Short-circuited Implication: A Negative Contribution

Laurence R. Horn and Samuel Bayer

1

In this paper we shall re-examine the problems involved in providing an elegant account of so-called negative transportation or neg-raising (NR), the availability (with certain predicates) of lower clause understandings for higher clause negations. The principal difficulty for non-syntactic treatments of these ‘NR understandings’ is the existence of lexical exceptions to neg-raising, i.e., pairs of virtual synonyms (within or across languages) of which one member allows the lower clause understanding and the other blocks it. We shall propose and attempt to motivate a new account of the neg-raising phenomenon which can deal successfully with the existence of these lexical asymmetries without invoking a minor transformational rule of controversial character. This new approach, leaning on the notion of ‘short-circuited implication’ (Morgan 1978), will be extended to other cases in which some fundamentally pragmatic mechanism has become partially grammaticized and taken on morphological or syntactic correlates.

2

The phenomenon of so-called neg-raising presents a textbook example of the view of scientific ‘progress’ embodied in Kuhn (1962): as the theoretical paradigm shifts course, it leaves in its wake a raft of problems which have been cast adrift or marooned rather than actually solved. The neg-raising phenomenon itself is real enough; it has disturbed philosophers from St Anselm (cf. Henry 1967: 193ff), who complained that (1a), whose literal meaning is ‘he needn’t marry’, is taken to signify the stronger (1b), an injunction to celibacy,

\begin{enumerate}
  \item a. non debet ducere uxorem, lit. ‘NEG [he should take a wife]’
  \item b. debet non ducere uxorem, lit. ‘he should NEG [take a wife]’
\end{enumerate}


\begin{enumerate}
  \item a. \( x \) does not believe that \( p \).
  \item b. \( x \) believes that not-\( p \).
  \item c. it is not the case that \( x \) believes that \( p \).
\end{enumerate}

\begin{enumerate}
  \item a. I don’t think he has come.
  \item b. I think he has not come.
\end{enumerate}

For Jespersen, the ‘strong tendency in many languages to attract to the main verb a negative which should logically belong to the dependent nexus’ (1917: 53) is to be subsumed under a more general lexico-semantic account of contrary negation, but within the development of early transformational grammar, linguists led by Charles Fillmore (1963) and Robin Lakoff (1969) argued for a lexically governed cyclic transformational rule of neg-raising whereby the relevant strong reading of (3a) would be derived from a source with embedded negation, i.e., that which directly underlies (3b). Evidence for co-deriving (3a) and (3b) included the interaction of NR with negative polarity: strict polarity items like in weeks or for some time, normally requiring a tautological negative trigger (or durative predicate), are acceptable in (3a) where this requirement is not met at surface structure:

\begin{enumerate}
  \item a. \( \{ \text{think} \} \) he has not come in weeks.
  \item b. I don’t \( \{ \text{think} \} \) he has come in weeks.
\end{enumerate}

The transformational approach to neg-raising was rejected as otiose and theoretically suspect by interpretivists led by Jackendoff (1971), who proposed a semantic interpretive rule to associate the matrix negation of (3a) with the embedded clause.\(^2\) While Jackendoff conceded the ambiguity of (3a), he found parallel third-person examples (e.g. Quine’s (2a) above, where \( x \) does not refer to the speaker) unambiguous and ‘vaguely non-committal’. Indeed, the apparent synonymy between (2b)
and one reading of (2a) 'has nothing to do with the syntactic component – it may even have nothing to do with the semantic component' (Jackendoff, 1971: 291).

Jackendoff’s semantic line on the relation of (3a) and (3b) was later recast in pragmatic terms by Renate Bartsch. Her idea is basically the following:³ (2a) is unambiguous, its meaning equivalent to that of (2c). Under certain pragmatic conditions, the weaker statement (2a) can convey the stronger negative proposition (2b). In particular, there is a natural pragmatic assumption that the subject x has given some thought to the truth of the proposition p and come to some conclusion about it. Given this unmarked assumption that x has considered p and taken a stance towards its truth, (2a) and (2b) – while sharing no semantic representation – will express the same information relative to the discourse context. Mirroring the entailment (semantische Implikation) from (2b) to (2a), we have a contextually-determined pragmatische Implikation from (2a) to (2b). The schema of pragmatic inference can thus be given as in (5):

(5) a. $F(x, p) \lor F(x, \neg p)$  (context-dependent pragmatic assumption)
   b. $\neg F(x, p)$  (utterance of (2a) by speaker)
   c. $F(x, \neg p)$  (i.e. (2b): from (5a) and (5b) by modus tollendo ponens)

The weakest link in Bartsch’s inference chain is the premise in (5a), which Bartsch assumes can be triggered whenever F is instantiated by a propositional attitude verb: think, believe, want, and so on. However, as observed in Horn (1978: 179), there are counter-examples in both directions to this assumption: neg-raisers triggering like likely, probable, and advise which are not obviously propositional attitude predicates in Bartsch’s sense, and propositional attitude predicates like hope, realize, know, be certain, etc. which are not neg-raisers – whose negations, that is, do not permit a lower clause understanding or trigger strict polarity items, as observed in (4b) above.

Further, as noted in Epstein (1976) and Horn (1978), even when the context is explicitly set up so that the disjunction in (5a) can be assumed, we still fail to get the predicted interpretation (5c) when $F = \text{say, hope, fear}$, or any other non-neg-raiser, as we see in the examples of (6):

(6) a. I asked him whether or not Mary’s been here, and he told me:
   [he said she hasn’t been here in weeks.]
   [he didn’t say she’s been here in weeks.]
   b. I have quite strong feelings about this plan of his:
   [I hope he doesn’t go there until July.]
   [I don’t hope he goes there until July.]

(Notice, in particular, the impossibility of the diagnostic polarity items – italicized above – in the versions where they are separated from their triggering negation by a non-neg-raising predicate: cf. He didn’t think she’s been here in weeks.) Along the same lines, Epstein observes that even if everyone in the context ‘has extremely strong feelings about the prospect of Wallace’s nomination, so that everyone either fears that Wallace will be nominated or fears that he will not be nominated’, an utterance of (6’a) does not ‘yield’ (6’b):

(6’) a. John doesn’t fear that Wallace will be nominated.
   b. John fears that Wallace will not be nominated.

‘This sort of “derived negative transportation”’, Epstein observes (1976: 158), ‘seldom if ever occurs in speech’. Yet nothing in Bartsch’s radical pragmatic account rules it out.

Moreover, the arguments first levelled by George Lakoff (1970) against a purely semantic treatment of neg-raising still militate against any of the variety of non-syntactic treatments which have been proposed since. As is well known, the availability of NR understandings is subject to semantically unmotivated lexical exceptions. Thus, suppose neg-raises on its parenthetical reading for all speakers, but guess does so only for some (cf. I don’t [suppose% guess] Max will arrive until midnight). Wish neg-raises freely, desire only with difficulty; the same pattern obtains for expect and anticipate.

Cross-linguistic disparities arise as well, although similarities in the class of NR triggers outnumber the differences. Among opinion verbs, Hebrew xosoq (‘think’) permits NR readings while maamin (‘believe’) does not; the opposite pattern obtains in Malagasy. While negation over hope cannot in general be associated with the lower clause or trigger strict polarity items (as seen in (6b); cf., however, I never saw a purple cow/I never hope to see one), its cross-Germanic counterparts (e.g. German hoffen) can. French souhaiter (‘hope, wish’) neg-raises, but its near-synonym espérer does not – although its Latin etymon sperare did, and so on (cf. Horn 1978: 183–7 for additional examples). The currently fashionable view that the lower clause understandings of certain higher clause negatives are to be assigned by a semantic or pragmatic device has yet to explain how a semantic interpretive rule or, a fortiori, a pragmatic principle can admit arbitrary lexical exceptions.

Despite these formidable problems, the alternative syntactic conception of neg-raising is even less appealing. Horn (1978) summarizes the arguments for a syntactic NR rule and concludes that these arguments, based on the interaction of neg-raising with other syntactic rules and processes, and with the complex panorama of negative polarity, are untenable, indecisive, or dependent on additional assumptions which prove to be theoretically
and/or empirically dubious at best. Of course, within the extended standard theory, or any of its hyperextensions, the question hardly gets to arise. Whether it would be formulated as something like (7)

\[
(7) \quad [X \hspace{1em} V \hspace{1em} Y \hspace{1em} \text{COMP} \hspace{1em} \text{[not} \hspace{1em} W]\text{,}]_a \\
1 \hspace{1em} 2 \hspace{1em} 3 \hspace{1em} 4 \hspace{1em} 5 \hspace{1em} 6 \\
\Rightarrow 5 + 1, \hspace{1em} 2, \hspace{1em} 3, \hspace{1em} 4, \hspace{1em} 0 \text{ or } 7, \hspace{1em} 6
\]

or simply treated as a mutant subcase of Move \(\alpha\), NR would be in arrant violation of at least the Tensed \(\mathcal{S}\) Condition (as noted in Wasow 1972) or any of its parametrized offspring, as well as of the aesthetic sensibilities of the theory’s practitioners.

In addition, even most of today’s non-formally extended standard theoreticians would be justifiably suspicious (if not downright horrified) of a rule which plays so much havoc with derivational morphology. Any coherent transformational programme must countenance a syntax in which NR feeds the incorporation rules which result in the formation of the lexical items in (8) (cf. Horn 1978: 170–1):

\[
(8) \quad \text{nobody} \hspace{1em} \text{-er than} \\
\text{neither ... nor} \hspace{1em} \text{few} \\
\text{neither/none of} \hspace{1em} \text{scarcely (any)} \\
\text{only} \hspace{1em} \text{doubt}
\]

Notice that the sentences of (9) are understood with the italicized incorporated negative taken in each case to be within the scope of the boldface neg-raising predicate to its right:

\[
(9) \quad \text{Nobody supposes that nuclear war is winnable.} \\
\text{Neither Mutt nor Jeff think(s) that this theory will last until tomorrow.} \\
\text{Neither/None of them is (are) likely to marry you.} \\
\text{Only John intends to vote for me.} \\
\text{Bill spent more than he should have.} \\
\{\text{Few} \hspace{1em} \text{Sarcely any}\} \hspace{1em} \text{of my friends think that you’d lift a finger for me.} \\
\text{I doubt that he wants to attend your talk.}
\]

Thus, the first example is taken to indicate that everybody supposes nuclear war is not winnable, the second that both Mutt and Jeff think the theory will not last until tomorrow, and so on. This consequence of a syntactic NR rule amounts to a reductio for those who subscribe to the currently prevailing view that rules of word-formation (as opposed to rules of inflectional morphology, on some accounts) are not to be interspersed with rules of syntax.

Horn (1978: 193–208) proposes a pragmatic line which attempts to provide a set of necessary criteria which must be satisfied by any would-be neg-raising predicate. Within this account, stative predicates whose lower bound is just above the midpoint on the relevant pragmatic scale (epistemic or deontic) may permit neg-raising. The ‘metascale’ incorporating epistemic and deontic values, but not quantifiers (cf. Horn 1978: 203–4), is reproduced as (10) (Table 73.1).

The basic idea is a functional one:

What is common to all NR predicates is the relative slimness of the functional difference between the preraised form with lower negation and the logical form with the upstairs negative taking wide scope. It is the closeness of the external (contradictory) readings of not likely, not believe, not advisable to likely not, believe not, advisable not, respectively, which renders the negated predicates potential neg-raisers, and the relative distance of not possible, not realize, not obligatory from possibly, realize, obligatory which removes these from contention . . . NR will not be conditioned when it would systematically result in the emergence of pernicious ambiguities (as when the higher-S and lower-S readings of main clause negations would carry a high functional load), leading to a breakdown in a communication.

(Horn 1978: 196–8)

<table>
<thead>
<tr>
<th>Table 73.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(10)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>be able</th>
<th>be possible</th>
<th>believe, suppose, think</th>
<th>know, realize</th>
</tr>
</thead>
<tbody>
<tr>
<td>be likely</td>
<td>probable</td>
<td>figure to seem, appear, look like</td>
<td>be clear, evident</td>
</tr>
<tr>
<td>not be</td>
<td>to be sure</td>
<td>be odd, significant</td>
<td>be sure, certain</td>
</tr>
<tr>
<td>able</td>
<td>advisable</td>
<td></td>
<td>be odd, significant</td>
</tr>
<tr>
<td>to be</td>
<td>desirable</td>
<td></td>
<td>be odd, significant</td>
</tr>
<tr>
<td>not</td>
<td>be good idea</td>
<td></td>
<td>be odd, significant</td>
</tr>
<tr>
<td>certain</td>
<td>choice</td>
<td></td>
<td>be odd, significant</td>
</tr>
<tr>
<td>must</td>
<td>have to make</td>
<td></td>
<td>be necessary</td>
</tr>
<tr>
<td>have to</td>
<td>be necessary</td>
<td></td>
<td>be necessary</td>
</tr>
<tr>
<td>order</td>
<td>be obligatory</td>
<td></td>
<td>be necessary</td>
</tr>
<tr>
<td>cause</td>
<td>force</td>
<td></td>
<td>be necessary</td>
</tr>
<tr>
<td>force</td>
<td>order</td>
<td></td>
<td>be necessary</td>
</tr>
<tr>
<td>demand</td>
<td>require</td>
<td></td>
<td>be necessary</td>
</tr>
</tbody>
</table>

Given that the weak scalar values in the leftmost column of (10), e.g. able or possible, whose negations are strong scalar elements on the corresponding negative scale (not able, impossible), and strong epistemics of the rightmost column, e.g. know or be certain, whose negations are weak scalar values, are both automatically stricken from the ranks of potential neg-raisers, we are left with the mid-scalars of the middle column, whose
negations are themselves mid-scalar elements on the corresponding negative scale.\(^5\)

While the mid-scalar criterion permits a generalization on the characteristics of this set of predicates, cross-cutting the semantic classes into which these predicates seem to fall, in English (as exemplified in (11)) and in other languages.

(11) [OPINION] think, believe, suppose, imagine, expect, reckon (%anticipate, %guess)
    [PERCEPTION] seem, appear, look like, look as though
    [PROBABILITY] be likely, be probable, figure to
    [INTENTION/VOLITION] want, intend, choose, plan
    [JUDGMENT/WEAK OBLIGATION] ought to, should, be supposed to, be a good idea, be desirable, advise, suggest

Horn (1978) is reluctantly forced to leave open the ultimate treatment of the phenomenon itself. His conclusion that neg-raising is a functionally based pragmatic process which has become grammaticized in English and other languages is more in the line of faute de mieux than a powerful and satisfying insight.

On balance, a non-syntactic account of neg-raising phenomena is clearly desirable, but is it possible? Can those pesky lexical exceptions be explained, or at least explained away? What is needed is a way of treating the lower clause understanding of certain higher clause negatives, and the syntactic correlates of that understanding, as a pragmatic association made possible by a certain semantic configuration, possibly involving the assumed disjunction of Bartsch (1973) and the mid-scalar criterion of Horn (1978) – an association which is, however, not freely available to all predicates with the proper semantic qualifications. That is, we need a device for allowing at least some pragmatic rules to admit lexical exceptions.

Fortunately, just such a device is now on the market, having been designed to handle a superficially quite different phenomenon, the Indirect Speech Act. It has long been known that apparently synonymous expressions may differ in their illocutionary act potential. Thus, the request in (12)

(12) Close the door.

(13a), but not (at least not without several additional degrees of indirectness) by asking the questions in (13b) and (13c):

a. Can you close the door?
b. Are you able to close the door?
c. Do you have the ability to close the door?

If, as the Gricean line on indirect speech acts laid down by Searle (1975) stipulates, (13a) conversationally implicates (12) via the maxim or relation (or relevance), this implicature should be non-detachable; its disappearance when we move from (13a) to the basically synonymous (13b) and (13c) is a mystery. It is just this mystery which motivated Sadock (1972) to reject the conversationalist line on indirect speech acts (as embodied in, e.g., Gordon and Lakoff 1975) in favour of a theory which posits a semantic ambiguity for (13a), with the relevant request reading analysed as constituting a ‘speech act idiom’. Searle (1975), however, observes that there can be conventions of usage which are not meaning conventions and that, by these conventions, ‘certain forms will tend to become conversationally established as the standard idiomatic forms for indirect speech acts’, e.g., (13a) but not (13b) or (13c) for indirect requests (Searle 1975: 76ff).

Carrying Searle’s idea one step further, Morgan (1978) motivates the notion of the short-circuited implicature (SCI) for such cases as (13a), where the implicature of (12) is in principle calculable (as are all conversational implicatures by definition) but it is not in practice actually calculated by speakers operating with the relevant usage conventions. Like Searle, Morgan rejects Sadock’s speech act idiom thesis as unpar- simonious, arguing that it is, however, conventional to use (13a), with its literal interrogative meaning, to convey (12). Morgan cites a plethora of additional candidates for short-circuited implicature status, including the conventional use of (14a), but not the semantically akin (14b), to wish a performer good luck before a performance:

(14a) Break a leg!
    b. Fracture a tibia!
       Break your leg!
       I hope you break a leg!

and the use of (15a), but not (15b), as an emphatic affirmative:

(15a) You can say that again!
    b. You can repeat that!
       You're able to say that again!
       You're capable of saying that again!
You're permitted to say that again! It's possible for you to say that again!

Morgan points out that while these conventions of usage do not themselves constitute idioms, as (16) shows:

(16) John really broke a leg last night! (≠ he really performed well)

such a development (i.e. from metaphor to idiom, from usage convention to meaning convention) is quite plausible, and indeed frequently attested. Grice himself pointed out that it may well be possible 'for what starts life . . . as a conversational implicature to become conventionalized' (Grice 1975: 58), and the SCI is a natural halfway house along this route. As instances in which such a shift has become complete, Morgan offers the valedictory in (17):

(17) (May) God be with you ⇒ Goodbye

(which was fully conventionalized the first time an atheist could utter it sincerely), and the euphemism in (18):

(18) go to the bathroom

(whose shift in literal meaning is demonstrated by the non-contradictory status of the complaint 'Your dog went to the bathroom on my living-room rug'.

The Searle–Morgan approach explicitly recognizes the existence of speakers' pragmatic conventions, alongside the better-understood conventions of syntax and semantics. Since the short-circuiting of implicatures is a matter of convention, we expect to find differences between speakers and between languages as to just which conventions of usage are operative. And indeed, Searle notes (1975: 76) that the Czech question corresponding to (13a) is not standardly used to issue a request. Green (1975) has amply demonstrated the range of cross-linguistic variation as to which questions are used to convey which requests. But Green also acknowledges that some indirect speech acts are more indirect than others, and that the most indirect requests – like (19a) and (19b) as hints for getting someone to close the window.

(19) a. the rain that's coming in the window is ruining the rug.
   b. It's freezing in here.

- do not exhibit the same variation. This is precisely to be expected if, as is plausible, hints involve non-short-circuited, and hence non-detachable, implicatures: only the literal meaning and extra-linguistic context can be relevant for determining what is hinted, not the choice of form of expression which is used to express that meaning.

One important morphosyntactic correlate of the short-circuiting of the request implicatum is the much-discussed preverbal or sentence-final please; this item appears readily in (13a), as in the literal request (12), but not in (13b) or (13c), as observed by Sadock (1974) (cf. Gordon and Lakoff 1975):

(20) a. Please close the door.
   b. Can you please close the door?
   c. Are you able to (*please) close the door?

The fact that preverbal or final please is also ruled out in the hints of (19), while readily explainable if (13a) but not (19) is a Sadockian speech act idiom, is also straightforwardly accounted for if please is triggered by the usage convention Searle and Morgan have proposed.

Similarly, as Alice Davison has reminded us, the 'speech act modifiers' of Davison 1975 also distinguish between direct and conventional (i.e. short-circuited) indirect speech acts on the one hand (cf. (21a) below) and non-conventional indirect speech acts (including hints) on the other (cf. (21b)):

(21) a. Shut the door
   Would you shut the door
   Could you shut the door
   because it's cold in here.
   b. Are you able to shut the door
   Are you willing to shut the door
   The heat is off
   I can't reach the door
   ??because it's cold in here.

4

The way is now clear for subsuming the conveyed meaning of neg-raising constructions under the treatment prescribed for indirect speech acts. While questioning the hearer-based preparatory condition on requests is a natural way to convey that request indirectly, only certain forms of those questions are conventionalized for that purpose (cf. Searle 1969, 1975). So too, while we have seen that predicates of a given semantico-pragmatic nature, e.g. mid-scalar propositional attitude verbs (a class including think, believe, want, hope and guess but not know or say), are candidates for the implicature permitting downstairs understandings of upstairs negation, whether a predicate within the appropriate class does in fact trigger the neg-raising
understanding depends on whether this implicature has been short-circuited into a usage convention. This explains why we get exactly the same sort of variation across languages, dialects and idiolects in neg-raising that Searle and Green have cited for indirect speech act potential. Thus, the notion of short-circuited conversational implicature captures just what is pragmatic (natural) and what is arbitrary (conventional) about the neg-raising phenomenon.

There is an even stronger parallel between indirect speech acts and NR understandings. One relatively straightforward way to accommodate the Bartschean pragmatic inference schema (cf. (5) above) within the Gricean framework for conversational implicature is to treat the crucial ‘strengthening’ step, which takes us from, e.g., (2a) to (2b), as being motivated by the maxim of relation (‘Be relevant’). But this is precisely the way Searle motivates indirect speech acts; in both cases, the literal meaning is strengthened and narrowed by the additional conveyed meaning. The problem posed by the fact that the relevance-based implicature is evidently detachable in the former case (possibly disappearing, for instance, when guess is substituted for suppose) reduces to the detachability problem in the latter case, both to be treated through the admission of pragmatic conventions applying to a proper subset of linguistic expressions appearing in the relevant frame.

5

We have noted that preverbal and final please, as well as speech act modifying because, co-occur with both direct and short-circuited indirect requests. We can view in the same light the phonological and morphosyntactic effects of the (short-circuited) indirect suggestions in (22), first discussed extensively in Gordon and Lakoff (1975):

(22) a. Why don’t you move to California
    \{ Why doncha
    Whyncha \} move to California
    [suggestion only]
    (cf. ?* \{ Why doncha
    Whyncha \} resemble your father.)

b. Why paint your house purple(?) [only a suggestion not to]

Following Searle (1975), we can assume that the full and reduced forms in (22) are all literal questions which may (and in the crucial cases conventionally will) implicate suggestions. We can then treat phonological reduction, as in (22a), and the ‘YOU + TENSE deletion’ of (22b), as epiphenomena of the short-circuited of the suggestion implicatum. The closely related full form corresponding to (22b), (23),

(23) Why are you painting your house purple?

is pragmatically ambiguous between a question and a (very indirect) suggestion (on top of that question); the absence of reduction and the rising contour signal that the implicature (if it applies at all, depending on the context) is not short-circuited in this case.

Similarly, WH-queclaratives (cf. Sadow 1974) like (24a) are literal questions, pace Sadow, which are conventionally used to convey the corresponding universally quantified negative assertion (equivalent, of course, to a negative existential), here (24b):

(24) a. Who (but a total idiot) would have said a thing like that?
    b. Nobody (but a total idiot) would have said a thing like that.

Correlated with this SCI is the appearance of but in the sense of ‘except’, an item which is normally restricted to universally quantified assertions, as indicated in (25):

(25) Everyone but Mary . . .
    Nobody but John . . .  \{ All
    Anyone but Carter . . . *Most
    *Somebody but Kim . . . *Many
    *Anywhere but here . . . *Three
    *Somewhere but here . . . *Some
    \} of my friends but Chris . . .
    None

Note, in particular, that non-queclarative questions, which may implicate declaratives in a given context but do not do so by convention, exclude this diagnostic:

(26) ?*Who but John is coming to the party?

Seen in this light, the extension of the distribution of strict negative polarity items like until or in weeks from environments containing a tautocausal negative to those in which the triggering negation is separated from the polarity item by a so-called neg-raising predicate correlates directly with the presence of the SCI associated with such a predicate, in the same way that preverbal please extends from direct to short-circuited indirect requests.

Given that all English speakers conventionally use verbs like think and suppose so as to transmit their negation to the downstairs clause, while only
some speakers operate with the analogous convention for *guess*, and none do for *hope*, or *say*, the distribution of polarity *until* follows accordingly.\(^6\)

\[
(27) \quad \text{I don't} \quad \begin{cases} \text{think} \\ \text{*guess} \end{cases} \quad \text{they'll hire you until you shave off your beard.}
\]

Put metaphorically, whether a given predicate will be transparent to negation depends on the presence of the SCI; the acceptability of the polarity item in contexts like (27) depends in turn on this transparency.

On the present analysis, (27a) is not rendered acceptable by being co-derived with (28)

\[
(28) \quad \text{(I think) they won't hire you until you shave off your beard.}
\]

nor by sharing any semantic representation with (28), but rather because it is conventionally used to implicate (28). Since no such usage convention exists for any speaker with *hope*, the polarity item in (27c) is ruled out, even when a negative proposition is indirectly conveyed in a given context (cf. the examples in (6) above). It must be borne in mind that while conversational implicature is basically (for Grice and subsequent writers) a matter of *parole*, the short-circuiting of implicature into a usage-convention takes place, as it were, on the boundary between *parole* and *linguistic rule*; it is this feature which enables SCIs to trigger the various linguistic correlates we have illustrated, including the strict polarity item *until* in (27a) and – in the relevant dialect – (27b).\(^7\)

Indeed, for a significant range of their distribution, strict polarity items seem to be licensed by the conventionally signalled presence of negation in the clause or proposition in which they occur, whether the relevant convention involves meaning or usage. Thus consider, alongside the pattern in (27), the contrasts in (29) and (30):

\[
(29) \quad \begin{align*}
\text{a.} & \quad \text{I'll be dammed if I'll hire you until you cut your hair.} \\
\text{b.} & \quad \text{*I'll be surprised if he hires you until you cut your hair.}
\end{align*}
\]

\[
(30) \quad \begin{align*}
\text{a.} & \quad \text{Why get married until you absolutely have to?} \\
\text{b.} & \quad \text{*Why are you getting married until you absolutely have to?}
\end{align*}
\]

In each pair, only the (a) construction is conventionally used to express the negative proposition which licenses *until*, viz.

\[
(29') \quad \text{I won't hire you until you cut your hair.}
\]

\[
(30') \quad \text{You shouldn't get married until you absolutely have to.}
\]

The non-short-circuited negative implicata which may be associated with

the (b) examples are insufficient to trigger such strict polarity items, although weaker (more lenient) negative polarity items may be acceptable in the same environments:

\[
(31) \quad \text{I'll be surprised if he \begin{cases} \text{hires anybody for that position.} \\
\text{lifts a finger to help you.}\end{cases}}
\]

The line on negative polarity sketched here seems to us to be consistent with recent theories presented in Ladasaw 1980 and Lineberger 1981, without actually following from either of them.

6

In the examples that we have been considering, the critical environments for determining the acceptability of a given item are those where an expression not only *can be used*, but is conventionally used (within the relevant dialect) to convey something not literally expressed. It appears that Searle's non-meaning conventions, Morgan's short-circuited implicatures, or some more precise and explanatory refinement of these notions which is still waiting to emerge, will play an essential role in any successful account of so-called neg-raising, and other pragmatic or functionally based phenomena which are less than fully productive across a given syntactic construction type or semantic class.

Before concluding, we find that an important methodological issue remains to be addressed. Given our current state of knowledge, it must be conceded that ascribing some linguistic phenomenon to the presence of a short-circuited implicature may amount more to labelling than to explaining that phenomenon. By pushing the problem of variation in indirect speech act potential back to the pragmatics, for example, we (along with Searle and Morgan) have in some sense reconstructed Sadow's speech act idiot analysis in different garb, rather than replacing it with a new, improved theory. There is the additional danger that short-circuited implicatures, like the related transferential 'conversational postulates' approach of Gordon and Lakoff (1975), are so powerful that they could (if allowed to run wild) sweep away all traditional forms of syntactic argumentation (cf. Sadow 1975). Our ultimate goal must be to discover just why the variation in usage conventions outlined above should exist, and why it should exist just where it does. One means of constraining the application of SCIs, as Alice Davison has pointed out to us, might be derivable from the similarity between the two principal cases we have investigated here: indirect speech acts and NR understandings for higher clause negation. This parallelism, as
we noted above, involves the fact that the implicatum being short-circuited in both cases is one derived via the maxim of relation.

More specifically, we can note the functional kinship between the indirect speech act example (32a) and the NR example (33a):

(32) a. I believe your answer is not wholly satisfactory.
   b. Your answer is not wholly satisfactory.
   (33) a. I don’t think your jumpsuit is entirely appropriate.
   b. I think your jumpsuit is not entirely appropriate.

In both cases, the effect of the SCI is to express a given proposition (= (32b), (33b)) in a qualified or ‘hedged’ way (cf. G. Lakoff 1972). The short-circuiting seems in each case (as well as in the related non-declarative indirect speech act examples of (13a), (21a) and (22a)) to serve the aims of politeness or face-saving. This feature has indeed often been acknowledged in connection with these constructions; cf. Searle (1975: 64), Fraser (1975), and various papers by Robin Lakoff on indirect speech acts as hedged assertions, requests, etc., and Prince (1976: 414ff.) on NR constructions as ‘metastatement hedges’. It seems plausible to derive the ‘politeness’ effect from the fact that in each case the weaker (a) sentence is in effect pragmatically ambiguous between two understandings which stand in a private relation (cf. Zwicky and Sadock 1975), with the stronger (b) version available to the addressee but not forced on her or him (unless the presence of diagnostics like those discussed in section 5 forces it). The present analysis, on which this stronger understanding is derived as an SCI, correlates with our sense that for most current speakers the politeness associated with examples like (32a) and (33a) is often felt as conventional only and not really ‘sincere’.

In any event, however the treatment of short-circuited implicature is to be narrowed and refined, we would argue that an improved classification of linguistic phenomena is not only useful in itself but represents an essential step on the road to full explanation. If we have correctly situated the issues underlying the ‘neg-raising’ phenomenon and strict polarity triggering, and correctly linked them up with other linguistic patterns not obviously related to them, perhaps we will have succeeded in building that first step.

Notes

* An earlier version of this paper was presented at the annual meeting of the Linguistic Society of America in San Diego, 3 December 1982. We wish to acknowledge those who commented on it at that meeting, especially Sue Schmeling, and readers of an earlier written version, including Alice Davison, Georgia Green, Jerry Morgan, and Jerry Sadock, for their helpful suggestions and critiques.

1 For further references and extensive discussion of the history and character of so-called neg-raising, cf. Horn (1978).
2 Cf. Pollack (1974) for the most extensive treatment along these lines, utilizing a formalized semantic interpretive rule of ‘neg-association’ whose effect mirrors the syntactic NR transformation of Fillmore and Lakoff. Partee (1970) presents an argument that (2a), (3a) and similar sentences with higher negation are not semantically ambiguous, on basically a priori grounds. These positions are challenged in Horn (1978).
3 Halpern (1976) arrives independently at a rather similar pragmatic analysis; cf. also Epstein (1976).
4 Unless the assumption is made that the negative element hops into and out of COMP (as in the successive-cyclic movement of WH-words); this assumption, which may well be motivated for Modern Irish (as Jim McCloskey has noted), is ad hoc and gratuitous for English, French and the majority of other languages for which NR would be posited.
5 Actually, strong scalars (but not weak scalars) of the deontic group, as well as mid-scalar deontics, are among the neg-raisers in many languages (e.g. falloir in French, behar in Basque, velat in Russian). Cf. Horn (1978: 200–2) for discussion and additional examples.
6 In the discussion following the oral presentation of this paper, Sue Schmeling pointed out an important asymmetry between the behaviour of negation over guess (in dialects like hers, which do not permit NR interpretations with this predicate) and the short-circuited implicature cases of Morgan (1978). Schmeling finds sentences like (27b) – or the ‘NR reading’ of I don’t guess they’ll win – immediately interpretable, although not utterable by her, despite the fact that (on our account) she lacks the ‘relevant convention of usage’. On the other hand, no implicature-bearing interpretation is at all available for her with (14b), (15b) and the related examples of Morgan (1978). We submit that this quite palpable difference is attributable to the fact that, in the language of Grice (1975), the former (NR) case involves the presence of a (short-circuited) generalized conversational implicature, while the latter cases require us to invoke a particularized conversational implicature. Thus, the default case for semantically qualified predicates (at least in English) is that they do trigger the NR reading, but the presence of the short-circuited implicature for particular constructions of the type illustrated in (14) and (15) must be learned ad hoc. The distinction between these two types of positive and negative ‘pragmatic exception features’, as it were, is analogous in obvious ways to the more traditional distinction between types of lexical exception features for syntactic rules first posited in G. Lakoff (1965).
7 Situated elsewhere along this same borderline, but not to be confused with the SCI, is Grice’s notion of conventional implicature. Unlike conversational implicata (short-circuited or not), conventional implicata are – as their name betokens – part of an expression’s conventional meaning; while not affecting truth-conditions, they constrain the appropriateness of use for a given lexical item or phrase (or, alternatively, constrain the set of contexts in which that expression may be appropriately used). Like SCIs, conventional implicata are fully detachable; unlike SCIs – or indeed, any conversational implicata – they are non-cancellable and non-cancellable (cf. Karttunen and Peters 1979 for a more detailed account of conventional implicature). Within the metaphor employed earlier, conventional implicatures mark a stage on the use–meaning thoroughfare somewhere between SCIs and literal (truth-conditioning) meaning components; it might be argued that some of the examples in section 5, especially (22a) and (22b), involve not just short-circuited conversational implicata but fully conventional implicata,
Indirect Speech Acts

where (at least for some speakers) an actual meaning shift has occurred (cf. Cole 1975 for a convincing example of just such a shift). Ockhamist considerations dictate that such a move be posited only when there is ample positive evidence to support it.

(i) Your jumpsuit is [not entirely appropriate. ]

Tag questions may provide evidence for the actual content of the assertion in such cases, as R. Lakoff (1969) points out in her notorious example (ii):

(ii) I don’t suppose the Yankees will win, will they?

where the tagged proposition must be

(iii) The Yankees will not win.

In these cases, as in (32a), a speaker may (in Searle’s terms) indirectly assert $p$ by asserting the speaker-based sincerity condition on assertions of $p$, i.e. that the speaker believes $p$. The earliest clear description of such indirect assertions may be Wittgenstein’s:

The statement ‘I believe it’s going to rain’ has a meaning like, that is to say a use like, ‘It’s going to rain’ . . . When I say ‘I believe’ I am describing my own state of mind – but this description is indirectly an assertion of the fact believed – As . . . I describe a photograph in order to describe the thing it is a photograph of. (Wittgenstein 1953: 190)

Just as the pragmatic ambiguity of indirect requests can be resolved by the insertion of preverbal or final please, the indirect assertion use of I believe, I suppose, I guess, etc., can be signalled by postposing or ‘niching’ this adverbial-like qualifier:

(iv) Your jumpsuit, I believe, is not wholly appropriate.
(v) The Yankees will not win, I (don’t) suppose.

The politeness factor alluded to in the text may itself be a subcase of a more general ‘weakening’ of illocutionary force associated with both indirect speech acts and NR understandings. Urnson notes that the ‘parenthetical use’ of verbs like believe constitutes a ‘warning device’ which serves ‘to modify or weaken the claim to truth which would be implied by a simple assertion’ (1952: 484). Analogously, Poutsma and Bolinger have observed that the NR understanding of, e.g., (2a) is weaker than the corresponding sentence with lower clause negation, (2b). In Poutsma’s words, ‘the shifting of not often has the effect of softening down the negative of a sentence’ (1928: 105). This ‘softening’ effect of NR is discussed further in Horn (1978: 131–3), where additional cross-linguistic examples are cited as well.

References


Postscript (1995): Neg-raising Revisited: Tinkering with the Short-circuit (Laurence R. Horn)

In Horn and Bayer 1984 (henceforth H&B), we reviewed the difficulties encountered by extant treatments of the so-called neg-raising phenomenon (NRP) – the lower clause understanding available for higher clause negations, e.g. I don’t believe that p, understood as I believe that not-p – and made a case for deriving such readings as instances of short circuited conversational implicature (cf. Searle 1975; Morgan 1978). I return to the scene of the crime here to touch on three aspects of the problem that we either neglected or were not in a position to address the first time around: (i) the precursors to our view of the NRP as a conventionalized or standardized instance of a pragmatic strengthening implicature, (ii) the intellectual history of the notion of short-circuiting, and (iii) the viability of one particular competing analysis of the NRP.

H&B picked up where Horn (1978) left off. The earlier paper investigated in somewhat grisly detail the various problems confronting both purely grammatical and purely semantico-grammatical analyses of the NRP, concluding somewhat tentatively that the rule of neg-raising represents a grammaticalization of a fundamentally pragmatic process. The thesis of the later paper was suggested by Sam Bayer in a term paper for my undergraduate Pragmatics course at Yale: the conventionalization of the pragmatic rule (see (5)) takes place within the pragmatics itself. Since there is no neg-raising rule in the syntax, the contortions any such rule must accommodate (see the discussion of (8) and (9))1 are no longer a problem; since the implicature capturing the functionally motivated pragmatic process is short-circuited as a convention of usage, the existence of lexical exceptions to the NRP (see the discussion of (5) and (6)) is no longer an embarrassment it would be for any purely pragmatic (or semantic) account.

In Horn 1989: ch. 5, the H&B account of the NRP is situated within a broader perspective, following Bosanquet’s recognition that the neg-raised interpretation of “I don’t think that” – which is really equivalent to “I think that – not” – represents the same strengthening process as the habitual use of phrases such as I do not believe it, which refer grammatically to a fact of intellectual state, but actually serve as negations of something ascribed to reality (1911: 319). In each case, a formal contradictory negation is strengthened – ‘filled in’ – to yield a stronger contrary opposite: ‘The essence of formal negation is to invest the contrary with the character of the contradictory’ (Bosanquet 1911: 281). Such contradictions in contrary clothing extend from simple litotes, in which the denial of an unmarked value (She isn’t happy, I don’t like it) may be pragmatically strengthened in a given discourse context to yield an affirmation of the contrary value (She’s sad, I dislike it), through the NRP, where this same pragmatic strengthening of negative force is fossilized as a short-circuited implicature (whence the unmediated nature of the inference and the lexical exceptions discussed in H&B), and finally to lexically incorporated affixal negation (unhappy, dislike), where the contrary readings induced by the negative prefix are typically lexicalized into the semantics. The question of just when such strengthening rules apply in each case – with which predicates, on which stems, in which contexts – is addressed in detail in Horn (1989: ch. 5).

While St Anselm’s role as the discoverer of the NRP is discussed in both Horn 1978: 200) and H&B (see (1)), no mention was made in those studies of the role of Tobler (1882) as the first modern maven of neg-raising. His description of the ‘logically unwarranted position’ (logisch ungerechtfertigte Stellung) of the negative in the French Il ne faut pas que tu meures (=Il faut que tu meures pas, i.e. ‘You mustn’t die’) is particularly astute (i.e. consistent with the analysis of H&B). Tobler observes that the stronger (lower clause) understanding of negation associated with verbs like the French falloir, devoir, vouloir or the German sollen, wollen, meinen cannot be assimilated to simple litotes, for which he provides Old French and German examples translating as the holy saints by whom God was not hated (i.e. by whom He was loved) or he will do you little good (i.e. all the harm he can). Crucially, Tobler notes, the conscious irony associated with these cases of simply litotic strengthening is absent from the processing of the NRP examples.

Curiously, this point seems to have escaped Tobler’s student Kulekpy (1891), who bequeathed us an otherwise quite valuable monograph on the
proposes a treatment of neg-raising within the frameworks of Simon Dik’s Functional Grammar and Nuyt’s own Functional Procedural Grammar. While Nuyt’s paper contains many insightful remarks and apt criticisms of previous work, I am not convinced that the grounds on which he rejects the H&B analysis are sufficiently well supported — granted, of course, that I am not an objective witness on this matter.

It appears to be a fundamental property of Nuyt’s conception of the NRP that the predicates involved are all ‘propositional attitudes’. But however broadly that (perhaps not entirely helpful) category is construed, it does not obviously extend to those quantificational determiners (e.g. most) and adverbs (e.g. usually) whose mid-scalar positions qualify them as analogues of believe and want, despite the fact that these operators aren’t even predicates, let alone propositional attitudes. (The arguments for this position, outlined in Horn 1978: 198–204, are not addressed by Nuyts.) Nor does Nuyts explain why he concludes (1990: 568–70), contra H&B and the work on which it leans, that the non-neg-raisers impossible and certain are just as apt as to introduce a ‘degree of uncertainty’ into the evaluation of the proposition as are the mid-scalar neg-raisers probable and likely. Contra Nuyts, (A2) and (A3) entail that the referent is not crazy, whence the contradictory nature of the continuation, while (A4) involves no such entailment:

(A2) It is impossible that he is crazy (#but perhaps he is).
(A3) It is certain that he is not crazy (#but perhaps he is).
(A4) It is improbable that he is crazy (but perhaps he is).

Perhaps we are dealing with a dialect difference here, but I see no reason to renounce the functional/scalar generalization of Kålekpy (1891), Horn (1978), and H&B.

In the end, though, the main problems for Nuyts’s analysis are those for any analysis of the NRP which does not adopt some version of a standardized or conventionalized pragmatic inference mechanism: the old bugaboo of ‘lexical exceptions’. As argued in H&B, no non-stipulative explanation has ever been proposed for how a general cognitive or notional theory of the NRP can deal with the fact that hope is not transparent to negation in English while its Dutch counterpart hopen is (or why French espérer has lost the neg-raisability of its Latin source sperare), or the fact that while all English speakers allow a parenthetical reading for I guess (parallel to those for I suppose/think/reckon), only some speakers—a typically those hailing from the southern United States—allow a neg-raising effect for I don’t guess. (See Horn 1978: 183–7 and H&B: section 1 for more on these and other intra- and cross-linguistic instances of arbitrariness.)

Finally, it is worth trying (again) to dispel any doubts—such as those vouchsafed by Nuyts 1990: 564—as to why the ‘functional-syntactic
approach’ of Horn 1978 was traded in for the ‘clearly pragmatic’ one of H&B: ‘The reasons for Horn’s switch are not clear’. To help clarify, I conclude this revisitation with a brief catalogue raisonné:

1 The invocation of a syntactic neg-raising rule to deal with the unexpected occurrences of strict negative polarity items (NPIs) embedded under certain predicates (as in (4b)) is not sufficient, since such strict NPIs occur elsewhere in environments (as in (29a), (30a)) where neg-raising is not involved but conventionalized indirection is.
2 The positing of such a rule is conceptually undesirable and empirically incoherent, as Nuys himself concedes, given the implications of the data in (9).
3 A syntactic account of the NRP fails to related to other cases of ‘R-based’ negative (contradictory→contrary) strengthening, including litotes and affixal negation (see Horn 1989: ch. 5).
4 A syntactic account of the NRP fails to relate it to functionally parallel cases of short-circuited implicature or conventionalized/standardized indirection, including those explored by Searle (1975), Morgan (1978), and Bach and Harnish (1979).

Notes

1 Parenthesized numbers of this type refer to the corresponding example in the H&B text reprinted above. Examples introduced in this addendum are signalled as (A1), (A2), etc.
2 I also diverge from Nuys's favourable judgments (1990: 584–5) on the acceptability of examples like those in (i) and (ii):

   (i) It’s impossible that he has done it – has he?
   (ii) It’s certain that he hasn’t done it – has he?

reinforcing my scepticism toward the conclusions he derives therefrom.

References
