

Deontic scope restrictions beyond polarity *

Igor Yanovich

Universität Tübingen, igor.yanovich@uni-tuebingen.de

Abstract

Deontic modals often have restricted scope with respect to clausemate negation. Recently attempts have been made to attribute all such restrictions to the polarity status of the respective modals ([Iatridou and Zeijlstra, 2013], [Homer, 2013]). I argue that once we examine the full range of cross-linguistic data, such a purely-syntactic account is hardly tenable. I suggest to complement polarity restrictions on deontics with semantic filters on scope. Conventions regarding such filters arise as speakers generalize from the range of actual uses that a pre-deontic construction used to have.

The aim of this paper is to show that the current syntactic accounts that view the scope restrictions of deontic modals as polarity restrictions cannot capture the whole range of empirical data, and need to be complemented with semantic-convention filters on scope configurations. After examining a wider range of data on deontic scope than it has been yet done in the formal literature, I argue that the observed diversity makes a purely syntactic account highly implausible. I therefore propose to complement syntactic polarity mechanisms with a semantic mechanism which may restrict the available range of scope construals for a given tense-aspect-mood form of the modal.

1 Deontics and clausemate negation: the state of the art

Deontic modals often have restricted scope with respect to clausemate negation. Possibility deontics (that is, permission modals) seem to universally scope under negation, see [van der Auwera, 2001, Sec. 5.6, 5.7], but necessity deontics (obligation modals) show a range of different behaviors. I have nothing to say about the scope restrictions of permission deontics, and set them aside for the purposes of this paper. From this point on, only obligation deontics will be discussed. [Iatridou and Zeijlstra, 2013] distinguish three types of such modals, given in 1:

- (1) I&Z's three polarity types of deontics:
 1. **PPI modals:** as other PPIs, need to be licensed by being in a non-downward-entailing (DE) context. **Examples:** *must, should*, Dutch *moeten*, Greek *prepi*
 2. **NPI modals:** as other NPIs, need to be licensed by being in a DE context
Examples: *need*, German *brauchen*, Dutch *hoeven*
 3. **Neutral modals:** are OK in upward-entailing contexts, but in the presence of negation scope under it. **Examples:** *have to*, German *müssen*

*The research reported here has benefitted from presentations at Ottawa University, UConn, MIT, Stanford and UCLA, and closely corresponds to a part of the Chapter 5 of my dissertation, [Yanovich, 2013]. The input of my advisers Kai von Fintel, Irene Heim and Sabine Iatridou has helped me to enormously improve the work reported, and is hereby most gratefully acknowledged. I also benefitted a lot from the discussions with Enoch Aboh, Jonathan Bobaljik, Cleo Condoravdi, Nathalie Dion, Olga Fischer, Sabine Iatridou, Magdalena Kaufmann, Stefan Kaufmann, Paul Kiparsky, Natalia Korotkova, Sven Lauer, Yael Sharvit, Sali Tagliamonte and Yakov Testelefs.

Iatridou and Zeijlstra employ two mechanisms to derive the empirically observed scope configurations from the surface structure. First, they argue that the modal appearing in the TP zone may reconstruct to a position within VP, and thus below negation. This is how I&Z derive the narrow scope for the NPI modal *need* in 2.

- (2) Mary needn't leave.
 = 'It is not that Mary needs to leave' OK $\neg > \square$
 ≠ 'Mary needs to not leave' * $\square > \neg$

Second, when a modal that occurs below negation in the surface syntax needs to scope above it, as in 3, I&Z posit covert, QR-like movement of the modal over the negation.

- (3) O Yanis dhen prepi na figi.
 John NEG \square -DEONTIC leave
 ≠ 'It is not that John has to leave' * $\neg > \square$
 = 'John has to not leave' OK $\square > \neg$

Thus reconstruction takes care of the cases when the modal needs to scope lower than it stays in the surface syntax, and covert QR-like movement applies when the modal needs to scope higher than its surface position. In the case of NPI or PPI modals, I&Z's analysis straightforwardly derives the facts: 1) due to the polarity requirements, only one of the two possible scope configurations allows the polarity-item-modal to be licensed; 2) if the only licensed scope configuration can be read off the surface structure, so be it; and if not, then either reconstruction or covert QR-like movement derives it. But for the third type of modals, which I&Z call "neutral", more must be said. Those modals are happy in upward-entailing environments, unlike NPIs. Yet when they occur in the same clause with negation, they obligatorily scope under it. So on the one hand, they are not polarity items, but on the other, they have fixed scope with respect to negation. English *have to* and German *müssen* are two examples:

- (4) Mary doesn't have to leave.
 = 'It is not that Mary has to leave' OK $\neg > \square$
 ≠ 'Mary has to not leave' * $\square > \neg$
- (5) Hans muss nicht abfahren.
 Hans \square -DEONTIC NEG leave
 = 'It is not that John has to leave' OK $\neg > \square$
 ≠ 'John has to not leave' * $\square > \neg$

In the surface structure, *have to* appears below negation, while *müssen* appears above it. I&Z take modals like *have to* to pose no problems in their system: "we need say nothing further for neutral modals that surface to the right of negation. They are simply interpreted where they appear in the overt syntax; that is, their syntactic and semantic scopes are identical", I&Z's p. 547. Now, that is not quite correct: as I&Z themselves discuss, modals have to have the option of undergoing covert QR-like movement for scope purposes; that option, other things being equal, should be available to *have to* in 4, and produce wide scope for the modal. But this is not what happens.

I&Z's account for non-NPI modals like German *müssen* or Spanish *tener que* which appear above negation in the surface structure is also problematic. For such modals, I&Z have to explain why their scope cannot be just read off the surface structure. To do that, I&Z introduce a principle forcing such modals to have narrow scope:

- (6) [Iatridou and Zeijlstra, 2013, (69b)]:
 Head-movement reconstructs unless reconstruction would result in a grammatical violation.

I&Z argue that given the principle in 6, the scopal properties of 5 follow: 1) as *müssen* is a non-polarity modal according to I&Z, there is no reason that would prevent it from scoping under the negation; 2) from 6, we derive that *müssen* obligatorily reconstructs.

There are two problems with this explanation. First, the principle in 6 only rules out one class of derivations that may lead to the $\square > \neg$ construal. There are at least two other types of derivations that need to be independently blocked: A) after reconstructing due to 6, a modal like *müssen* may undergo QR-like movement upwards across the negation; B) the modal may undergo QR-like movement from its surface position, ending up in a position above the negation; for modals like *müssen*, that would block head-movement reconstruction, rendering the principle in 6 irrelevant.

I do not claim that one absolutely cannot introduce constraints that would rule out all the derivations that lead to illicit scope construals. But for *have to*-type and *müssen*-type modals, we would have to introduce very different constraints that result in the same interpretational restrictions. For *have to* we need to prohibit QR-like movement from the base position of the modal. For *müssen* we need to prohibit QR-like movement from a raised position at T. At the same time, we cannot prohibit QR-like movement for modals in general, as it is needed to derive the observed scope configurations for other modals. There seems to be no principled way to carve out the relevant constraints on covert movement which would apply to *have to* and *müssen* without directly mentioning those modals—rather than some structural configuration—in the definition. In other words, the narrow scope of *have to* and *müssen* seems to be an idiosyncratic property of those modals, and not the consequence of how general syntactic mechanisms work.

The second problem with I&Z's explanation only reinforces that conclusion. In addition to non-polarity modals with fixed scope, there exist freely-scoping deontics. For example, I&Z themselves discuss French *devoir* which has free scope freely in simple present sentences like 7:

- (7) Il ne doit pas partir.
 He EXPL.NEG \square -DEONTIC NEG leave
 = 'It is not that John has to leave' OK $\neg > \square$
 = 'John has to not leave' OK $\square > \neg$

Modals like *devoir* do not fit into I&Z's classification in 1, but I&Z suggest that perhaps the principle 6 forcing obligatory reconstruction of head movement is language-specific, and does not exist in French. However, saying that French does not obey 6 is still not enough to account for the behavior of *devoir*, because French does not permit free scoping for *devoir* in all cases. As I&Z themselves note, in indicative perfectives, *devoir* is required to scope below the negation, 8.

- (8) Jean n' a pas du prendre l'autobus.
 Jean EXPL.NEG have NEG \square -DEONTIC take the bus
 = 'It is not that John had to take the bus' OK $\neg > \square$
 \neq 'John had to not take the bus' * $\square > \neg$

While in 7, *devoir* has moved over the interpretable negation *pas* (with higher *ne* being an expletive, omissible negative particle), in 8 the T position above *pas* is occupied by the auxiliary *avoir* that blocks head movement by *devoir*. I&Z hypothesize that the presence of the auxiliary that blocks head movement of the modal to T may be the reason why scope of *devoir* is restricted in 8,

Another important feature of the Russian system is that in addition to free-scope deontics, Russian also has both $\square > \neg$ and $\neg > \square$ fixed-scope necessity deontics. Free-scoping and restricted deontics have similar syntax in Russian, all occurring below negation in the surface structure. This further illustrates that scope restrictions are often specific to particular modals, rather than stem from the general properties of the syntactic system (contra I&Z, who propose the language-specific principle of obligatory reconstruction of head movement in order to derive the fixed scope of German *müssen* and Spanish *tener que*).

2.1 Russian free-scope deontics

Morphologically and syntactically, most deontics in Russian are predicative adjectives taking as arguments a Nominative or a Dative subject, and an infinitive clause. Predicative adjectives in Russian require the presence of copula *byť* ‘be’. In the present, the copula’s form is \emptyset , so it is not visible on the surface.

Russian modals *dolžná* (that can have the deontic, teleological and epistemic modal flavors) and *núžno* (need/deontic/teleological) have free scope with respect to clausemate negation:

- (12) Ona ne dolžna upominatj o svojom znakomstve s Anej.
 she NEG \square -DEONTIC mention about her acquaintance with Anya
 $\square > \neg$: ‘She mustn’t mention she’s acquainted with Anya.’
- (13) Maša objasnila, što Anja ne dolžna pisatj otčot.
 Masha explained that Anya NEG \square -DEONTIC write report
 $\neg > \square$: ‘Masha explained that Anya does not have to write a report.’
- (14) Ej ne nužno segodnja prinositj svoj obed:
 she.DAT NEG \square -DEONTIC today bring her lunch
 ‘She {mustn’t / doesn’t have to} bring her lunch today:’
- a. ... xolodiljnik slomalsja, i dekan poprosila poka ne prinositj svoju edu.
 fridge broke, and chair asked yet NEG bring one’s food
 $\square > \neg$: ‘... the fridge broke down, and the chair asked (everyone) to not bring their food until further notice.’
- b. ... na fakuljtete budet furšet.
 on department will.be catered.food
 $\neg > \square$: ‘... there will be catered food in the department.’

The same freedom of scoping is retained in the past tense:

- (15) Ona ne dolžna byla upominatj o svojom znakomstve s Anej.
 she NEG \square -DEONTIC was mention about her acquaintance with Anya
 $\square > \neg$: ‘She had to keep silent about her acquaintance with Anya.’
 $\neg > \square$: ‘She didn’t have to keep silent about her acquaintance with Anya.’
- (16) Ej ne nužno bylo upominatj o svojom znakomstve s Anej.
 she.DAT NEG \square -DEONTIC was mention about her acquaintance with Anya
 $\square > \neg$: ‘She had to keep silent about her acquaintance with Anya.’
 $\neg > \square$: ‘She didn’t have to keep silent about her acquaintance with Anya.’

Both *dolžna* and *nužno* appear below sentential negation in the surface syntax: negation in Russian always occupies a high position, cliticizing on the left to the highest finite element in its clause.² Thus the scope configuration $\neg > \square$ may be read off the surface structure, but the $\square > \neg$ interpretation has to be derived by covert QR-like movement of the modal, or its analogue.

2.2 Russian fixed-scope deontics

Not all Russian modals scope freely with respect to clausemate negation. In particular, *stoit* (an “advice” modal) always scopes over negation, while *objazana* (a deontic) always scopes below it:

- (17) Tebe ne soit begatj po utram.
 you.DAT NEG \square -ADVICE run on mornings
 $OK \square > \neg$: “You should not run in the mornings.”
 * $\neg > \square$: “It is not that you should run in the mornings.”
- (18) Ona ne objazana byla pisatj otčot o poezdke.
 she NEG \square -DEONTIC was write report about trip
 * $\square > \neg$: “She was required to not write a report about the trip.”
 $OK \neg > \square$: “She did not have to write a report about the trip.”

Objazana belongs to the same morphosyntactic category as *dolžna*, so the comparison between the two shows that it really depends on the lexical item alone which scope construals are allowed. *Stoit* is an impersonal verb, not a predicative adjective, but its surface position is under negation, just as for predicative adjectives, so there is no reason to think the morphosyntactic differences between *stoit* and *nužno* or *dolžna* should have any consequence with respect to their interaction with clausemate negation.

²The role of that “highest finite element” in 15 and 16 is taken up by the predicative adjective: on the surface, it looks as if the adjective has head-moved across the copula, in a pattern similar to the well-known “long head movement” pattern in South and West Slavic. It is still an open question what exact underlying structure corresponds to the linear order (*Neg Adj Aux* in different Slavic languages, the several analytical options being: (1) [Rivero, 1994] argues for a long head movement analysis, with the adjectival element (a participle or an actual adjective) appearing in C; (2) [Bošković, 1995] argues for incorporation analysis, in which the adjectival element merges into the copula head; (3) [Ackema and Čamdžić, 2003] arguing for base-generating the adjectival element in the higher position. This does not exhaust the analytical options, and some authors, esp. [Embick and Izvorski, 1997], caution against lumping together all instances of the order *Adj Aux*, within the same language or across different Slavic languages (cf. also [Borsley and Rivero, 1994] and [Broekhuis and Migdalski, 2003]).

Russian data, previously not described in connection to “long head movement”/“Adj movement across the copula” in other Slavic, should bear on that debate, though much future research is needed. For example, the fact that the Russian sentence in (i) is grammatical shows that either the incorporation analysis of Bošković does not work for Russian, or it does not work for BCS. Bošković argues that if there is a need for the incorporated adjectival element to move further up, it necessarily exorporates, stranding the copula below. So when Adj has an independent reason to move higher from the complex *Neg-Adj-Aux*, it strands *Neg-Aux* — as can be observed in BCS. But in the Russian (i), negation is not stranded, but taken along by the moving adjective. Thus either Russian orders *Adj Aux* do not feature Bošković-style incorporation, or Bošković is wrong about the excorporation requirement, which would destroy his account of the BCS data.

- (i) Ne nužno Maše bylo tuda ezditj.
 NEG \square -DEONTIC Masha.DAT was there go
 $\square > \neg$: ‘Masha should not have gone there.’
 $\neg > \square$: ‘It is not that Masha should have gone there.’

New Russian data above add further evidence that scope restrictions of modals are so diverse that they are likely to be associated directly with individual modals, or even forms thereof, rather than stem from major principles regulating the work of the syntactic component.

3 Semantic-convention filters on modal scope construals

In view of the evidence we discussed above, I argue for the following:

- (19) Modals may be PPIs (like *must*) or NPIs (like *need*), or they may be not polarity-sensitive at all (like English *have to*, French *devoir*, Russian *nužno*).
- (20) The syntactic component allows a non-polarity-sensitive modal to have either scope with respect to clausemate negation.
(After [Iatridou and Zeijlstra, 2013], we can take the mechanisms responsible for this to be reconstruction and scope-changing covert movement.)
- (21) Not every modal uses all the possibilities made available by the syntax.
Namely, there may be further constraints in place that rule out a particular scope configuration for a given modal word (English *have to*, Russian *objazana*), or for a modal word within a particular environment (French *devoir* in indicative vs. irrealis perfectives).

Thus polarity-sensitivity is not the only mechanism that may constrain how a modal scopes with respect to clausemate negation. In particular, the case of *devoir* suggests that scope restrictions may be tied to particular “constructions” rather than attached to lexical items. It is hardly possible to derive such construction-specific constraints using general-purpose syntactic mechanisms. But do we have any way to account for scope constraints not using syntactic mechanisms?

I argue that we do, and that fixed-scope constraints may be imposed by the lexical and construction-specific semantics and pragmatics of the language. We know independently that certain meanings may be indexed to particular constructions rather than follow from the compositional semantics of the lexical items alone. For example, *Can you pass me the salt?* conventionally conveys a request, while *Are you capable of passing me the salt?* is not (cf. [Horn and Bayer, 1984]). Even the latter sentence may give rise to the implicature of a request, but only the first does so conventionally, without requiring much pragmatic reasoning. There is no *a priori* reason why restrictions on relative scope could not be conventionalized, too.

A semantic convention imposing fixed scope would be learned in the same way speakers learn the lexical meanings of words. After language learners hear a word used a large number of times in a similar way, they abstract from those occurrences a semantic representation for the word. The semantics of a word gets generalized from individual instances in such a way as to be capable of explaining each of those. Statistically significant absence of positive evidence works as negative evidence in the creation of such conventions: we know that *rabbit* cannot denote a frog because we never heard people use *rabbit* that way. Finally, meanings may be associated not with individual words, but with larger chunks of structure, as the existence of idioms shows.

The acquisition of a semantic filter on scope construals would proceed similarly. For example, the learners would hear surface string such as \square *Neg*, and due to the existence of the constraint in the speech of competent speakers, that string would only be used in sentences conveying the *Neg* \square reading. With only a few examples of this sort encountered, the learners could have not noticed the pattern. But the more frequent the surface string \square *Neg* is, the more striking it becomes that it is only used to convey the *Neg* \square reading. As learners are sensitive to such statistical evidence,

interpreting it as a sign that something should be ruled out by the grammar, they acquire a scope constraint. If the constraint can be tied to the workings of a general syntactic mechanism (e.g., to the licensing of polarity items), then learners may acquire a syntactic constraint. But if the constraint seems to be idiosyncratically tied to a particular modal, or even to its combination with a particular tense-aspect-mood form, a semantic filter may become established in the grammars being acquired.

Once a semantic convention is established, it will perpetuate itself, other things being equal. The usage of all members of the linguistic community will be constrained by the convention, and new speakers will learn to conform to the same convention as they acquire language, unless there is pressure for language change. So the explanatory burden associated with positing a particular semantic convention restricting modal scope is to demonstrate how it got conventionalized: once it is established, the speakers will use the restriction until they have a good reason not to; it is the rise of the restriction that is not a trivial matter. For reasons of space I omit here case studies showing how we can trace a scope filter's rise; in [Yanovich, 2013, pp.203-226], I discuss the conventionalization of the wide scope of Russian *stoit* 'should' and the narrow scope of English *have to*.

References

- [Ackema and Čamdžić, 2003] Ackema, P. and Čamdžić, A. (2003). LF complex predicate formation: The case of participle fronting in Serbo-Croatian. In *UCL Working Papers in Linguistics*, volume 15, pages 131–175.
- [Borsley and Rivero, 1994] Borsley, R. D. and Rivero, M. L. (1994). Clitic auxiliaries and incorporation in Polish. *Natural Language and Linguistic Theory*, 12:373–422.
- [Bošković, 1995] Bošković, Željko. (1995). Participle movement and second position cliticization in Serbo-Croatian. *Lingua*, 96:245–266.
- [Broekhuis and Migdalski, 2003] Broekhuis, H. and Migdalski, K. (2003). Participle fronting in Bulgarian as XP-movement. *Linguistics in the Netherlands*, 20:1–12.
- [Embick and Izvorski, 1997] Embick, D. and Izvorski, R. (1997). Participle-auxiliary word orders in Slavic. In Brown, W., Kondrashova, N., and Zec, D., editors, *Formal Approaches to Slavic Linguistics: The Cornell meeting 1995*, pages 210–239, Ann Arbor, MI. Michigan Slavic Publications.
- [Homer, 2013] Homer, V. (2013). Neg-raising and positive polarity: The view from modals. To be published with *Semantics and Pragmatics*, pending minor revisions. Accessed at <http://tinyurl.com/d82c5cr> in January 2013.
- [Horn and Bayer, 1984] Horn, L. R. and Bayer, S. (1984). Short-circuited implicature: a negative contribution. *Linguistics and Philosophy*, 7:397–414.
- [Iatridou and Zeijlstra, 2013] Iatridou, S. and Zeijlstra, H. (2013). Negation, polarity and deontic modals. *Linguistic Inquiry*.
- [Rivero, 1994] Rivero, M. L. (1994). Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory*, 12(1):63–120.
- [van der Auwera, 2001] van der Auwera, J. (2001). On the typology of negative modals. In Hoeksema, J., Rullmann, H., Sánchez-Valencia, V., and van der Wouden, T., editors, *Perspectives on negation and polarity items*, pages 23–48. John Benjamins, Amsterdam.
- [Yanovich, 2013] Yanovich, I. (2013). *Four pieces for modality, context and usage*. PhD thesis, MIT, Cambridge, MA. Directed by Kai von Stechow, Sabine Iatridou (committee chairs) and Irene Heim.