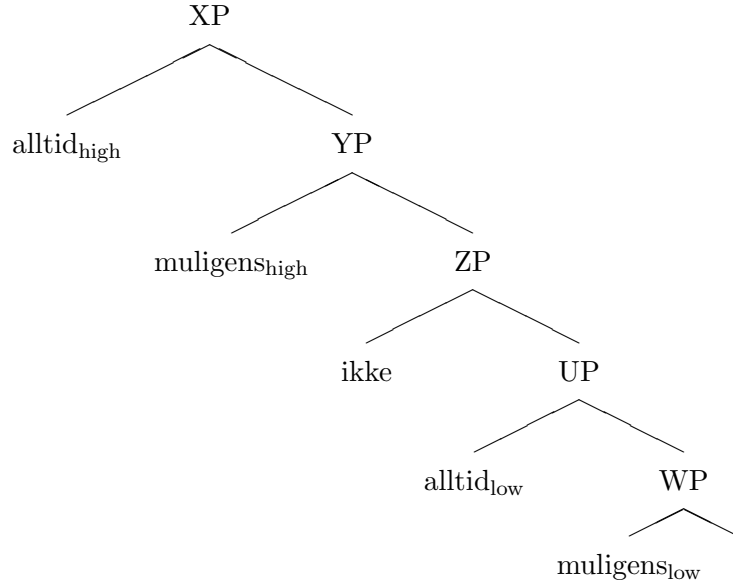
The ungrammaticality of (75) probably relates to the fact discussed in Beghelli and Stowell (1997), that universals generally don’t like to immediately outscope negation. By linearity, we now expect that *muligens* ‘possibly’ has to precede *alltid* ‘always’. But we have already seen that this is not true. We repeat the relevant example here.

- (76) Dette er et morsomt, gratis spill hvor spillerne alltid
this is a fun free game where the-players always
muligens er et klikk fra å vinne \$1000!
possibly are one click from to win \$1000

Thus the ordering of *muligens*, *ikke*, *alltid* isn’t linear. That is to say, we cannot account for their relative ordering by assigning positions in a linear sequence to them. Now assume that they are to be analyzed as specifiers of functional heads. If these heads are ordered by selection, we would be forced to assume multiple positions for the adverbs. To see this, suppose that we allow Norwegian *alltid* ‘always’ to occupy a “high” position, and let this be responsible for the examples where this adverb precedes *muligens* ‘possibly’. Given that, as we have seen, *muligens* can follow the negation (*ikke*) just in case *alltid* intervenes, we might consider adding an extra, low position for *muligens* as well. This results in the tree in (77)

- (77)

²³ Given binary branching, functional heads ordered by selection necessarily give rise to a linear sequence of specifiers.



This accommodates all the grammatical examples, but also many ungrammatical ones: in particular, this tree leads us to expect sentences like (75b) and ones where *muligens* immediately follows the negation to be possible, contrary to fact. Hence, it follows that some extra theory, such as the one advocated in the present paper, is needed to rule these out. But the extra theory seems to do quite well on its own, i.e. without extra positions and selectional sequences. In other words, ordering phenomena such as these don't provide arguments for selectional sequences.

An alternative, if one wants to say that adverbs are specifiers of unique functional heads, would be to say that these heads are freely ordered. This would amount to assuming some kind of IP-shell analysis. A particularly intriguing phenomenon to handle for such an approach is the following, discussed by (Cinque, 2000, this volume). In the Romance languages, there is considerable variation with respect to which adverbs a given verb form can precede or follow. For instance, Cinque shows that a French past participle can follow manner adverbs like *bien* 'well', whereas Italian past participles can't follow *bene* 'well'.

- (78) a. Il en a bien compris à peine la moitié.
 he of-it has well understood hardly the half
 b. Gianni ha (*bene) capito (bene).
 G has (*well) understood (well)

Furthermore, Cinque notes that the following generalization appears to be

true: If in a language L, the past participle (or another verb form) can follow (precede) a given adverb, then it can follow all other adverbs that can precede (follow) that adverb. In other words, verb-adverb ordering is transitive.²⁴ If true, I take this to show that whatever is the right account for adverb-adverb ordering, it should somehow carry over to verb-adverb ordering. In other words, if we account for adverb distribution in terms of the semantic scopal requirements of the individual adverbs, we should somehow be able to account for “short verb movement” phenomena as scopal in nature as well. Although working this out in detail is beyond the scope of the present paper, I would like to argue that it is not as implausible as it may seem at first blush.

First, if short verb movement is to be treated as a scopal phenomenon, it had better show some scopal effects. In fact, (Cinque, 1999, p49, n8) points out that, whereas (79a) entails that Gianni still has long hair, (79b) is compatible with him now having short hair.²⁵

- (79) a. Gianni ha sempre avuto i capelli lunghi.
 G has always had the long hairs
 b. Gianni ha avuto sempre i capelli lunghi
 G has had always the long hairs

Suppose that past participial morphology is treated as a “modifier” on a par with adverbs like *sempre*, but that it attracts the verb.²⁶ Suppose, furthermore, that it conveys anterior tense of the constituent it modifies.²⁷ Then the position of the verb in (79) would be a function of where the past participial morphology is merged. If it is merged above *sempre*, as would be the case in (79b) anterior tense applies to the constituent [always HAVE long hair], thus

²⁴ If our result that adverbs like *still*, *already*, *always*, *often*, etc. can precede “high” adverbs (SOA) carries over to Italian and other Romance varieties, this is problematic for Cinque’s transitivity claim, since he shows that the past participle generally cannot precede SOA in Romance, but they generally can precede the Romance translations of *still*, etc.

²⁵ My French informants report that there is no such semantic contrast in this language.

²⁶ such attraction may be made to follow on phonological grounds. The affix and the verbal stem are both “phonologically incomplete” in the sense that they do not make up phonological words on their own.

²⁷ Iatridou et al. (2002) argue that one should not, strictly speaking, analyze the perfect as “anterior”, i.e. that the Reichenbachian R_E interval is irrelevant for the perfect. See Borik (2002) for a sophisticated implementation of Reichenbachian ideas, which appears to take care of the objections raised in Iatridou et al. (2002) (although Borik herself does not relate her analysis to that of Iatridou et al.). The precise semantics of the perfect need not concern us in this informal discussion, although ultimately it is, of course, important.

allowing for the possibility that the state of affairs denoted by this constituent no longer holds. If it applies below *sempre*, the adverb would result in what has come to be known as the “universal” perfect (Dowty, 1979; Vlach, 1993; Iatridou et al., 2002), which does entail that the state of affairs still holds. In such a theory, one would have to derive the fact that perfect morphology cannot outscope certain adverbs by means of the scopal requirements of the perfect and the adverbs.

In fact, the claim that Romance participles can’t precede “high” adverbs may not even be true. An internet search (google) for exact strings like “dato fortunatamente” (‘given fortunately’), “avuto probabilmente” (‘have-PTC probably’), etc. returned several hundred hits among which were the following sentences.

- (80)
- a. Due incendi che non hanno **avuto fortunatamente**
Two fires that not have-3PL had fortunately
conseguenze rilevanti si sono sviluppati
consequences relevant SI are developed
 - b. le analisi hanno **dato fortunatamente** esito
the analyses have-3pl had fortunately output
negativo
negative
 - c. è stato probabilmente stampato a Roma
is-3SG been probably printed in Rome

There are two scenarios, depending on which status we assign to examples like (80). I discuss them in turn.

Suppose that these are, in fact, straightforwardly grammatical in Italian. Then past participial movement is allowed around very high adverbs. Then, given that the finite verb can also precede or follow high adverbs, we face the following problem which was pointed out by Bobaljik (1999) for argument ordering versus adverb ordering. Auxiliaries (and arguments) come in a specific order, so, in general, we have that, given a sequence of, say, three auxiliaries aux_1, aux_2, aux_3 and one adverb a of the relevant type, a can occupy any position in the sequence of auxiliaries, as long as the relative ordering of auxiliaries remains constant. Conversely, we have that given a sequence of adverb a_1, a_2, a_3 , and one auxiliary aux , aux can occupy any position in the sequence of adverbs, as long as the relative ordering of the adverbs remains the same. It follows that there can be no single selectional sequence accommodating the relative ordering of both adverbs and auxiliaries.

Suppose, on the other hand, that examples like (80) can be disregarded, for example because the adverb in question is in a so-called comma-intonation.

Then, we can maintain Cinque’s view that the participle can’t move around these adverbs.²⁸ We have seen above that phase quantifiers like *still* can precede “high” adverbs in English. Cinque shows that Italian participles *can* precede such adverbs. If our result that *still* can precede, say, *probably* carries over to Italian *ancora* and *probabilmente*, we have another failure of transitivity: PTC can precede *ancora*, *ancora* can precede *probabilmente*, but (in the current scenario) PTC can’t precede *probabilmente*. The following examples were found on the internet, suggesting that our result *does* carry over to Italian.

- (81) a. La risposta è ancora probabilmente no
 the answer is still probably no
 b. Gli americani sono ancora probabilmente in maggioranza,
 the Americans are still probably in majority
 ma non per molto.
 but not for long

A possible defense of functional sequences might be that one can accommodate this by assuming an extra, high position to be available for *ancora*. But adding extra positions for different elements significantly expands the expressive power of the system, and, leads to overgeneration in some cases. Secondly, although it may be made to work for the current problem, we have already seen that it will not do for the transitivity failure we noted for Norwegian.²⁹

A question raised by the suggestion that verb placement can be treated as a scopal phenomenon in the manner outlined here, raises the following question: Why doesn’t every language have short verb movement? In fact, I think there is evidence that even languages like Norwegian *do* have short verb movement. Consider (82a) which exhibits crossing scope dependencies between adverbs and verbs. In the sharply ungrammatical (82b) the adverbs immediately precede the constituent they modify.³⁰

- (82) Norwegian
 a. ...at det ikke lenger alltid helt kunne ha
 ...that it not any.longer always completely could have

²⁸ A relevant question is, of course, what the precise status of the “comma-intonation” is.

²⁹ The difference between the two cases is that, in the Norwegian case, we made use of the relation “x *must* precede y”, while the current problem concerns failure of transitivity of the relation “x *can* precede y”. This is why extra positions can rescue the latter, but not the former.

³⁰ the scope of *alltid* and *lenger* is not easy to determine, so one could also let the former adverb outscope *blitt* ‘been’ or the latter be outscoped by this verb.

- blitt ordnet.
 been fixed
- b. * ...at det ikke kunne lenger ha alltid blitt
 ...that it not could any.longer have always been
 helt ordnet.
 completely fixed

Crossing scope dependencies are entirely unexpected if absence of SVM is analyzed by letting the verbs remain in-situ. Therefore, (82a) may be taken as evidence for a rather elaborate system of short verb movements in Norwegian, inspired by [Koopman and Szabolcsi \(2000\)](#) on verbal complexes in West Germanic and Hungarian.³¹ In particular, this pattern would arise if scope reflects the order of merger, and each adverb attracts the projection of the closest verb, and each verb attracts the projection of the closest adverb:

- (83) [completely [fixed]]
 MOVE VP→
 [fixed [completely]]
 MERGE *been* →
 [been [fixed [completely]]]
 MOVE AdvP→
 [completely [been [fixed]]]
 MERGE *always*→
 [always [completely [been [fixed]]]]
 MOVE VP→
 [[been [fixed]] [always [completely]]]
 MERGE *have* →
 [have [[been [fixed]] [always [completely]]]]
 MOVE AdvP→
 [[always [completely]] [have [been [fixed]]]]
 MERGE *any.longer*→
 [any.longer [[always [completely]] [have [been [fixed]]]]]
 MOVE VP→
 [[have [been [fixed]]] [any.longer [always [completely]]]]
 MERGE *could*→
 [could [[have [been [fixed]]] [any.longer [always [completely]]]]]
 MOVE AdvP→
 [[any.longer [always [completely]]] [could [have [been [fixed]]]]]

³¹ [Bentzen \(2002\)](#) shows that SVM is visible in some varieties of Norwegian, including my own. For example, examples like (i) are perfectly grammatical.

(i) Jeg har ikke spist **ofte** tran.

I have not eaten often cod.liver.oil

However, in my variety, this is limited to a few adverbs, thus I could not substitute *alltid* ‘always’ for *ofte* ‘often’ in (i).

Such an approach could be extended to phenomena like “affix hopping” in English and other languages. In Nilsen (2002) such a “generalized verb raising” approach to SVM is developed in more detail.

5.2 Semantic selection

In Bartsch (1976); Ernst (2002), analyses of adverb distribution have been promoted which share with the present paper the assumption that it should be given a semantic treatment. The execution of the analyses is quite different, however. While these authors make use of rich semantic ontologies and selectional restrictions, no use has been made, in the present paper, of notions like “event”, “proposition”, “fact” etc. to account for the distribution of different adverbs. For example, Ernst (2002) assumes that there are different ontological entities like “events”, “propositions” and “facts” that adverbs can apply to, and that these ontological categories enter into the following system of type-conversion (his “FEO-calculus”)

$$\text{event} \Rightarrow \text{spec-event} \Rightarrow \text{proposition} \Rightarrow \text{fact} \Rightarrow \text{speech act}$$

In this setup, some adverbs are event modifiers (e.g. *completely*), some propositional modifiers (e.g. *not*) etc. The fact that *not* must precede (outscope) *completely* now follows from their (semantic) selectional requirements and the FEO-calculus. The structure of the account is thus remarkably similar to Cinque’s. The FEO-calculus corresponds to Cinque’s hierarchy of functional projections, and the ontological categories to Cinque’s heads. As far as I can see, ontological categories and his FEO-calculus are in no less need of explanation and motivation than Cinque’s sequence of heads. In practice, Ernst always assigns identity types to his adverbs, that is, they always map a constituent of semantic type X to another constituent of type X. If this is taken to be a general restriction on his system, he also predicts adverb ordering to be linear, which we have seen that it is not. If Ernst allows adverbs to have non-identity types, i.e. maps from type X to type Y, he no longer predicts linear ordering of adverbs. However, this move jeopardizes the intuitive “plausibility” his approach might have. For example, the Norwegian non-transitive triplet of adverbs *muligens* (‘possibly’), *ikke* (‘not’), *alltid* (‘always’) can be accommodated in the FEO-calculus if (and only if; see Nilsen (2001)) *muligens* takes a fact and returns a proposition, *ikke* takes an event and returns a fact, and *alltid* takes a proposition and returns an event. I refer the reader to Nilsen (2001, 2002) for discussion of this point.

I have assumed that adverbs like *always* are temporal and that, say *possibly* applies to propositions, but the fact that the two can occur in either order, with different scopal effects, suggests that this should not be brought to bear on the unavailability of certain adverb orderings. In fact, that would be a sur-

prising claim, which would require an explanation. For example, the fact that, say, *everybody* quantifies over individuals does not prevent it from outscoping (or being outscoped by) modal adverbs or verbs, negation and other non-individual operators. It would be surprising, then, if it were to turn out that the fact that some adverbs operate on times or events should prevent them from behaving similarly. As we have seen, they do not seem to be prevented in this way.³²

5.3 Summary

I have argued that SOA should be treated as PPIs, in the sense of being excluded from DE environments and that this suffices to derive their syntactic distribution. Positive phase quantifiers are also PPIs, but only excluded from AA environments. We have seen that the status of the adverb *possibly* as a PPI follows from its lexical semantics if we derive its meaning from the meaning of that of *possible* by means of an operation Ω on its epistemic accessibility relation \mathcal{E} , and making Ω subject to an anti-weakening constraint, sensitive to the scalar implicatures of the modified proposition. We also saw that intervention effects follow in a manner entirely parallel to Chierchia's (2001) derivation of such effects with NPIs.

References

- Alexiadou, A., 1997. Adverb placement a case study in antisymmetric syntax. John Benjamins, Amsterdam. 1, 5
- Alexiadou, A., 2001. On the status of adverb in a grammar without a lexicon., ms. University of Stuttgart. 1
- Bartsch, R., 1976. The Grammar of Adverbials. Elsevier, North-Holland. 5.2
- Beaver, D., Clark, B., 2002. Always and only – why not all focus sensitive operators are alike, ms. Stanford. 17, 21
- Beghelli, F., Stowell, T., 1997. Distributivity and negation. In: Szabolcsi, A. (Ed.), *Ways of Scope Taking*. Kluwer, Dordrecht, pp. 71–109. 17, 23
- Bellert, I., 1977. On semantic and distributional properties of sentential adverbs. *Linguistic Inquiry* 8, 337–51. 1, 2.1
- Bentzen, K., Feb 2002. Independent V-to-I movement without morphological clues, handout from paper presented at *Grammatik i Fokus*, Lund, 2002. 31
- Bernardi, R., 2002. Reasoning with polarity in categorial type logic, University of Utrecht. 7, 8

³² The question discussed here should not be confused with the distinct question why adverbs, as opposed to nominal quantifiers, tend to take surface scope.

- Bobaljik, J., 1999. Adverbs: the hierarchy paradox. *Glott International* 4 (9/10), 27–28. 27
- Borik, O., 2002. Aspect and reference time. Ph.D. thesis, Utrecht institute of Linguistics OTS. 27
- Chierchia, G., 2001. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface., ms. University of Milan – Biocca. 1, 3.3, 4.3, 4.3.1, 20, 20
- Chomsky, N., 1965. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Massachusetts. 5
- Chomsky, N., 1999. Derivation by phase, ms. MIT. 4.3
- Chomsky, N., 2001. Beyond explanatory adequacy, ms. MIT. 4.3
- Cinque, G., 1999. *Adverbs and Functional heads – A Crosslinguistic Perspective*. Oxford University Press, Oxford. 1, 2, 3.1.1, 13, 5, 5.1, 24
- Cinque, G., 2000. On Greenberg’s U20 and the Semitic DP, ms. University of Venice. 23
- Dowty, D., 1979. Word meaning and Montague grammar : the semantics of verbs and times in generative semantics. Reidel, Dordrecht. 27
- Dowty, D., 2000. The dual analysis of adjuncts/complements in categorial grammar. *ZAS Papers in Linguistics* 17. 1
- Ernst, T., 2000. Manners and events. In: Tenny, C., Pustejovsky, J. (Eds.), *Events as Grammatical Objects*. CSLI, Stanford, pp. 335–358. 13, 5
- Ernst, T., 2002. *The syntax of adjuncts*. Cambridge University Press, Cambridge. 1, 5.2
- Giannakidou, A., 1997. The landscape of polarity items. Ph.D. thesis, University of Groningen. 7, 21, 22
- Groenendijk, J., Stokhof, M., 1984. Studies on the semantics of questions and the pragmatics of answers. Ph.D. thesis, University of Amsterdam. 8
- Groenendijk, J., Stokhof, M., Veltman, F., 1996. Coreference and modality. In: Lappin, S. (Ed.), *Handbook of Contemporary Semantic Theory*. Blackwell, Oxford, pp. 179–216. 4.1
- Harper, W. L., 1976. Ramsey test conditionals and iterated belief change (a response to Stalnaker). In: Harper, W. L., Hooker, C. A. (Eds.), *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science*. Reidel, Dordrecht. 4.2
- Iatridou, S., Anagnostopoulou, E., Izvorski, R., 2002. Observations about the form and meaning of the perfect. In: Kenstowicz, M. (Ed.), *Ken Hale. A Life in Language*. MIT Press, Cambridge, Massachusetts, pp. 189–238. 27
- Julien, M., 2000. Syntactic heads and word formation : a study of verbal inflection. Ph.D. thesis, University of Tromsø. 1
- Kadmon, N., Landman, F., 1993. Any. *Linguistics and Philosophy* 16, 353–422. 2, 8, 3.3
- Koopman, H., Szabolcsi, A., 2000. *Verbal Complexes*. MIT Press, Cambridge, Massachusetts. 30
- Kratzer, A., 1977. What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy* 1, 337–355. 4.1

- Kratzer, A., 1991. Modality. In: von Stechow, A., Wunderlich, D. (Eds.), *Semantics: An international handbook of contemporary research*. de Gruyter, Berlin, pp. 639–650. 4.1
- Krifka, M., 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25, 209–257. 2, 8, 3.3, 3.3, 20
- Lahiri, U., 1997. Focus and negative polarity in Hindi. *Natural Language Semantics* 6, 57–123. 2, 3.3
- Linebarger, M.-C., 1987. Negative polarity and grammatical representation. *Linguistics and Philosophy* 10, 325–387. 8, 8, 4.3.1, 20
- Löbner, S., 1999. Why German *schon* and *noch* are still duals: A reply to van der Auwera. *Linguistics and Philosophy* 22, 45–107. 11, 12
- Nilsen, Ø., 2001. Adverb order in type logical grammar. In: van Rooy, R., Stokhof, M. (Eds.), *Proceedings of the Amsterdam Colloquium 2001*. Amsterdam, pp. 156–161.
URL <http://let.uu.nl/~oystein.nilsen/personal/ac.pdf> 5.1, 5.2
- Nilsen, Ø., 2002. Eliminating positions, forthcoming Ph.D. thesis, Utrecht University. 31, 5.2
- Nilsen, Ø., Vinokurova, N., 2000. Generalized verb raisers. In: Jang, Y. J., Kim, J.-S. (Eds.), *Proceedings of the 2000 International Workshop on Generative Grammar*. Hansung university, Seoul, pp. 167–176.
URL <http://let.uu.nl/~oystein.nilsen/personal/gvr.pdf> 1
- Pollock, J.-Y., 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20 (3), 365–424. 5.1
- Szabolcsi, A., 2001. Hungarian disjunctions and positive polarity., ms. NYU. 16
- Szabolcsi, A., 2002. Positive polarity–negative polarity., ms. NYU. 10, 12, 4.3.1, 20, 21
- van der Wouden, A., 1997. *Negative contexts : collocation, polarity and multiple negation*. Routledge, London. 1, 7
- van Rooy, R., 2001. *Attitudes and context change*, ms. of book ILLC, Amsterdam. 4.2, 4.2
- van Rooy, R., 2002. *Negative polarity items in questions*, ms. ILLC, Amsterdam. 5, 8, 3.3, 21
- Vlach, F., 1993. Temporal adverbials, tenses and the perfect. *Linguistics and Philosophy* 16, 231–283. 27
- Westerståhl, D., 1988. Quantifiers in formal and natural languages. In: Gabbay, D., Günthner, F. (Eds.), *Handbook of Philosophical Logic*. Kluwer, Dordrecht, pp. 1–131. 2, 3.3
- Zwarts, F., 1995. Nonveridical contexts. *Linguistic Analysis* 25, 286–312. 7
- Zwarts, F., 1998. Three types of polarity. In: Hamm, F., Hinrichs, E. (Eds.), *Plurality and Quantification*. Kluwer, Dordrecht, pp. 177–238. 7