Neg-raising through question-raising?

1 A Conjecture

Negation often has a positive side-effect. (1a), for example, denies that we will have pizza for dinner, but also conveys that we will have something else.

(1) a. We will not have [pizza]$_F$ for dinner.

b. $\rightsquigarrow$ We will have something other than pizza for dinner.

I want to suggest that the mechanism behind this effect may also explain the phenomenon of "neg-raising", exemplified in (2).

(2) a. Bill doesn't think that Sue is smart.

b. $\rightsquigarrow$ Bill thinks that Sue is not smart.

I'm going to focus on the following idea about how the inference in (1) comes about. When a hearer encounters (1a), he needs to reconstruct a question the utterance might address. The most plausible candidate is what we will have for dinner, which presupposes that we will have something for dinner. In the same way, I suggest, a hearer who encounters (2a) would normally assume that the utterance elaborates on what Bill thinks about Sue's intellectual abilities, which presupposes that Bill has an opinion. In section 5, I will briefly look at an alternative explanation based on a grammatical exhaustification operator.

Let me put a little more flesh on the proposed explanation. I assume, in the tradition of [Carlson 1983], [Roberts 1996], [Beaver and Clark 2008], and others, that the interpretation of declarative utterances involves the reconstruction of a question the utterance is assumed to address. A felicitous utterance must be *congruent* with this question. To a first approximation, an utterance is congruent with a *wh*-question if the question can be obtained by replacing focal constituents in the utterance with suitable *wh*-words, as in (3a) and (3b).

(3) a. What will we have for dinner?

b. We'll have [pizza]$_F$ for dinner.

c. We will not have [pizza]$_F$ for dinner.

Let's call this *direct* congruence. To explain the inferences in (1) and (2), we need to assume that there are also indirect forms of congruence. In particular, we need to assume that (3c) (= (1a)) is

also congruent with (3a). This is intuitively plausible: (3c) would be a felicitous response to (3a). In general, negation tends to preserve congruence: if S is congruent with Q, then $\neg S$ is congruent with Q.

Now suppose a hearer encounters (1a). He needs to figure out what question the speaker might want to address. A plausible candidate is (3a). Let's represent this question by the set of its complete (possible) answers:

(4) { x is what we'll have for dinner | x is a kind of food }

The question *presupposes* that we will have something for dinner, in the sense that all its elements entail that we will have something for dinner. A speaker who addresses a question can be assumed to presuppose that one of its elements is true. If some element of (4) is true, but the literal meaning of (1a) is also true, we must have something other than pizza for dinner. (Compare [Rooth 1996: 293-295], [Beaver and Clark 2008: 45-49] for similar explanations of (1).)

The inference is pragmatic: we would not judge (1a) to be false if there'll be nothing for dinner. On the other hand, the inference is hard to cancel, at least if *pizza* has a fall-rise contour:

(5) ? We will not have $[\text{pizza}]_{\text{f-r}}$ for dinner. We are fasting.

In some respects, the effect looks like a presupposition, but it doesn't project the way presuppositions normally do. For example, (6a) and (6b) don't presuppose that we'll see anyone in the park.

(6) a. We might see $[\text{Jane}]_{\text{F}}$ in the park.

b. If we'll see $[\text{Jane}]_{\text{F}}$ in the park, we can give her the keys.

After all, neither (6a) nor (6b) addresses the question who we'll see in the park.

All these properties are shared by NR-effects. Let's return to (2).

(2) a. Bill doesn't think that Sue is $[\text{smart}]_{\text{F}}$.

b. $\rightsquigarrow$ Bill thinks that Sue isn't smart.

As before, we assume that a hearer who encounters (2a) needs to reconstruct a question the utterance might address. A plausible candidate is what Bill think about Sue's intellectual abilities, which we might represent as (7).

(7) { Bill thinks Sue is x | $x \in$ { smart, not smart } }

(2a) is indirectly congruent with (7), just as (1a) is indirectly congruent with (4). (7) presupposes that Bill has an opinion about Sue's intellectual abilities: he either thinks that she is smart or that she is not smart. Since (2a) denies the first possibility, the hearer can infer that the second obtains.

The explanation seems to carry over to NR-adjacent phenomena like the "homogeneity presupposition" triggered by definite plurals, as in (8).

(8) a. It is not the case that the pigs are $[\text{in the barn}]_{\text{F}}$.

b. $\rightsquigarrow$ None of the pigs are in the barn.

The explanation would be that (8a) is assumed to address a question like (9) that presupposes that all the pigs are in the same location.

(9) { The pigs are in x | x is a location }

Since the literal meaning of (8a) rules out the barn location, one can infer that the pigs are (all) somewhere else.

Here's a schematic model. We want to explain why some utterances of the form $\neg P(\phi)$ systematically convey $P(\neg\phi)$. My conjecture is that the utterance is interpreted as addressing a question $\{ P(\xi_i) \}$, in which all elements compatible with $\neg P(\phi)$ entail $P(\neg\phi)$ (and there is at least one such element). The literal meaning $\neg P(\phi)$ together with the question's presupposition $\bigvee_i P(\xi_i)$ then entails $P(\neg\phi)$. If P is upward monotonic, as neg-raisers typically are, the above condition – that there is at least one element compatible with $\neg P(\phi)$ and all such elements entail $P(\neg\phi)$ – is satisfied whenever the alternatives ξ_i are mutually exclusive and include ϕ.[1]

I have been deliberately vague about the details. A full account would need to explain the role of the "addressed question", formalize the concepts of questions and congruence, specify rules for determining congruent questions, and so on. There are many ways to do this. I want to explore what an approach of this kind would have to look like if it is to explain NR effects. Some of the details don't matter. But some do. Let's turn to what I see as the greatest stumbling block for the explanation I have outlined.

2 Unspecific answers

Let's have another look at (8).

(8) a. It is not the case that the pigs are in the barn.

b. $\rightsquigarrow$ None of the pigs are in the barn.

I have suggested that the inference arises because (8a) is assumed to address a question like (9).

(9) { The pigs are in x | x is a location }

Given that the focus in (8a) lies on *in the barn*, this is in line with, for example, the "Focus Principle" of [Beaver and Clark 2008: 37].[2] But there is a problem.

Suppose the "locations" in (9) are all fairly specific: *in the barn*, *in the field*, *in the sty*, etc. (8a) is then predicted to convey that all the pigs are together at a specific place. But arguably it doesn't. There's nothing wrong with uttering (8a) if the pigs are spread between the field and the sty.

For another example of this problem, suppose (10a) addresses (11). We could then derive not only the NR inference to (10b), but also the more dubious inference to (10c).

(10) a. Bill doesn't think that Sue is from Italy.

1 I assume that a question contains at least two elements. Suppose $\{P(\xi_i)\}$ contains some element $P(\xi)$ besides $P(\phi)$. If the elements of $\{\xi_i\}$ are mutually exclusive, ξ entails $\neg\phi$. If P is downward monotonic, it follows that $P(\xi)$ entails $P(\neg\phi)$.

2 Following [Rooth 1992], Beaver and Clark assume that each sentence is associated with a set of "alternatives" that (loosely speaking) vary the denotation of focussed constituents. The alternatives to *The pigs are [in the barn]$_F$* would be (9). The Focus Principle says that a felicitous utterance must have a part whose alternatives contain all elements of the "Current Question". If the Current Question is (9), (8a) satisfies this condition, as the alternatives to its part *The*

b. $\rightsquigarrow$ Bill thinks that Sue is not from Italy.

c. $\rightsquigarrow$? There is a specific country other than Italy that Bill thinks Sue is not from.

(11) { Bill thinks Sue is from x | x is a country }.

That's the stumbling block I mentioned at the end of the previous section. The problem isn't that we can't derive the NR-effects. It's that the derivation seems to support another, more dubious inference as well.

I have a few things to say in response. First, I have not introduced any new machinery to derive either NR effects or the dubious inferences. I have simply used popular machinery that others have introduced to explain a range of focus-related effects like (1). If this machinery – say, the model of [Beaver and Clark 2008] – predicts the dubious inferences, then this is a problem not just for my account of neg-raising, but for the model itself. In other words, there are independent reasons to think that the problem can be solved.

A second point. I haven't marked the intonation of (8a) and (10a). If there's a particular intonational (rise-fall) contour on *barn* and *Italy*, the "dubious" inferences look more acceptable. This suggests that (9) and (11) are the right questions for a particular intonation of (8a) and (10a), but not for others.

Third. It's easy to see how the dubious inferences could be blocked. We'd have to assume that the question addressed by $\neg P(\phi)$ can involve unspecific alternatives to ϕ. If the "locations" in (9) include unspecific locations like *outside the barn*, we no longer predict an inference that the pigs are all at the same specific location.

Fourth. It is independently plausible that the question addressed by (9) allows for unspecific locations as answers. The question is *Where are the pigs?*. In normal contexts, this does not presuppose that the pigs are together in some specific location. Likewise, (10b) might address *Where does Bill think Sue is from?*, which may well be answered by *from East Asia*.

It's tempting to assume that the addressed question allows for both specific and unspecific answers: *in the field* as well as *outside the barn*. But these aren't mutually exclusive, and we might want the complete answers to *Where are the pigs?* to exclude one another. Worse, if we allow for unspecific answers that are entailed by the complement ϕ in the observed utterance $\neg P(\phi)$ then we can no longer derive the NR-inference to $P(\neg\phi)$.

For example, suppose the hearer takes (10a) to address a question that contains both *Bill thinks Sue is from Italy* and *Bill thinks Sue is from Europe*. We then can't predict the inference to (10b). The question presupposes that Bill has an opinion, but the opinion might be unspecific: Bill might think that Sue is from some European country or other. The further information that he lacks the opinion that Sue is from Italy then doesn't imply that he thinks she is from outside Italy.

In sum, we need a model according to which the addressed question satisfies two conditions: it must contain unspecific answers, to block the dubious inferences, but it must not contain answers $P(\xi)$ distinct from $P(\phi)$ whose complement ξ is entailed by ϕ, as otherwise we can't derive the NR-inference.

One way to explain this is to assume that the alternatives ξ_i in the addressed question $\{P(\xi_i)\}$ are mutually exclusive by default, but that context underdetermines their grain size. That is, for all the

hearer can tell, an utterance of (10a) might address any of the questions in (12), as well as other candidates of the same kind.

(12) a. { Bill thinks Sue is from x | x is a country }

b. { Bill thinks Sue is from x | x is a country in Europe or a continent other than Europe }

c. { Bill thinks Sue is from x | $x \in$ { Italy, outside Italy } }

No matter which of these questions is addressed, the hearer can infer (10b). The "dubious" inference to (10c), by contrast, would assume that the question is (12a), and for this the hearer has not enough evidence.

Another possibility is that the answer normally reconstructed for (10a) is the binary question (12c). More generally, we'd assume that an assertion of $\neg P(\phi)$, with a neg-raising predicate P, would normally be construed as addressing the question whether $P(\phi)$ or $P(\neg\phi)$. One would, of course, like to know why this should be so.

To this end, note that the binary question $\{ P(\phi), P(\neg\phi) \}$ can be determined in two ways: we could hold fixed P and let the complement range over ϕ and $\neg\phi$. But we could also hold fixed the complement ϕ and let the predicate range over P and $P\neg$. Intuitively, the first option construes the question addressed by (10a) as (13a), the second as (13b).

(13) a. Does Bill believe that Sue is from Italy or that she is from somewhere else?

b. Does Bill believe or disbelieve that Sue is from Italy?

The difference is somewhat elusive, given that disbelieving ϕ means to believe $\neg\phi$. I'm not sure how the two options would differ for definite plurals. The second option, however, might explain why the addressed question has the form $\{ P(\phi), P(\neg\phi) \}$ rather than $\{ P(\phi), P(\phi'), \ldots \}$: No sensible alternative to *believe*, when applied to *that Sue is from Italy*, yields *believes that Sue is from Spain*.

Varying the predicate also looks attractive for NR-adjacent phenomena like (14).

(14) a. Bill doesn't like beer.

b. $\rightsquigarrow$ Bill dislikes beer.

We can derive this inference if we assume that (14a) addresses the question { Bill likes beer, Bill dislikes beer }. Since the complement is a noun phrase (*beer*), it's not obvious that it can be negated. (Although it's not obvious either that it can't be negated.)

Again, though, one would like an explanation of why the addressed question is $\{ P(\phi), P(\neg\phi) \}$, even if there is no lexicalized form of $P\neg$ as in the case of definite plurals, or that of *want*.

(15) Bill doesn't want beer.

There may be no deep explanation. [Gajewski 2007] suggests that the lexicon encodes the alternatives to any expression and that neg-raising predicates P have $P\neg$ as an alternative, while other predicates don't. We could adopt this account. To complete the model of neg-raising, we'd then only need to invoke standard assumptions about alternatives and addressed questions, such as Beaver and Clark's Focus Principle.

In some respects, the resulting model would resemble Gajewski's. Gajewski, however, does not appeal to reconstructed questions. Following [Abusch 2005], he instead postulates a quasi-pragmatic principle by which any sentence, even under embeddings, triggers a special kind of ("soft") presupposition to the effect that at least one of its (lexically encoded) alternatives is true. These assumptions have been criticized: see especially [Romoli 2013: 297-303]. The Focus Principle version of the account might get around some of the objections.

Still, I would prefer a less stipulative explanation. As [Homer 2015: 58, 65] points out, Gajewski's model leaves unexplained why almost the same expressions are neg-raising across all human languages. Let's have a closer look at this issue.

3 Distribution

[Horn 1989] lists the following examples of (apparently[3]) neg-raising predicates: *think*, *believe*, *suppose*, *imagine*, *expect*, *reckon*, *feel*, *seem*, *appear*, *look like*, *sound like*, *feel like*, *be probable*, *be likely*, *figure to*, *want*, *intend*, *choose*, *plan*, *be supposed to*, *ought*, *should*, *be desirable*, *advise*, *suggest*. Horn observes that neg-raising predicates are typically "mid-scalar". For example, *believe* and *be certain* arguably have a common scale, on which *be certain* lies higher than *believe*. Accordingly, *believe* is neg-raising but *be certain* is not. But this doesn't fully explain the distribution. Plural predications aren't mid-scalar. There is also small variation across languages: *Hope* isn't neg-raising in English, but its German equivalent *hoffen* is, even though there is no recognizable difference in scale position.

The approach I have outlined may help explain the data. Let's begin with a comparison of *think* (NR) and *be certain* (not NR).

(16) a. Bill doesn't think that Sue is smart.
b. Bill isn't certain that Sue is smart.

I have suggested that (16a) triggers an NR inference because it is assumed to address a question like (17).

(17) { Bill thinks Sue is x | $x \in$ { smart, not smart } }.

The question presupposes that Bill has an opinion about Sue's intellectual abilities. (16a) denies that the opinion is favourable. So a hearer can infer that the opinion is unfavorable. We would predict a parallel inference for (16b) if it were assumed to address (18).

(18) { Bill is certain that Sue is x | $x \in$ { smart, not smart } }.

It is independently plausible, however, that we normally wouldn't interpret (16b) as addressing (18). At least not without special intonation.

pigs are [in the barn]$_F$ coincide with the Current Question.

3 There is no uncontroversial test for whether a predicate is neg-raising. The final classification must arguably draw on theoretical considerations, and might exclude some items on Horn's list. [Collins and Postal 2014], for example, propose a model of neg-raising that excludes *should*. [Homer 2015] excludes *usually*.

(16a) is congruent with (17) only if the focus lies on *is smart*. If instead the focus (only) lies on *think*, we would expect (16a) to address a question about Bill's attitude towards Sue's smartness, as in (19).

(19) { Bill x that Sue is smart | $x \in$ { be certain, think, doubt, … } }.

This fits the observation, mentioned in [Gajewski 2005], that NR effects can be cancelled by putting intonational focus on the predicate.[4]

So here's a conjecture. (16b) does not give rise to an NR inference because we tend to interpret *certain* as having focus. As a result, (16b) is not congruent with (18). The addressed question must be something else. The best candidate is (19). And (19) doesn't presuppose that Bill has an opinion as to whether Sue is smart.

Generalizing, NR effects arise for predicates that don't naturally carry focus. This makes sense of Horn's mid-scalar observation: strong expressions like *must* and *be certain* are more likely to carry focus than weaker expressions like *should* and *think*.

The same is true for unusual expressions. Compare (20a) and (20b).

(20) a. Bill doesn't believe that Sue is smart.

b. Bill doesn't harbour the belief that Sue is smart.

Believe is ordinary and inconspicuous. By contrast, the unusual choice of *harbouring the belief* attracts focus. If someone uses a strong or unusual expression,, it's likely that they want to address a question to which this expression is an answer, rather than part of the background.

For another illustration, compare (21a) and (21b), where I've used *doubt* to make the negation more natural.

(21) a. I doubt that the pigs are in the barn.

b. I doubt that all the pigs are in the barn.

One would expect the two sentences to be equivalent, but they have a markedly different effect. The explanation, I suggest, is that *all* attracts focus in a way *the* does not. As a consequence, (21a) is naturally interpreted as addressing a question like (22a), while (21b) addresses (22b).

(22) a. { The pigs are in x | x is a location }

b. { x of the pigs are in the barn | $x \in$ {all, some-but-not-all, none } }

Only the former question supports an NR inference.

An interesting case is that of *want* and *desire*.

(23) a. I don't want to leave.

b. I don't desire to leave.

To my ears, (23a) is more likely to trigger an NR inference than (23b). How come? I suggest that

4 The situation is a little complicated because intonational focus can have an emphasis function as well. *I can't BELIEVE that you did this* has a highly salient NR reading. In this context, we don't take the utterance to address a question about the speaker's attitude.

we normally hear (23a) with focus on *to leave*. Informally, (23a) elaborates on what the speaker wants to do: *want* is in the background, *to leave* in the foreground. In (23b), the more unusual choice of *desire* attracts focus. The question addressed by (23b) is more likely to be what attitude the speaker has towards leaving. But the situation isn't clear-cut. I can also hear an NR reading for (23b), especially if the sentence is pronounced in a "posh" accent. We can see why this could make a difference: for a posh speaker, *desire* is no longer an unusual choice of word.

What about the difference between English *hope* and German *hoffen*? Is *hoffen* more inconspicuous in German than *hope* in English? This isn't obvious. An element of conventionalization plausibly plays a role.

If my explanation is on the right track, understanding an utterance of the form $\neg P(\phi)$ requires knowing whether P is "conspicuous" or "inconspicuous". In terms of their statistical distribution, *hope* and *hoffen* might both be borderline cases. If you learn German, you might therefore be unsure whether you should infer *Hoffen*$(\neg P)$ from an utterance of $\neg$ *Hoffen*(P). But suppose you repeatedly hear people say things like (24).

(24) Ich hoffe nicht, dass er bestraft wird.
[I don't hope that he will be punished.]

Without the NR inference, this would be an oddly weak thing to say. The observation of utterances like (24) therefore suggests that *hoffen* is "inconspicuous" in German. Thus a borderline expression can *acquire* the status of "inconspicuous" by having an NR use.

All this is obviously only the beginning of a systematic theory. But the beginnings looks promising.[5]

Let's briefly return to the issue of the previous section. We saw that the proposed derivation of NR inferences would have implausible consequences unless the relevant utterances of $\neg P(\phi)$ are normally taken to address a question that contains unspecific alternatives like $P(\neg\phi)$. I mentioned that this could be explained by assuming that the addressed question holds fixed the complement ϕ and lets the predicate range over P has $P\neg$. The observations of the present section make this response somewhat less attractive. The difference between NR predicates and non-NR predicates, I have suggested, is precisely that NR predicates are normally backgrounded, so that the addressed question varies the complement and holds fixed the predicate.

(But the case is not clear. Here's an alternative explanation of the distribution data. Inconspicuous expressions like *believe* belong to smaller "registers" from which their alternatives are drawn. *Be certain* invokes the idea of graded belief. Its alternatives specify different grades. The only genuine alternative to *believe*, by contrast, might be *disbelieve*.)

4 Neg-Raising Symptoms

Let's go through some other symptoms associated with neg-raising and see what the explanation I have suggested might say about them.

5 [Zeijlstra 2018: 430] mentions *to be of the opinion* as a neg-raiser. This is a little surprising for the approach I have outlined, as the expression is intuitively unusual. Zeijlstra also mentions, however, that *to be of the opinion* differs

Defeasibility. [Bartsch 1973] points out that neg-raising inferences are context-dependent and defeasible:

(25) a. Bill isn't sure whether Brutus and Caesar lived at the same time. So, naturally, Bill doesn't think Brutus killed Caesar.

b. ⇝ Bill thinks Brutus didn't kill Caesar.

An example from [Homer 2015]:

(26) *Context*: At a job interview.

a. I don't want to make a lot of money.

b. ⇝ I want to not make a lot of money.

On the other hand, explicit cancellation is often difficult or impossible, as [Gajewski 2005] and [Romoli 2013: 300] note:

(27) Bill doesn't think that it's raining, # he is not sure.

How can we reconcile these observations?

The account I have outlined suggests an explanation. A sentence of the form $\neg P(\phi)$ can address different questions. $\neg Want(\phi)$, for example, might address $\{ Want(\phi), Want(\neg\phi) \}$ or $\{ Want(\phi), \neg Want(\neg\phi) \}$. Intonational focus usually helps to clarify which question is intended. I have not marked the above examples for intonation. The most natural focus location in (25a) and (26a) is on the negation, or possibly on the attitude verb. With focus on negation (and, to a lesser extent, on the attitude verb), (27), too, becomes acceptable. If the focus is on negation, it is clear that $\neg P(\phi)$ is meant to elaborate on whether or not $P(\phi)$ obtains: the question is $\{ P(\phi), \neg P(\phi) \}$, not $\{ P(\phi), P(\neg\phi) \}$.

Projection Failure. NR-effects are sometimes called "homogeneity presuppositions", but they don't behave like standard presuppositions. In particular, the excluded-middle/homogeneity assumption, $P(\phi) \vee P(\neg\phi)$, doesn't project from conditional antecedents or epistemic modals (see, e.g., [Gajewski 2007: 295], [Romoli 2013: 302f.].) For example, (28) doesn't suggest that Bill has an opinion about whether Sue is smart.

(28) If Bill thinks/doesn't think that Sue is smart, he'll make this clear.

On the approach I have suggested, this is unsurprising. An utterance of $\neg P(\phi)$ triggers the homogeneity presupposition $P(\phi) \vee P(\neg\phi)$ because it is assumed to address a question like $\{ P(\phi), P(\neg\phi) \}$ whose union entails $P(\phi) \vee P(\neg\phi)$. An utterance of $P(\phi) \to Q$, however, normally doesn't elaborate on whether $P(\phi)$ or $P(\neg\phi)$. *Negation* tends to preserve the addressed question, arbitrary embeddings do not.

NR with VP Ellipsis. [Jacobson 2018, 2020] and [Crowley 2019: 3] consider sentences like

from other NR predicates in an important respect: it doesn't license strict NPIs. So maybe it's a different kind of beast.

(29a) and (29b).

(29) a. Bill didn't think it would snow but Sue did.

b. Bill thought that it would snow but Sue didn't.

In either case, the negated conjunct has a natural NR reading. This is a problem for syntactic movement accounts: if the negation is located under *think* in the LF of *Bill didn't think*, we would expect *Sue did* in (29a) to mean *Sue did not think it would snow*. But it does not.

The VP data do not look problematic for the proposal I have suggested. We would predict the NR-effect in (29a) and (29b) from the assumption that the two conjuncts address which of *Snow* and ¬*Snow* is believed by Bill and Sue.

NPI Licensing. [Lakoff 1969] observed that strong NPIs like *until tomorrow* are licensed in the complement of NR predicates, but not in corresponding NR predicates, as illustrated in (30)

(30) a. Bill doesn't think that Sue will leave until tomorrow.

b. Bill isn't certain that Sue will leave until tomorrow.

This has initially been thought to support a syntactic account of neg-raising, on the assumption that NPIs are licensed only under local (clausemate) negation. However, there are good reasons to think that the licensing conditions for strong NPIs are more liberal. [Chierchia 2013] and [Gajewski 2011] suggest that strong NPIs are licensed in all Strawson-DE environments. If this is on the right track (see [Zeijlstra 2018], [Crowley 2019], [Jacobson 2020] for discussion), the approach I have put forward has a good chance of explaining the data.

If an utterance of $\neg P(\phi)$ is assumed to address the question $\{P(\phi), P(\neg\phi)\}$, the question is accommodated as a pragmatic kind of presupposition. On this presupposition, $\neg P(\phi)$ is equivalent to $P(\neg\phi)$. If P is monotonic, ϕ is a DE environment in $P(\neg\phi)$. So ϕ is Strawson-DE if the addressed question is taken to be {$P(φ), P(¬φ)$}.

As we would expect, the felicity of NPIs depends on intonational focus and other clues about the addressed question:

(31) # Bill isn't aware that Sue has to leave. So, naturally, he [doesn't]$_{\mathrm{F}}$ think that Sue will leave until tomorrow.

Partial Cyclicity. When neg-raising predicates are stacked under (overt) negation, as in $\neg P_1(P_2(\phi))$, they often have a "cyclic" reading $P_1(P_2(\neg\phi))$ in which the negation is interpreted as located below the lowest predicate. This happens, for example, with *think* and *want* in (32a) – but not with the inverse order in (32b), as [Horn 1971] observed.

(32) a. I don't think Sue wants to help.

b. I don't want Sue to think I'm angry.

Could the difference lie in the addressed question? The approach I have outlined predicts the cyclicity in (32a) if the addressed question is something like { Think(Want(ϕ), Think(Want($\neg\phi$)) }. We would

predict the non-cyclicity in (32b) if the addressed question here is something like { Want(Think(ϕ)), Want($\neg$Think(ϕ)) }. But why this difference?

I don't have a full answer, but it is independently plausible that *I think* normally doesn't address a question about the speaker's attitude. Informally, the focus in (32a) isn't on what the speaker thinks, but on what Sue wants: whether she wants to help or not help. By contrast, the focus in (32b) is on the speaker's attitude (or its complement), on whether she wants Sue to think that she is angry.

This line of explanation is supported by the distribution of cyclicity across NR predicates (see, e.g., [Staniszewski 2017]). The clearest examples of cyclicity use epistemic verbs like *think*, *seem* or *suppose* as the higher predicate P_1, which are usually not the focus of the utterance. It's also notable that the cyclic reading is strongest if the higher predicate combines with a first-person subject. Compare (32a) with (33).

(33) Bill doesn't think Sue wants to help.

A cyclic reading of (33) is still possible, but it is less robust.[6]

Quantified Neg-Raising. So far, we've looked at neg-raising with simple, individual subjects. There are also cases with quantified subjects, as in (34) and (35).

(34) a. No guest wants to drink alcohol.
b. $\rightsquigarrow$ Every guest wants to not drink alcohol.

(35) a. Not every guest wants to drink alcohol.
b. $\rightsquigarrow$ Some guest wants to not drink alcohol.

Here we have to be careful. Let's begin with a simpler example:

(36) a. Some guests don't want to drink alcohol.
b. $\rightsquigarrow$ Some guests want to not drink alcohol.

One might assume that the question addressed by (36a) is (37), analogous to (38c) that we'd assume is addressed by (38a).

(37) { Some guests want y | $y \in$ { drink, not drink } }

(38) a. Bill doesn't want to drink [alcohol].

6 [Gajewski 2005: 53ff] explains the difference between (32a) and (32b) in terms of the different projection of presuppositions under *think* and *want*. Suppose, with Gajewski, that *Want(φ)* triggers the presupposition *Want(φ)* $\vee$ *Want($\neg\varphi$)*. Presuppositions under *Think* project into the doxastic alternatives. (32a) therefore implies that the speaker's doxastic alternatives satisfy either *Want(Help)* or *Want($\neg$Help)*. Since the former is ruled out by the utterance, one can infer that she thinks Sue wants to not help. Presuppositions under *Want*, however, project not into the bouletic alternatives, but into the doxastic alternatives. Given that *Think* presupposes *Think(φ)* $\vee$ *Think($\neg\varphi$)*, it follows that (32b) only implies that Sue has an opinion about whether the speaker is angry in all doxastic alternatives. But does (32b) really convey a belief that Sue has an opinion about whether the speaker is angry? I don't think so. Also, this explanation of the non-cyclicity in (32b) draws on a specific fact about *want*. It's not how it carries over, for example, to cases with *should* as the higher predicate.

b. ⇝ Bill wants to not drink [alcohol].

c. { Bill wants y | $y \in$ { not drink, drink } }

But then we couldn't predict the inference in (36). (36a) presupposes that *some* guests have a preference with respect to drinking. If all *these* guests prefer alcohol and all the others have no preference, (36a) is true and (36b) false, even though the question's presupposition is satisfied.

The same problem would arise for the inference in (39), if we assume that the addressed question is (40).

(39) a. Some guests don't eat $[\text{pizza}]_F$.

b. ⇝ Some guests eat something other than pizza.

(40) { Some guests eat x | x is a food }

This gives us independent reason to think that utterances like (36a) and (39a) address a stronger question. Intuitively, (39a) gives a partial answer to the question *who* ate *what*. The cells (elements) of this question would assign to each guest the food they consumed. (Alternatively, if the identity of the guests is not relevant, the cells might specify, say, how many guests consumed which kinds of food.)

Now return to (36). If (36a) addresses the question which guests want what kind of drink, we can explain the inference. Of course, we would also run into the problem from section 2: we get a dubious further inference, that each guest wants a particular drink. To avoid this, we have to assume that the addressed question is only about alcohol: it's about which guests want to drink alcohol and which prefer not to drink alcohol. If this is the addressed question, (34) and (35) raise no further problems (apart from the problem of accounting for the scalar implicature in (35), which is a different issue; see e.g. [Breheny et al. 2018].)

5 Neg-raising through presuppositional exhaustification?

I have suggested that NR-effects might arise through the same mechanism that gives rise to the inference in (1).

(1) a. We won't have $[\text{pizza}]_F$ for dinner.

b. ⇝ We will have something other than pizza for dinner.

I have assumed a particular approach to this inference, drawing on the idea that the interpretation of declarative utterances involves the reconstruction of a question. Other explanations of (1) have been proposed. [Herburger 2000], for example, suggests that the logical form of (1a) involves existential quantification over events: the negation is applied locally to the claim that the event involves pizza. It might be worth exploring whether this syntactic approach could be extended to neg-raising inferences, and how it would relate to syntactic accounts of neg-raising such as [Collins and Postal 2014].

I will not pursue this question. Instead, I want to briefly explore whether the inference in (1), as well as NR-effects, might arise from a tacit exhaustification operator *Exh*.

The *Exh* operator has figured prominently in recent work on scalar implicatures. (See, e.g., [Chierchia et al. 2012].) Its function is to strengthen its argument by conjoining it with the negation of its alternatives. More precisely, *Exh(ϕ)* conjoins ϕ with the negation of its *innocently excludable* alternatives, where an alternative is innocently excludable if it is contained in every set of alternatives whose members can be consistently denied while affirming ϕ. (All this assumes that semantics, perhaps with the help of pragmatics, supplies a base set of alternatives for each sentence ϕ.) Let IE(φ) be the set of innocently excludable alternatives to φ. A standard semantics of *Exh* is then given by (41).

(41) $[\![Exh(\phi)]\!] = [\![\phi]\!] \cap \bigwedge \neg[\![\psi]\!] : \psi \in \mathrm{IE}(\phi)$

To see how this might be relevant to (1), start with a positive version:

(3) We will have [pizza]$_F$ for dinner.

(3) suggests that we will have only pizza for dinner: we won't, for example, have pizza *and pasta and burgers*. This exhaustivity inference is surprisingly hard to explain with neo-Gricean resources (see [Cremers et al. 2023]). It can be explained if we assume that assertions are generally embedded in *Exh* (as suggested, for example, in [Magri 2009]). The logical form of (3) would then be (42).

(42) Exh(We will have [pizza]$_F$ for dinner).

Given that the alternatives to the embedded sentence include *We will have pizza for dinner* and *We will have pasta for dinner*, and these are innocently excludable, (42) entails that we'll have none of these other foods.

Now return to (1). Since *Exh* is a strengthening operator, its effect under negation would be to *weaken* the assertion. If (1a) is the negation of (42), it would assert that *either* we won't have pizza *or* we will have something other than pizza. This would be true if, say, we will have both pizza and pasta. Needless to say, (1a) is not normally interpreted in this way.

Why not? One hypothesis is that *Exh* does not like to scope under negation. (See, e.g. [Fox and Spector 2018].) In recent work, [Bassi et al. 2021, Del Pinal 2021, Del Pinal et al. 2024] have argued for a different explanation. They argue, on other grounds, that *Exh(ϕ)* merely *asserts* ϕ and *presupposes* the negation of the alternatives:

(43) $[\![Exh(\phi)]\!] = \begin{cases} \text{assertion: } [\![\phi]\!] \\ \text{presupposition: } \neg\mathrm{Excl}(\phi) \end{cases}$

(Here, and in what follows, I use $\neg Excl(\phi)$ as shorthand for $\bigwedge \neg[\![\psi]\!] : \psi \in \mathrm{IE}(\phi)$.)

Since presupposed content is preserved under negation, $\neg Exh(\phi)$ is not predicted to weaken $\neg\phi$. We still get a bad prediction for (1a), however: The utterance is predicted to assert that we won't have pizza and presuppose that we won't have anything other than pizza. The combined effect would be to convey that we'll have nothing for dinner. This is the opposite of what we observe: in fact, (1a) conveys that we'll have *something* for dinner, just not pizza.

[Bassi et al. 2021] present other data suggesting that the presuppositional content of *Exh* is actually weaker than what (43) suggests. To account for their data, they make the presupposition conditional:

(44) $[\![Exh(\phi)]\!] = \begin{cases}\text{assertion: } [\![\phi]\!] \\ \text{presupposition: } [\![\phi]\!] \rightarrow \neg\text{Excl}(\phi)\end{cases}$

Another possibility, which they don't consider but which would equally fit their data, is that the presupposition is biconditional:

(45) $[\![Exh(\phi)]\!] = \begin{cases}\text{assertion: } [\![\phi]\!] \\ \text{presupposition: } [\![\phi]\!] \leftrightarrow \neg\text{Excl}(\phi)\end{cases}$

To motivate this, note that the presupposed content in (45) is equivalent to the conjunction of the presupposed content in (44) with the hypothesis that at least one element of IE(ϕ) is true. In fact, if the alternatives in IE(ϕ) are mutually exclusive, then $P(\phi) \leftrightarrow \neg\text{Excl}(P(\phi))$ is equivalent to the disjunction of the alternatives in IE(ϕ). It's intuitively plausible that assertions should carry this presupposition. (Compare [Abusch 2005].) The assumed presupposition in (44), by contrast, looks somewhat arbitrary.

With (45), we could explain the inference in (1). Schematically, $\neg Exh(\phi)$ is predicted to assert $\neg\phi$ and presuppose $\phi \leftrightarrow \neg\text{Excl}(\phi)$. Presupposition and asserted content together imply $\text{Excl}(\phi)$. That's what we wanted: (1a) conveys that we will have something other than pizza. With a focus-sensitive account of alternatives, as in [Fox and Katzir 2011], we could also predict why the inference in (1) depends on intonational focus.

Now return to the neg-raising inference in (2).

(2) a. Bill doesn't think that Sue is smart.
b. $\rightsquigarrow$ Bill thinks that Sue isn't smart.

Suppose we read (2a) as (46):

(46) $\neg$Exh(Bill thinks that Sue is smart).

This asserts that Bill doesn't think that Sue is smart, and presupposes that Bill thinks that Sue is smart iff he has no other opinion about Sue's intellectual abilities. Together, the two components imply that Bill has some other opinion about Sue's intellectual abilities.

This looks promising. Again, we could explain the focus-sensitivity of the effect by the focus-sensitivity of alternatives. The inference arises if the alternatives to *Bill thinks that Sue is smart* are { Bill thinks that Sue is smart, Bill thinks that Sue isn't smart }, but not if they are { Bill thinks that Sue is smart, Bill isn't sure that Sue is smart }.

The issue of section 2 would still arise, of course: we'd have to explain why the base alternatives to $P(\phi)$ include alternatives with unspecific complements like $P(\neg\phi)$, in order to prevent "dubious inferences" like the one from (10a) to (10c) in section 2. Essentially the same options are available here as there.

Above I mentioned that if the alternatives to $P(\phi)$ are mutually exclusive (as they plausibly should be to avoid the dubious inferences), the presupposition $P(\phi) \leftrightarrow \neg\text{Excl}(P(\phi))$ is equivalent to the disjunction of the alternatives. The proposed derivation therefore resembles that of [Gajewski 2005, 2007], mentioned at the end of section 2, according to which NR effects arise from a presupposition that at least one of the alternatives to $P(\phi)$ is true. On the present model, however, this

would not be triggered by a (possibly problematic) quasi-pragmatic principle that applies even under embeddings, but by the Exh operator.

This brings me to a problem. We know that "homogeneity presuppositions" don't project like normal presuppositions:

(47) a. If Bill thinks/doesn't think that Sue is smart, he'll make this clear.

b. $\not\rightsquigarrow$ Bill has an opinion about whether Sue is smart.

If the antecedent of (47a) involves a tacit *Exh* occurrence that works as suggested in (45), we would predict that (47a) presupposes (47b).

This, however, seems to be a general problem for presuppositional exhaustification. Suppose, as suggested in [Bassi et al. 2021], that the scalar strengthening of *some* is due to the presence of *Exh*, whose semantics is given by (44). That is, (48a) implies (48b) because is logical form is *Exh(Some)*, which presupposes $Some \rightarrow \neg All$.

(48) a. Some students took the exam.

b. $\rightsquigarrow$ Not all students took the exam.

If the antecedent of (49a) is embedded under *Exh*, we would predict the inference in (49). (Note that $Some \rightarrow \neg All$ is equivalent to $\neg All$.)

(49) a. If some/any students took the exam, I have marking to do.

b. $\rightsquigarrow$ Not all students took the exam.

But this inference looks, well, dubious.

6 Conclusion

I have suggested that neg-raising effects might be an instance of a broader phenomenon by which negative utterances have a positive side-effect. More concretely, I have suggested that neg-raising might be explained by resources familiar from the literature on focus. Such an explanation might be able to explain otherwise puzzling data, such as the distribution of NR predicates and the partial cyclicity of NR inferences.

References

Dorit Abusch [2005]: "Triggering from Alternative Sets and Projection of Pragmatic Presuppositions"

Renate Bartsch [1973]: ""Negative Transportation" Gibt Es Nicht". *Linguistische Berichte*, 27(7): 1–7

Itai Bassi, Guillermo Del Pinal and Uli Sauerland [2021]: "Presuppositional Exhaustification". *Semantics and Pragmatics*, 14: 11:1–48

David I. Beaver and Brady Z. Clark [2008]: *Sense and Sensitivity: How Focus Determines Meaning*. John Wiley & Sons

Richard Breheny, Nathan Klinedinst, Jacopo Romoli and Yasutada Sudo [2018]: "The Symmetry Problem: Current Theories and Prospects". *Natural Language Semantics*, 26(2): 85–110

Lauri Carlson [1983]: *Dialogue Games: An Approach to Discourse Analysis*. Dordrecht, Netherlands: Reidel

G. Chierchia [2013]: *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford University Press

Gennaro Chierchia, Danny Fox and Benjamin Spector [2012]: "Scalar Implicature as a Grammatical Phenomenon". In *Semantics: An International Handbook of Natural Language Meaning,* de Gruyter

Chris Collins and Paul M. Postal [2014]: *Classical NEG Raising: An Essay on the Syntax of Negation*, vol 67. MIT press

Alexandre Cremers, Ethan G. Wilcox and Benjamin Spector [2023]: "Exhaustivity and Anti-Exhaustivity in the RSA Framework: Testing the Effect of Prior Beliefs". *Cognitive Science*, 47(5): e13286

Paul Crowley [2019]: "Neg-Raising and Neg Movement". *Natural Language Semantics*, 27(1): 1–17

Guillermo Del Pinal [2021]: "Oddness, Modularity, and Exhaustification". *Natural Language Semantics*, 29(1): 115–158

Guillermo Del Pinal, Itai Bassi and Uli Sauerland [2024]: "Free Choice and Presuppositional Exhaustification". *Semantics and Pragmatics*, 17(3)

Danny Fox and Roni Katzir [2011]: "On the Characterization of Alternatives". *Natural Language Semantics*, 19: 87–107

Danny Fox and Benjamin Spector [2018]: "Economy and Embedded Exhaustification". *Natural Language Semantics*, 26(1): 1–50

Jon R. Gajewski [2011]: "Licensing Strong NPIs". *Natural Language Semantics*, 19(2): 109–148

Jon Robert Gajewski [2005]: *Neg-Raising: Polarity and Presupposition*. PhD Thesis, Massachusetts Institute of Technology

— [2007]: "Neg-Raising and Polarity". *Linguistics and Philosophy*, 30(3): 289–328

Elena Herburger [2000]: *What Counts: Focus and Quantification*. Cambridge, MA: The MIT Press

Vincent Homer [2015]: “Neg-Raising and Positive Polarity: The View from Modals”. *Semantics and Pragmatics*, 8: 4:1–88

L. Horn [1971]: “Negative Transportation: Unsafe at Any Speed?” In *Proceedings of Chicago Linguistics Society,* vol 7. Chicago: CLS, 120–133

Laurence R. Horn [1989]: *A Natural History of Negation*. Chicago: University Of Chicago Press

Pauline Jacobson [2018]: “Some People Think There Is Neg Raising, and Some Don’t: Neg Raising Meets Ellipsis”. *Linguistic Inquiry*, 49(3): 559–576

— [2020]: “Neg Raising and Ellipsis (and Related Issues) Revisited”. *Natural Language Semantics*, 28(2): 111–140

Robin Lakoff [1969]: “A Syntactic Argument for Negative Transportation”. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society,* vol 5. Chicago Linguistic Society, 140–147

Giorgio Magri [2009]: “A Theory of Individual-Level Predicates Based on Blind Mandatory Scalar Implicatures”. *Natural Language Semantics*, 17(3): 245–297

Craige Roberts [1996]: “Information Structure in Discourse: Toward a Unified Theory of Formal Pragmatics”. *Ohio State University Working Papers in Linguistics*, 49: 91–136

Jacopo Romoli [2013]: “A Scalar Implicature-Based Approach to Neg-Raising”. *Linguistics and Philosophy*, 36(4): 291–353

Mats Rooth [1992]: “A Theory of Focus Interpretation”. *Natural language semantics*, 1(1): 75–116

— [1996]: “Focus”. In Shalom Lappin (Ed.) *The Handbook of Contemporary Semantic Theory,* Oxford: Blackwell

Frank Staniszewski [2017]: “Partial Cyclicity and Restrictions on Neg-raising”. In *Proceedings of the 34th West Coast Conference on Formal Linguistics, Ed. Aaron Kaplan et Al,* 490–497

Hedde Zeijlstra [2018]: “Does Neg-Raising Involve Neg-Raising?” *Topoi*, 37(3): 417–433