VENEETA SRIVASTAV DAYAL

SCOPE MARKING AS INDIRECT *WH*-DEPENDENCY*

In certain languages, scope-marking structures are used to express long-distance *wh*-dependencies along with or instead of the more familiar extraction structure. The existence of these two strategies raises an interesting question for the mapping between syntactic structure and semantic representation. Should apparent semantic equivalence be taken as a guide and syntactic parallelism posited at an abstract level of syntax? Or should the surface syntactic distinction between them be maintained and an alternative explanation sought for the similarity in meaning? In this paper I show that theoretical as well as empirical considerations argue against the first approach. I present a syntactic analysis of scope-marking structures in which the dependency between *wh*-expressions is indirect (in contrast to extraction structures which encode direct *wh*-dependencies). I draw attention to certain differences between scope marking and extraction structures which show that they are not really equivalent. The interpretive procedure given for indirect *wh*-dependencies derives the considerable similarity in meaning between the two structures while maintaining the necessary distinctions.

1. THE PHENOMENON OF SCOPE MARKING

1.1. *Extraction vs. Scope Marking*

It was first noted by van Riemsdijk (1983) that in certain dialects of German there exists an alternative to extraction for expressing long-distance *wh*-dependencies. The following examples illustrate the two strategies:

(1) a. **Mit wem**   glaubst du   dass Maria gesprochen hat?
     with whom think   you that Maria spoken   has

   b. **Was** glaubst du, **mit wem** Maria gesprochen hat?
     what think   you with whom Maria spoken   has

   'Who do you think Maria has spoken to?'

(1a) instantiates the standard extraction structure; the *wh*-expression in Spec of matrix CP *mit wem* signals that it is a direct question and identifies what the question is about. (1b) instantiates the so-called scope-marking structure; here, too, a *wh*-expression in Spec of matrix CP *was* signals that it is a direct question, but it is the *wh*-expression in Spec of embedded CP *mit wem* that provides semantic content. The sole function of *was* seems to be to extend the scope of *mit wem*, hence the name 'scope marker'.

Since van Riemsdijk's original observation, the existence of scope-marking structures has been attested in languages such as Romani (McDaniel 1989), Hindi (Davison 1984, Mahajan 1990, Srivastav 1989, 1991a), Bangla (Bayer 1990), and Iraqi Arabic (Wahba 1991).[1] The examples below illustrate scope-marking structures in these languages, in the order of mention:

(2) a. **Kas** o   Demiri mislinol Arifa dikhla?
       who does Demir think   Arifa saw

    b. **So** o   Demiri mislinol **kas**   Arifa dikhla?
       what does Demir think   whom Arifa saw

       'Who does Demir think Arifa saw?'

(3)   jaun **kyaa** soctaa hai meri **kis-se**   baat karegii?
       John what thinks   Mary who-with will-talk

       'Who does John think Mary will talk to?'

(4)   tumi **ki**   bhebe-cho **ke**   baaRi kore-che?
       you what think   who house built

       'Who do you think built the house?'

(5)   **sh**-tsawwarit Mona Ali raah **weyn**?
       what thought Mona Ali went where

       'Where did Mona think Ali went?'

Notice that Romani, like German, has overt *wh*-movement and allows for both the scope-marking and the extraction structure. Hindi and Bangla are different in that they have *wh*-in-situ. To the extent that overt *wh*-movement is possible, it is usually analyzed as an instance of topicalization

---

[1] Some of these languages have a structure in which the *wh*-expression in the embedded clause is repeated in the higher clause. As McDaniel (1989, 569, fn. 5) notes, such structures cannot be equated with scope marking. See also Szabolcsi and Zwarts (1993, 278—79) for a discussion of Hungarian scope marking.

or scrambling, not as a standard question-formation strategy. It is also well documented that these languages lack extraction out of finite complements at LF. Barring topicalization, therefore, scope marking is the only possible way to express long-distance *wh*-dependencies. Iraqi Arabic has optional *wh*-movement in the syntax but appears to disallow extraction. Scope marking is thus needed to question out of subordinate clauses. These differences notwithstanding, it is obvious that all these languages have an alternative to the more familiar extraction strategy for expressing long-distance *wh*-dependencies, an alternative that can be described in the same terms as (1b).

1.2. *Other Properties of Scope Marking*

In this subsection I present certain other properties of scope marking in order to familiarize the reader with the structure. The examples I give are from German and Hindi, respectively.

One characteristic of the scope-marking structure is that although the scope marker is always the same, there is no restriction on the type of *wh*-expression that can occur in the embedded clause.[2] So, in addition to the examples given earlier, we can have the following:

---

[2] A *NALS* reviewer points out that embedded yes/no questions are marginal in the German scope-marking structure in (i). The Hindi counterpart in (ii), however, is quite acceptable:

(i)   **Was** glaubst du ob die Maria mit dem Hans gesprochen hat?
         what believe you if the Maria with the Hans spoken   has

(ii)   tum **kyaa** socte ho ki   meri-ne haans-se   baat kiyaa yaa nahiiN?
         you what think,   that Mary   Hans-with talked   or   not

         'Do you think Mary talked to Hans?'

Possible answers to (ii) are of the form *I think Mary talked to Hans* or *I think Mary didn't talk to Hans*.

Anna Szabolcsi informs me that such questions are possible in Hungarian. As (iii) shows, however, these may not be genuine cases of embedding, since *hogy*, the Hungarian subordinator, is not licit in them. Also, the yes/no operator is optional, as in matrix clauses:

(iii)   **mit** gondolsz *(hogy) alszik (-e)?
         what think-2sg SUB   sleep-3sg-YES/NO

         'What do you think? Is she sleeping?'

I do not see a simple way of reconciling the German/Hungarian facts with the Hindi facts in any of the analyses I am aware of, so I leave this as an open problem here. The analysis I will develop in section 3 will yield the right result for (ii) but not for (i) and (iii). It should be noted, however, that German yes/no complements do not occur with *es* in the matrix, a correlation that is expected under the account to be developed here.

(6) a. **Was** glaubst du, **wo**   Maria getanzt hatte?
    what think   you where Maria danced

    'Where do you think Maria danced?'

   b. jaun **kyaa** soctaa hai meri **kahaaN** jaayegii?
    John what thinks   Mary where   will-go

    'Where does John think Mary will go?'

There is also no restriction on the number of *wh*-expressions that can occur in the embedded clause. The examples presented so far contained only one such expression, but in (7) we have examples with two embedded *wh*-expressions where the scope marker extends the scope of both:

(7) a. **Was** glaubst du, **wann** Hans an **welcher Universität**
    what think   you when Hans at which   University

    studiert hat?
    studied has

    'When do you think Hans studied at which university?'

   b. jaun **kyaa** soctaa hai **kaun kahaaN** jaayegaa?
    John what thinks   who where   will-go

    'Who does John think will go where?'

In fact, there can be as many *wh*-expressions in the embedded clause as the language allows in multiple *wh*-questions. The scope marker extends the scope of all of them.

Scope-marking structures can be used to express unbounded dependencies, as shown below:

(8) a. **Was** glaubst du, **was** Peter meint,   **mit wem** Maria
    what think   you what Peter believes with who Maria

    gesprochen hat?
    spoken   has

    'With who do you think Peter believes Maria has spoken?'

   b. jaun **kyaa** soctaa hai, anu **kyaa** kahegii, meri **kis-se**
    John what thinks   Anu what will-say Mary with-who

    baat karegii?
    will-talk

    'Who does John think Anu will say Mary will talk to?'

When there are multiple embeddings, however, each intermediate clause must have a scope marker:

(9) a. *****Was** glaubst du, dass Peter meint,   **mit wem** Maria
    what think   you that Peter believes with who Maria

    gesprochen hat?
    spoken   has

    'With who do you think Peter believes Maria has spoken?'

   b. *jaun **kyaa** soctaa hai, anu kahegii, meri **kis-se**
    John what thinks   Anu will-say Mary with-who

    baat karegii?
    will-talk

    'Who does John think Anu will say Mary will talk to?'

The distribution of scope marking is also interesting. Though the matrix verb must be able to take [−WH] complements, the actual complement must be [+WH]:

(10) a. *****Was** glaubst du, dass Maria mit Hans gesprochen hat?
    what think   you that Maria with Hans spoken   has

   b. **Was** glaubst du, **mit wem** Maria gesprochen hat?
    what think   you with whom Maria spoken   has

   c. *****Was** fragst du, **mit wem** Maria gesprochen hat?
    what ask   you with whom Maria spoken   has

(11) a. *jaun **kyaa** jaantaa hai meri ravi-se   baat karegii?
    John what knows   Mary Ravi with will-talk

   b. jaun **kyaa** jaantaa hai meri **kis-se**   baat karegii?
    John what knows   Mary who-with will-talk

   c. *jaun **kyaa** puuchhtaa hai meri **kis-se**   baat karegii?
    John what asks   Mary who-with will-talk

The (a) sentences above are ruled out because the complement clauses are not [+WH], where as the (c) sentences are ruled out because the matrix verb cannot take [−WH] complements.

Although the examples I have given are from German and Hindi only, the facts seem to hold in all languages that have a scope-marking structure. There is thus a clear sense of what the core syntactic and semantic

properties of this structure are, and it is expected that the analysis of scope marking in one language should carry over to scope marking in other languages as well.

1.3. *The Challenge*

The primary challenge for linguistic analysis posed by the existence of both scope marking and extraction in natural language, it seems to me, is to reconcile surface syntactic differences with semantic equivalence. In dealing with the paradigm in (1), for example, a choice has to be made between two possible approaches. One is to take meaning as a guide and posit syntactic parallelism at the level of representation that serves as input to interpretation; the other is to take syntactic differences seriously and find a way of interpreting the two distinct structures that makes them come out equivalent. A priori both approaches are valid, so one may be inclined to view the choice between them as largely a matter of personal taste. In point of fact, however, the choice of approach has nontrivial implications for the conception of how the syntactic and the semantic modules interact.
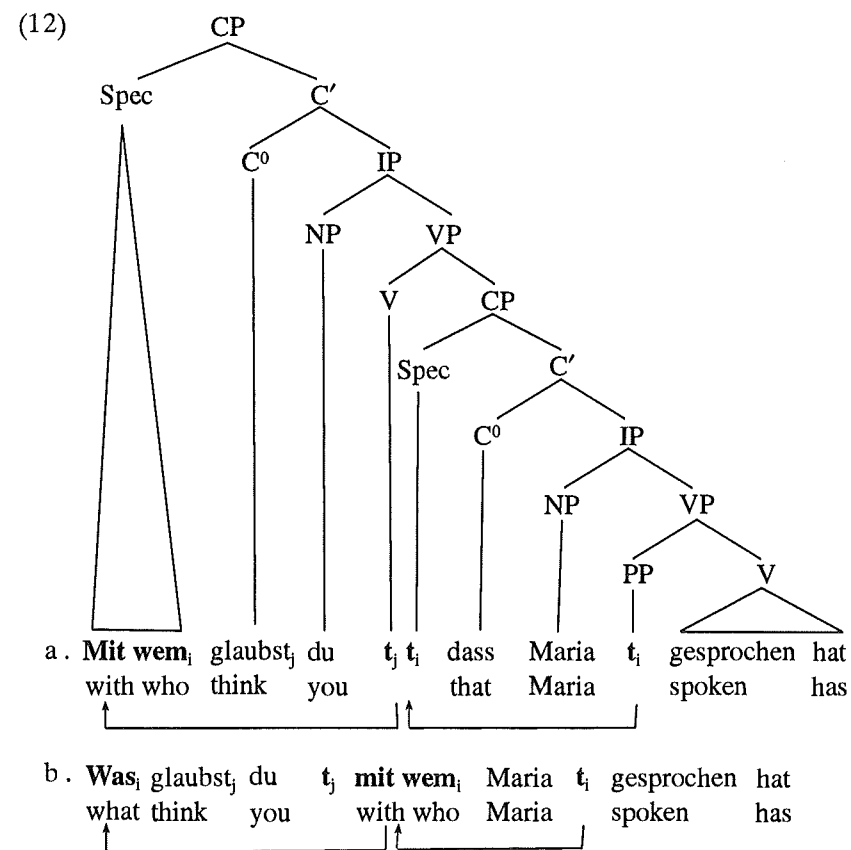
The view that there is a close connection between syntactic and semantic components is fairly standard in current linguistic theory. The principle of Full Interpretation (Chomsky 1986a), for example, enjoins syntactic analyses to be semantically accountable, while the principle of Compositionality (Frege 1960) requires semantic procedures to be defined systematically on syntactic structures. Working within the contraints imposed by these principles, however, is not easy, and at times one comes across data that seem to call their validity into question. Scope-marking structures appear to constitute just such a case. However, I will propose an analysis for them below that achieves full interpretation in a compositional manner, and I will also show that such an analysis makes the right empirical generalizations. In this way, the present paper provides support for Full Interpretation and Compositionality as sound methodological principles which can be used as objective criteria of evaluation in choosing between particular analyses. It also endorses the view of the syntax-semantics interface that these principles imply.

2. THE DIRECT DEPENDENCY APPROACH

2.1. *Earlier Analyses*

The practice of analyzing questions in terms of the answers they allow is standard, at least since Belnap and Steel (1963). On this approach, if two

question types allow the same kinds of answers, it seems reasonable to hypothesize that they must have the same structures; it is not surprising, therefore, that scope markers have typically been treated as variants of extraction structures. Thus McDaniel (1989), following van Riemsdijk (1983), claims that the scope marker is an expletive *wh*-expression which is base generated in Spec of matrix CP and forms a chain with the *wh*-expression in the embedded CP. The representations of (1a–b), under her analysis, are as in (12a–b):

(12)

a. **Mit wem**$_i$ glaubst$_j$ du   t$_j$ t$_i$ dass Maria   t$_i$ gesprochen hat
   with who   think   you   that Maria   spoken   has

b. **Was**$_i$ glaubst$_j$ du   t$_j$ **mit wem**$_i$ Maria t$_i$ gesprochen hat
   what think   you   with who   Maria   spoken   has

The only difference between the two structures is the source of the *wh*-dependency. In the extraction structure (12a) it results from movement, but in the scope-marking structure (12b) from coindexing. In either case, a direct *wh*-dependency is established between the position where theta role is assigned (i.e. the embedded argument position) and the position where scope is fixed (i.e. the matrix Spec position). Scope markers, in this view

then, are just a special type of *wh*-operator employed by some languages; the relationships they enter into are the ones standard for such operators. McDaniel provides the following formulation of the Chain Condition, designed to cover scope-marking as well as extraction structures:

(13)   A Chain C = $(a_1, a_2, \ldots a_n)$ is a *wh*-chain iff:

    a. $\forall a_i, 1 \leqslant i < n$, a_i locally A′-binds a_{i+1},

    b. $\forall a_i, 1 \leqslant i < n$, a_i is a *wh*-element,

    c. a_n is a variable in IP-internal position, and

    d. for any scope marker a_i, $1 \leqslant i < n$, $(a_{i+1}, \ldots a_{n-1})$ contains a true *wh*-phrase.

This approach goes some way in accounting for the characteristics of scope-marking structures noted in section 1. For example, it accounts for the fact that there is no restriction on the type of *wh*-expression in the embedded clause, that scope markers can encode unbounded dependencies, and that scope marking only occurs with matrix verbs that can take [–WH] complements even though the actual complement must have a *wh*-expression. I will not go into the explanations here but refer the reader to McDaniel (1989) for a fuller discussion. I should also emphasize that though I have presented McDaniel's analysis here for concreteness, the view that scope marking is to be treated in the same terms as extraction is also held by van Riemsdijk (1983), Davison (1984), Mahajan (1990), and Wahba (1991).[3] With some modifications, the discussion below holds for all these versions of the direct dependency approach.

As will be obvious from this summary, a consequence of the direct dependency approach is that the syntactic distinction between scope-marking and extraction structures is reduced to a status of relative triviality. Given that the two structures appear to be equivalent this appears a welcome result. Though the direct dependency analyses that have been proposed do not give an explicit semantics for scope marking, the assumption seems to be that whatever theory one adopts for interpreting extraction structures will also apply straightforwardly to scope-marking structures, given an appropriate interpretation of coindexing in A-bar chains.

---

[3] Bayer (1990) gives a parsing account of scope marking in which the matrix *wh*-phrase primes the parser to look for a *wh*-phrase in the structure that follows. I do not feel completely confident in classifying his account with the direct dependency approach, but I do believe that the empirical and theoretical consequences are similar.

2.2. *Negation in Scope Marking*

Rizzi (1992) notes a potential problem for the view that scope-marking structures are a simple variant of extraction structures. As he shows, the equivalence of scope-marking and extraction structures breaks down under negation, as shown by (14):

(14) a. **Mit wem** glaubst du nicht, dass Maria gesprochen hat?
       with whom think   you not   that Maria spoken   has

     b. ***Was** glaubst du nicht, **mit wem** Maria gesprochen hat?
        what think   you not   with whom Maria spoken   has

    'Who don't you think Maria has spoken to?'

While such facts pose a problem for the direct dependency approach, they do not provide definitive evidence against it. As Rizzi points out, it is well known that negation can block certain instances of *wh*-movement. The negative island effect is illustrated in (15), where *wh*-movement of adjuncts but not arguments is affected by negation:

(15) a. Who do/don't you think Mary will hire?
     b. How do/*don't you think Mary will behave?

Rizzi explains this contrast in terms of relativized minimality. Arguments, he suggests, can carry a referential index and bind their trace; thus both the affirmative and the negative versions of (15a) are licit. Adjuncts, on the other hand, cannot carry a referential index and need to antecedent govern their trace. The presence of a potential A-bar negative operator blocks the relevant government relation in (15b).

Following McDaniel's analysis of scope-marking structures, Rizzi extends the explanation for the negative island effect to the paradigm in (14). In the standard extraction structure in (14a), *mit wem* is an operator that carries a referential index and binds its trace in the lower Spec. Matrix negation has no blocking effect. *Was* in (14b), on the other hand, is an A-bar expletive which cannot carry a referential index. It must therefore form a chain with *mit wem*, which it must antecedent govern. Matrix negation interferes with that relation, hence the ungrammaticality.

The relativized minimality account of negative islands is not unproblematic, and I will return to it in section 4. Nevertheless, Rizzi's basic idea — that the problem of negation in scope marking structures can be reduced to the more general problem of negative islands — is worth considering. At least it provides a way of dealing with the negation facts within the direct dependency approach.

2.3. *Problems with the Direct Dependency Approach*

Appealing though an approach may seem which takes semantic equiva-
lence as a guide in positing syntactic parallelism between scope-marking
and extraction structures, there are some serious theoretical reasons for
questioning its validity.

A major problem for the direct dependency approach is posed by
embedded questions with more than one *wh*-expression, as in (7a),
repeated below:

(7) a. **Was** glaubst du, **wann** Hans an **welcher Universität**
       what think   you when Hans at   which   university

       studiert hat?
       studied has

       'When do you think Hans studied at which university?'

The problem in establishing a direct dependency here between the scope
marker and the embedded *wh*-expressions is obvious: the *wh*-chain would
have one head and two tails. In his original analysis (in which he did not
consider questions with embedded multiple *wh*-questions) van Riemsdijk
concluded that scope markers could be treated like other operators of the
language and that their distribution was uniformly constrained by the
ECP. Given his formulation of the ECP, this ensured a one-to-one
correspondence between scope markers and embedded *wh*-expressions.
McDaniel's analysis essentially recasts van Riemsdijk's insight in terms
of Chomsky (1986b). Thus questions like (7a) are as problematic in
McDaniel's approach as in van Riemsdijk's.

McDaniel is aware of the problem and offers a solution. She argues that
the two embedded *wh*-expressions undergo absorption, in the sense of
Higginbotham and May (1981), and that the index of the absorbed
expression is copied onto the scope marker. Note, however, that this
requires the absorbed *wh*-term $i+j$ in the embedded Spec to be an
intermediate link in the chain, whereas under standard assumptions,
absorption creates an operator and operators can only head a chain. Thus
McDaniel is forced to suggest that the absorption in the complement of
scope-marking structures differs from the standard type of absorption in
that it does not create an operator. This, of course, is quite ad hoc and
robs the notion of absorption of all content.

A second nontrivial problem with the direct dependency account is that
it does not allow for a compositional mapping from LF representations to

meaning. Since the scope marker has no semantic content and there are
no restrictions on the type and number of *wh*-expressions in the em-
bedded clause, the kind of quantification involved in a particular scope-
marking structure cannot be determined at the point where the scope
marker has to be interpreted. Using a semantics for questions such as
that of Hamblin (1973), for example, the denotation of (1a—b) would be
(16) — the set of possible answers allowed by these questions:

(16) $\lambda p \exists x[\text{person}'(x) \,\&\, p = {}^\wedge\text{think}'(\text{you}, {}^\wedge\text{talk-to}'(m, x))]$

In this way of interpreting questions, *wh*-expressions are existential
quantifiers whose restriction is provided by a common noun inside the
wh-expression whenever such a noun is present. Thus, *mit wem* in the
matrix Spec of (1a), under an LF representation like (12a), is an existential
quantifier over individuals. Turning to (1b), under an LF representation
like (12b), we see that in order to derive something like (16), the scope-
marking *wh*-expression should not be permitted to contribute directly to
meaning by providing an existential quantifier. Alternatively, we could
concede that it provides an existential quantifier, but say that the restric-
tion on that quantifier has to come from the *wh*-expression in the lower
Spec. That is, the meaning of the scope marker itself would be underdeter-
mined, while at the same time, the meaning of the embedded *wh*-element
would be kept in store, so to say, till the scope marker is reached, in order
to fill in the missing element in the quantification. This is particularly
obvious when one considers cases where the embedded *wh*-element has
semantically relevant material, for example NPs like *which girl* or *which
girl's book*.

Clearly, the assumption that the relevant coindexing guarantees a
systematic mapping to semantics is not well founded. It cannot be imple-
mented without compromising compositionality. A way of maintaining
compositionality would be to posit successive operations to replace the
scope marker with the embedded *wh*-expression(s). The only principled
way of doing this would be in terms of expletive replacement at LF
(Chomsky 1986a). Note, however, that Rizzi's explanation for the negation
facts would then be lost, since the scope-marking and the extraction struc-
ture become isomorphic at LF. The prediction would be that embedded
wh-expressions that are arguments, but not those that are adjuncts, should
be able to cross over negation and replace the expletive. As we have seen
in (14b), this is not the case.

Finally, the direct dependency approach does not extend to in-situ

languages like Hindi and Bangla in a perspicuous manner.[4] In Hindi, for example, the scope marker *kyaa* occurs in the preverbal position. This is the canonical position for direct objects, as shown by (17):

(17)   anu   ravi-ko   kitaab degii.
        Anu   Ravi-DAT book   will-give

        'Anu will give Ravi a book.'

In order to establish a direct dependency between scope marker and embedded *wh*-expression in Hindi, then, the scope marker would crucially have to move at LF to the matrix Spec, its link with the object position erased, and then a new chain formed with the *wh*-element in the embedded Spec position, as illustrated below:

(18)
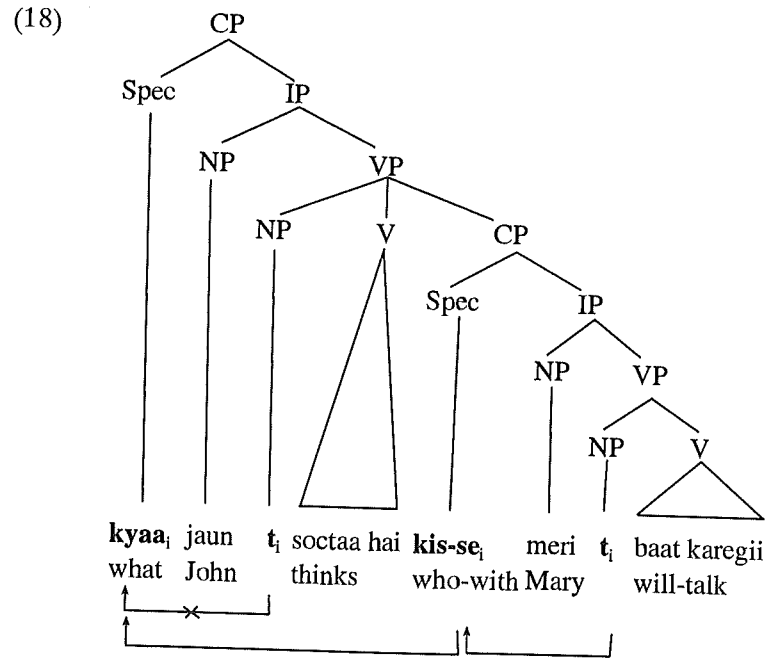
The moves required are clearly non-standard and cannot be independently motivated. One way of deriving the same effect is given in Mahajan

---

[4] McDaniel's account is only concerned with cases of overt *wh*-movement; it is not clear to me whether she would even predict the existence of scope marking in in-situ languages and if so, what the predictions would be. See Srivastav (1991a) for an attempt to tease out these issues.

(1990). Details aside, Mahajan suggests that the finite complement adjoins to the scope marker, as an instance of expletive replacement. The wide-scope reading of the embedded *wh*-expression follows from this structure. The problem with this analysis is that it makes scope-marking structures isomorphic to extraction structures at LF. Once again, the explanation for the negation facts is lost.

So it seems clear that the direct dependency approach is wrong, or at least flawed, in that additional stipulations are needed to make it work. Obviously, an alternative approach to the scope-marking phenomenon that manages to avoid the problems inherent in the direct dependency approach is to be preferred. In the next section I explore such an alternative.

3. THE INDIRECT DEPENDENCY APPROACH

3.1. *The Syntax of Indirect Dependency*

We saw in the previous section what is involved in taking similarity of meaning as a guide in syntactically analyzing scope-marking structures along the lines of extraction structures. Let us now try the opposite tack. If we take the surface syntax of scope-marking structures to reflect their representation at LF, the task would be to show that standard procedures of interpretation can be applied in such a way that a meaning similar to that of extraction structures is obtained.

I will base my syntactic analysis of scope-marking structures on Hindi and then generalize it to other languages. In Srivastav (1991a) I proposed that Hindi is an SOV language which does not allow CPs in argument position due to the Case Resistance Principle (Stowell 1981). Thus finite complementation in Hindi yields a SO_iVX_i construction where the finite complement is a syntactic adjunct coindexed with the matrix preverbal direct object position. This position may be phonetically null or be filled by the expletive pronoun *yeh*, as illustrated by (19a). Following Huang (1982), I assumed movement of in-situ *wh*-phrases to Spec positions and posited the LF representation for this sentence given in (19b):[5]

---

[5] Movement of *wh*-expressions out of the embedded clause is prevented by an interaction of subjacency and ECP, resulting in an unambiguous indirect-question interpretation. See Srivastav (1991b) for the basic facts and the essentials of the analysis given in Srivastav (1991a).

(19) a. jaun (**yeh**) jaantaa hai ki   meri   **kis-se**   baat karegii.
    John this   knows   that Mary who-with will-talk

    'John knows (this) who Mary will talk to.'

b.
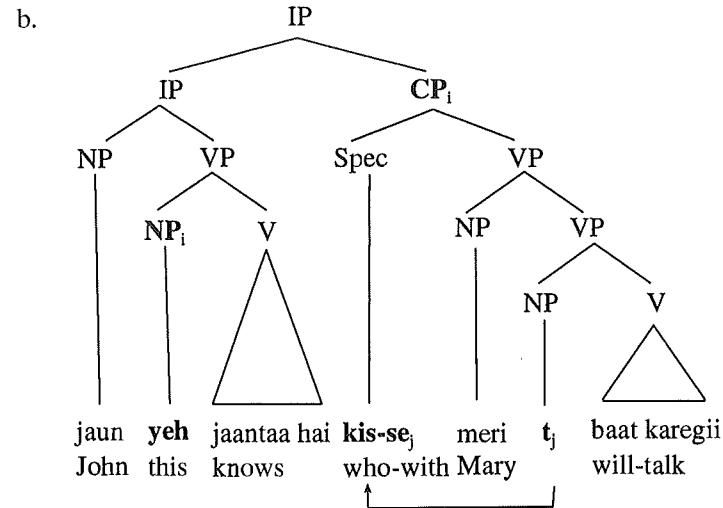
On analogy with structures like (19), I argued that Hindi scope markers are in object position at S-structure and are coindexed with the complement CP, which is adjoined at the IP level. At LF, the *wh*-expressions move to Spec, resulting in the following structure for (3):

(20)
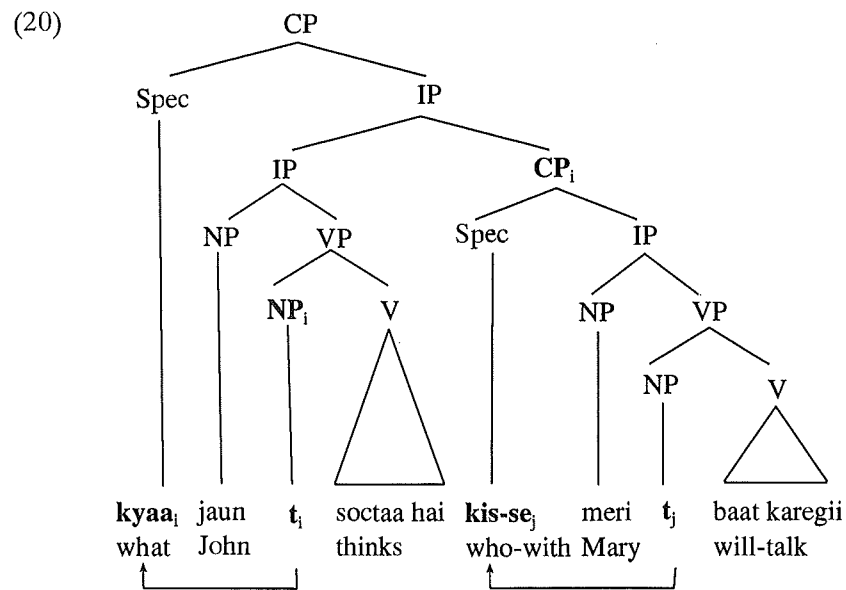
In this structure there is no direct dependency between the scope marker and the actual *wh*-phrase. There are, instead, two local *wh*-dependencies. The effect of long-distance dependency arises from the fact that coindexation of the dominating nodes links up the two local dependencies.

I should mention that Davison (1984) and Mahajan (1990) hold similar views on finite complementation in Hindi and posit similar S-structure representations for scope-marking structures. The crucial difference is that I maintain the syntactic distinction between scope-marking and extraction structures at all levels of syntactic representation, whereas they eliminate it at the level which feeds into the semantics.

The analysis under discussion differs fundamentally from those of van Riemsdijk (1983), McDaniel (1989), and Wahba (1991) in the characterization of the scope marker as an expletive generated in argument position rather than as an expletive generated in Spec of CP.[6] It is worth pointing out that there is no independent evidence given in the literature for generating the scope marker in Spec position except its clause-initial position in German, Romani, and Iraqi Arabic. In contrast, though the primary motivation for generating the scope marker in argument position comes from Hindi and Bangla word order, I believe this idea can extend to languages in which the scope marker is in clause-initial position.

The analysis of German phrase structure in Cardinaletti (1990, 71—100), for example, is quite compatible with the present account. She argues that finite CPs in German may either be complements of V or IP adjuncts coindexed with the expletive *es*, which is in complement-of-V position. Applying this to scope-marking, one could say that the scope marker *was* originates in the same position as *es* (i.e. as complement of V) and is coindexed with the adjoined finite clause.[7] Though surface word order does not give any evidence of this, the following examples, fashioned after those provided by a *NALS* reviewer, are telling:

(21) a. **Mit wem**   glaubt jeder Student$_i$, dass er$_i$ gesprochen hat?
    with whom thinks every student   that he spoken   has

---

[6] Wahba's terminology is somewhat different in that the scope marker is referred to as a Quantifier Phrase, but I think the idea is the same.

[7] Josef Bayer points out (personal communication) that the ungrammaticality of the following is unexpected under the view that the scope marker is the complement of V:

(i)   *Wer hat was$_i$ gedacht, wen$_i$ wir anrufen sollten?
      who has what thought   whom we call-up should

   'Who thinks what, who should we call up?'

At this point I have no explanation for this fact. It is worth noting, however, that the alternative order with *was* in initial position is not grammatical either.

b. **Was** glaubt jeder Student$_i$, **mit wem** er$_{*i}$ gesprochen hat?
what thinks every student with whom he spoken   has

'With whom does every student think he has spoken?'

Both (21a) and (21b) are grammatical, but the bound variable reading for the pronoun in the subordinate clause is only available in (21a). This is easily explained, under the present proposal, since the finite complement is c-commanded by the matrix subject in the extraction structure but not in the scope-marking structure.

The primary difference between German and Hindi, then, is that German has overt *wh*-movement so that the type of *wh*-dependency represented in (20) is present at S-structure and LF, while Hindi has *wh*-in-situ and only displays this dependency at LF. I have not investigated the phrase structure of all the languages known to have scope marking, but I take it as encouraging that there seems to be a plausible way of extending the analysis to German.

Before concluding this discussion I would like to minimally modify the analysis I have just presented. Following Bittner (in press, 38—42), as well as the suggestions of a *NALS* reviewer, I will allow the subordinate clause in a scope-marking structure to be adjoined to the matrix CP. Note that though this move is not standard it does not violate any constraints. As McCloskey (1992) points out, the prohibition against adjunction to arguments in Chomsky (1986b) does not translate into a general prohibition against adjunction to CP. Matrix CPs, for example, should allow adjunction because they are non-arguments. If this is right, it is expected that adjunction in scope-marking structures can be at the level of CP as well as IP. (22), then, is a permissible LF structure for (3):

(22)
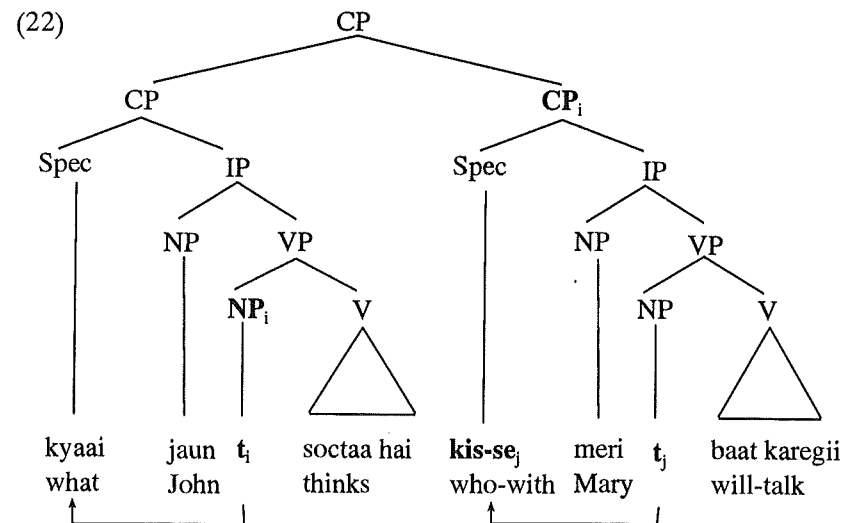
I will use CP-adjoined structures like (22) to illustrate the interpretation procedure for scope marking, though IP-adjoined structures like (20) are also well-formed (see footnote 10 for discussion regarding the interpretation of structures like (20)).

To wrap up this point, the basic syntactic claim of the indirect dependency approach is that the scope marker is generated in argument position and is coindexed with a CP in adjoined position. Depending on the language, it moves to Spec of CP at S-structure or LF. Crucially, at the level of representation which is the input to semantics, the so-called long-distance dependency of scope-marking structures in actuality consists of two local *wh*-dependencies linked via coindexation of the dominating nodes.

3.2. *Semantic Analysis*

Since the syntax I am adopting maintains the difference between scope marking and extraction at all levels of syntactic representation, the question that immediately arises is how scope marking can be interpreted in a way that makes it equivalent to extraction.[8]

Let us assume the semantics for questions in Hamblin (1973), mentioned earlier, which takes the denotation of a question to be the set of propositions that constitute possible answers to it. Thus the denotation of (23a) is (23b), resulting in sets such as (23c):

(23) a. Who does John think Mary will talk to?
   b. $\lambda p \, \exists x \, [p = {}^\wedge \text{think}'(j, {}^\wedge \text{will-talk}'(m, x))]$
   c. $\{{}^\wedge \text{John thinks Mary will talk to Bill, } {}^\wedge \text{John thinks Mary will talk to Sue} \ldots\}$

In this approach to the meaning of a question, the *wh*-expression is interpreted as an existential quantifier. In the case of *who*, quantification is over (animate) individuals.

Turning to scope-marking structures, the first issue to be settled is what the scope marker quantifies over. Examining other contexts in which the Hindi scope marker *kyaa* occurs, we find that it can quantify over propositions, as shown by possible answers to (24a):

---

[8] In Srivastav (1991a) I had treated the scope marker as a pleonastic polyadic quantifier that binds the complement CP and, through Spec-head agreement, the *wh*-expressions in the embedded Spec. However, that analysis cannot in any straightforward way handle the negation facts pointed out by Rizzi (1991). It is also problematic from the point of view of compositionality, as was pointed out to me by Polly Jacobson (personal communication).

(24) a. jaun **kyaa** soctaa hai?
John what thinks

'What does John think?'

b. john soctaa hai ki vo tez hai.
John thinks that he smart is

'John thinks that he is smart.'

Applying Hamblin-type semantics to (24a) gives us (25a) as its denotation; the resulting sets are as in (25b):

(25) a. $\lambda p \exists q[p = \,^\wedge think'(j, q)]$
b. $\{\,^\wedge John\ thinks\ that\ he\ is\ smart,\ ^\wedge John\ thinks\ that\ Mary\ will\ talk$
to Bill, $^\wedge John\ thinks\ that\ Mary\ will\ talk\ to\ Sue \ldots\}$

Now, this is a completely standard analysis of questions like (24a). The point I want to draw attention to is that the matrix clause of the scope-marking structure in (22) is identical to it. The null hypothesis, then, would be that they have the same denotation. Let us explore this hypothesis a bit.

Let us assume that the matrix clause of (22) is interpreted as (25a). We know, of course, that (22) should not allow the first proposition in (25b), since that proposition does not deal with John's knowledge about who Mary will talk to; clearly something more needs to be done. Consider now the complement clause. This too is a question, so it is going to denote a set of propositions. In this case it will denote (26a), resulting in sets like (26b):

(26) a. $\lambda p' \exists x[p' = \,^\wedge will\text{-}talk'(m, x)]$
b. $\{Mary\ will\ talk\ to\ Bill,\ Mary\ will\ talk\ to\ Sue,\ Mary\ will\ talk\ to$
Jane...$\}$

Intuitively, what we want to do is combine the denotations of the two questions in such a way that we end up with (27a). This will give us sets such as (27b) as an answer for (22):

(27) a. $\lambda p \exists q[\lambda p' \exists x[p' = \,^\wedge will\text{-}talk'(m, x)]\ (q)\ \&\ p = \,^\wedge think'(j, q)]$
$\Rightarrow \lambda p \exists q[\exists x[q = \,^\wedge will\text{-}talk'(m, x)]\ \&\ p = \,^\wedge think'(j, q)]$
b. $\{John\ thinks\ that\ Mary\ will\ talk\ to\ Bill,\ John\ thinks\ that\ Mary$
will talk to Sue...$\}$

(27a) says that the scope-marking structure denotes a set of propositions p such that there exists a proposition q such that for some individual x, q is $^\wedge Mary\ will\ talk\ to\ x$ and p is $^\wedge John\ thinks\ q$. The first condition in the formula, we see, excludes all propositions that are not possible answers to the complement set. The set of propositions that the scope-marking structure allows, then, is the same as the set of propositions that the corresponding extraction structure would allow.

Let us now try to achieve this in a principled manner. In order to do so, we need to step back a bit and focus afresh on simple questions involving quantification over propositional variables. The English question corresponding to (24a), *What does John think?* is adequate for sharpening our intuitions. This question asks for the set of propositions that John stands in the *think'* relation to, but it doesn't really seem as if the questioner is interested in every proposition that is part of John's thoughts; I'm not even sure that it is possible to list the full set. Of course, in a given context, the question is generally understood to be something like *What does John think about X?* where *X* provides a reasonable delimitation for the set of propositions under consideration. This would make the question computable. One might think of this delimitation as the topic of the question.

Building this kind of delimitation into the denotation of the question is simple. We treat the *wh*-expression *kyaa* in the standard way, as an existential quantifier over propositional variables. Since natural language quantification is known to be restricted, one expects there to be a restriction in this case also. We may therefore posit a covert restriction on the variable q bound by the existential, say T (for topic), whose type is $\langle\langle s, t\rangle, t\rangle$, i.e. a set of propositions. Thus the denotation of such questions contains a condition $T(q)$. This would be fully parallel to the restriction to the set of animate individuals in the case of questions with *who*. The difference between questions over propositions and those over individuals is that whereas the former have no basic lexical item that can give overt realization to T, the latter can use common nouns to spell out the restriction, as in *which person* or *which girl*.

Now let us carry this intuition over to the scope-marking structure in (3) and its syntactic analysis in (22). Borrowing somewhat eclectically from Bittner (in press, 38–42) and from the comments of a *NALS* reviewer, I show here one way of compositionally incorporating Hamblin's treatment of questions into the syntactic framework of Chomsky (1986b):

(28) a.

$$CP \; \lambda T_i[\lambda p \exists q[T_i(q) \& [p = {}^\wedge think'(j, q)]] \; (\lambda p' \exists x[p' = {}^\wedge will\text{-}talk'(m, x)])$$
$$\Rightarrow \lambda p \exists q[\lambda p' \exists x[p' = {}^\wedge will\text{-}talk'(m, x)](q) \& [p = {}^\wedge think'(j, q)]]$$
$$\Rightarrow \lambda p \exists q[\exists x[q = {}^\wedge will\text{-}talk'(m, x)] \& [p = {}^\wedge think'(j, q)]]$$

CP
$$\lambda Q \exists q[T_i(q) \& Q(q)](\lambda q[p = {}^\wedge think'(j, q)])$$
$$\Rightarrow \exists q[T_i(q) \& \lambda q[p = {}^\wedge think'(j, q)](q)]$$
$$\Rightarrow \exists q[T_i(q) \& [p = {}^\wedge think'(j, q)]]$$

$CP_i \lambda p' \exists x[p' = {}^\wedge will\text{-}talk'(m, x)]$

Spec
$$\lambda Q \exists q[T_i(q) \& Q(q)]$$

C' $\lambda p'[p = p']$ $({}^\wedge think'(j, q))$
$$\Rightarrow p = {}^\wedge think'(j, q)$$

$C^0 \lambda p'[p = p']$ IP $think'(j, q)$

kyaa$_i$ e   **jaun t$_i$ soctaa hai**   **kis-se$_j$**   **meri t$_j$ baat karegii**
what   John   thinks   who-with Mary   will-talk
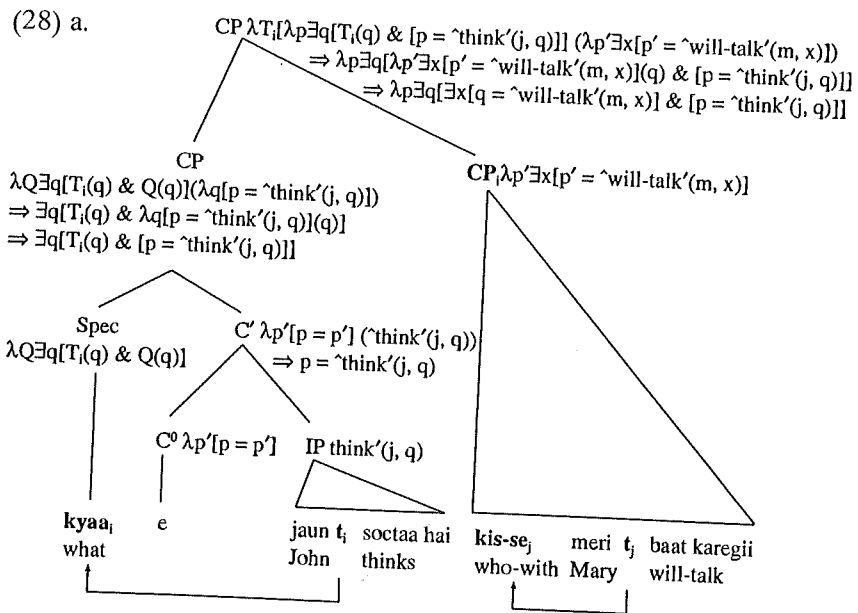
b. {${}^\wedge$John thinks that Mary will talk to Bill, ${}^\wedge$John thinks that Mary will talk to Sue}

Let us consider the interpretation of the matrix clause first. I take the denotation of *kyaa*, the scope marker, to be an existential *wh*-quantifier over propositions, $\lambda Q \exists q[T_i(q) \& Q(q)]$. The variable T_i represents the covert restriction on q, the variable bound by the scope marker. The trace of the *wh*-phrase is interpreted as a variable and the IP as the open sentence *think' (j, q)*. The [+WH] C^0 introduces the variable p, which is to be identified with the IP and denotes a function from propositions to propositions $\lambda p'[p = p']$. Applying this function to the IP denotation yields $p = {}^\wedge think'(j, q)$ as the meaning of C'. Once the variable q is abstracted over, it can combine with the *wh*-operator via functional application. Finally, the free variable p is abstracted over, resulting in the question denotation $\lambda p \exists q[T_i(q) \& p = {}^\wedge think'(j, q)]$. The derivation of the subordinate clause $\lambda p' \exists x[p' = {}^\wedge will\text{-}talk'(m, x)]$ is quite straightforward, so I have skipped over the intermediate steps here.

Turning now to the highest CP node, we see that there is a simple way for the subordinate clause meaning to combine with the main clause meaning. Note that the variable T_i in the *wh*-expression of the matrix Spec, with which the subordinate clause is coindexed, is free at this level. It is therefore possible to abstract over T_i, giving us a function from question denotations to question denotations $\lambda T_i[\lambda p \exists q[T_i(q) \& p =$

${}^\wedge think' (j, q)]]$. The subordinate clause, being a question and therefore of the right semantic type, can combine with it through functional application, giving us $\lambda p \exists q[\exists x[q = {}^\wedge will\text{-}talk' (m, x) \& p = {}^\wedge think' (j, q)]]$, the interpretation we had aimed for.[9]

To verify that this works, consider the proposition ${}^\wedge$*John thinks that he is smart*. Here q would be *that he is smart*. This proposition will not be included because there is no individual x such that q is identical to ${}^\wedge$*Mary will talk to x*. But when we consider the proposition ${}^\wedge$*John thinks that Mary will talk to Sue*, we have the value of q set at *that Mary will talk to Sue* and, indeed, there is an x (namely Sue) such that q is identical to *that Mary will talk to x*. The interpretive procedure includes all and only the propositions that it should.[10]

To summarize this discussion of the semantics of scope marking: I have proposed that there is always a covert restriction in the case of questions involving quantification over propositions. This restriction is a variable over sets of propositions and provides the means by which the complement clause denotation can combine with the main clause denotation. The procedure outlined here takes as input a representation that is close to surface syntax and interprets it compositionally, assigning a determinate

---

[9] What we have, in effect, is an analysis of the scope-marking structure as *What does John think about who Mary will talk to?*. Interestingly, McDaniel (1989, 570, fn. 6) considers this possibility but dismisses it as counterintuitive. She also points out a construction parallel to scope marking in Romani relative clauses to which her account for questions is intended to extend. I have not had the opportunity to test whether the present theory would extend to the kinds of relative clauses McDaniel mentions, but I expect that it would given an appropriate semantics for relative clauses.

[10] Recall that scope marking can also involve adjunction at the IP level, as in (20). In an earlier version of this paper I had the variable T be part of the trace of the scope marker. However, the resulting structure does not translate into the formula in (28a) when interpreted compositionally, as I had assumed, but rather translates as (i) below. I am indebted to Maria Bittner and two *NALS* reviewers for pointing this out:

(i) $\lambda p \exists q[p = {}^\wedge[\exists x[q = {}^\wedge will\text{-}talk'(m, x)] \& think'(j, q)]]$

This clearly does not denote the same set of propositions that (28a) denotes. Irene Heim and Angelika Kratzer point out (personal communication) that this denotation could be considered acceptable if one appealed to the pragmatics of question-answer dialogues. The key observation is that the condition ${}^\wedge \exists x[q = {}^\wedge will\text{-}talk'(m, x)]$ in (i) is not a contingent one since for any given value of q it is either true in all worlds or false in all worlds. In the former case, this proposition would have no effect; in the latter it would admit the contradictory proposition into the set as well. One could argue that the contradictory proposition is simply ignored and therefore irrelevant. Alternatively, one might argue that although IP-adjoined structures are syntactically well-formed, they are semantically deviant. Scope-marking structures are attested in languages because there exists one well-formed and fully interpretable LF, namely the one with CP adjunction.

meaning to every lexical item and every phrasal category in the process. The indirect dependency approach to scope marking thus achieves semantic equivalence with extraction structures without eliminating the syntactic distinction between the two.

3.3. *Theoretical Payoffs*

The main theoretical advantage of the indirect dependency approach, as I see it, is its extreme simplicity and generality. It treats each *wh*-expression in Spec position at LF as an operator, which is standard practice in the analysis of ordinary questions. It creates the effect of long-distance dependency by interpreting the complement as a restriction on the variable in the higher clause. I would like to show now that these basic features of the analysis derive all of the properties noted in section 1.2 without further stipulations.

Consider the fact that any type or number of embedded *wh*-expression(s) are allowed in a scope-marking structure. Below I repeat some examples from section 1.2 and provide the translations that would be given under the indirect dependency approach:

(6) b.   jaun **kyaa** soctaa hai meri **kahaaN** jaayegii?
         John what thinks   Mary where   will-go

         'Where does John think Mary will go?'
         $\lambda p \exists q [\exists x [place'(x) \& q = \text{^will-go-to}'(m, x)] \& p = \text{^think}'(j, q)]$

(7) b.   jaun **kyaa** soctaa hai **kaun kahaaN** jaayegaa?
         John what thinks   who   where   will-go

         'Who does John think will go where?'
         $\lambda p \exists q [\exists x \exists y [person'(x) \& place'(y) \& q = \text{^will-go-to}'(x, y)] \& p = \text{^think}'(j, q)]$

As should be obvious, the indirect dependency approach has no problems with compositionality since each *wh*-expression is fully interpreted in its LF position. If there are two such expressions, as in the case of (7b), the normal procedure for interpreting multiple *wh*-questions can be followed. For example, if absorption is the assumed mechanism, it would yield a binary operator here as elsewhere. Thus interpretation of multiple embedded questions in scope-marking structures is completely straightforward. Recall from the discussion in section 2.3 that such multiple *wh*-structures were particularly problematic for the direct dependency approach, which had to resort to a special definition of absorption.

Since a scope-marking structure is itself a question, it follows that it can form the restriction on a propositional variable in a higher clause, creating the effect of unbounded *wh*-dependencies. The coindexing between each complement clause and the object position in the clause above it eliminates from the higher-clause denotation the propositions that do not belong in its complement:

(8) b.   jaun **kyaa** soctaa hai, anu **kyaa** kahegii, **meri kis-se**
         John what thinks   Anu what will-say **Mary** with-who

         baat karegii?
         will-talk

         'Who does John think Anu will say Mary will talk to?'
         $\lambda p \exists q [\exists r [\exists x [r = \text{^will-talk}'(m, x)] \& q = \text{^will-say}'(a, r)] \& p = \text{^think}'(j, q)]$

Let us demonstrate informally how this works. Suppose, for example, that the unrestricted matrix clause question denotes sets like $X = \{\text{^}John$ *thinks he is smart,* ^ *John thinks Mary will talk to Bill,* ^*John thinks Anu will say Mary will talk to Sue* . . . $\}$, the unrestricted intermediate clause denotes something like $Y = \{Anu$ *will say Mary will talk to Sue,* ^*Anu will say Mary will talk to Bill,* ^*Anu will say Bill is smart* . . .$\}$ and the lowest clause something like $Z = \{\text{^} Mary$ *will talk to Bill,* ^ *Mary will talk to Sue,* ^ *Mary will talk to Harry* . . .$\}$. As the interpretation proceeds and the lowest CP combines with the intermediate clause, Z fills in for the restriction on Y and we get $(Y \cap Z) = \{Anu$ *will say Mary will talk to Sue, Anu will say Mary will talk to Bill*$\}$. When this conjunct combines with the matrix, $(Y \cap Z)$ will fill in for the restriction on X and we get $(X \cap (Y \cap Z)) = \{\text{^}John$ *thinks Anu will say Mary will talk to Sue* . . .$\}$. Of course there are many more propositions in each set, but the schema should make it clear how certain classes of propositions are excluded as each clause provides the restriction for the one immediately above it.

Regarding multiple embeddings, every intermediate clause must have a scope marker in order to be interpreted as a question. Thus, (9b) is ungrammatical:

(9) b.   *jaun **kyaa** soctaa hai, anu kahegii, meri **kis-se**
          John what thinks   Anu will-say Mary with-who

          baat karegii?
          will-talk

          'Who does John think Anu will say Mary will talk to?'

The explanation for this in the present analysis is that the complement of a scope-marking structure must be of the same type as the variable T, i.e. a set of propositions. In (9b) the intermediate clause is not a question but a proposition. Since it is not of the same type as the variable T in the higher clause, it cannot combine with it. The structure is ruled out as a violation of Full Interpretation (Chomsky 1986a).

A related question is whether it is possible for a scope marker to move long distance. That is, we need to find out why the following derivation is not possible for the ungrammatical German example in (9a):

(9) a.

$$\lambda Q \exists q[T_i(q) \ \& \ Q(q)]$$

CP $\lambda p' \exists x[p' = speak'(m, x)]$

Was$_i$ glaubst$_j$ du t$_j$ t$'_i$ dass Peter t$_i$ meint **mit wem**$_k$ Maria t$_k$ gesprochen hat?
what think you   that Peter believes with who Maria spoken has
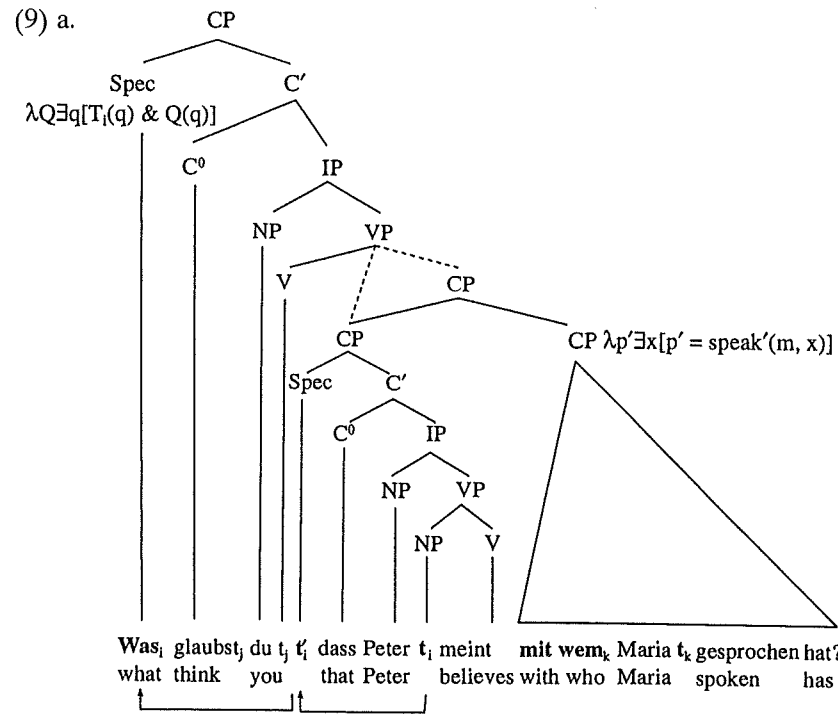
Recall that Cardinaletti (1990) allows finite subordinate clauses to be generated in complement position or in adjoined position. In order for extraction out of subordinate clauses to be licit, they must be in complement position. Thus movement of *was* to matrix Spec does not violate syntactic constraints on extraction. However, once we allow the CP to be a complement of the verb, the kind of adjunction needed for scope marking will not be possible. Further, there is a problem with interpreting this structure. Note that once *was* is in matrix Spec, the variable T is introduced there — the intermediate trace is not expected to have any

semantic content. But then T is too high for the [+WH] complement adjoined to the intermediate clause to combine with it and we cannot assign a meaning to the CP-adjoined structure. Thus (9a) is both syntactically and semantically ill-formed.

Finally, the analysis explains why the matrix verb in a scope-marking structure must be able to take a [−WH] complement but the complement itself must be [+WH]:

(11) a. *jaun **kyaa** jaantaa hai meri ravi-se   baat karegii?
         John what knows   Mary Ravi with will-talk

b. jaun **kyaa** jaantaa hai meri **kis-se**   baat karegii?
   John what knows   Mary who-with will-talk

c. *jaun **kyaa** puuchhtaa hai meri **kis-se**   baat karegii?
   John what asks   Mary who-with will-talk

In the present system, the complement combines with the main clause by functional application. Thus it must be of the same type as T, the restriction on the variable bound by the scope marker. In (11a) T is of type $\langle\langle s, t \rangle, t \rangle$ while the complement is of type $\langle s, t \rangle$, so functional application is not possible and we end up with a violation of Full Interpretation. In (11b) both T and the complement are of type $\langle\langle s, t \rangle, t \rangle$, so interpretation proceeds smoothly. In (11c), the matrix verb *ask* allows only for sets of propositions in its answer. The variable q bound by the scope marker is of type $\langle\langle s, t \rangle, t \rangle$, and T therefore has to be of a higher type $\langle\langle\langle s, t \rangle, t \rangle, t \rangle$; however, the complement is of type $\langle\langle s, t \rangle, t \rangle$. This blocks functional application and results in ungrammaticality due to the violation of Full Interpretation.

In concluding this discussion let me reiterate what I see as the chief theoretical advantage in adopting the indirect dependency approach to scope-marking structures. The indirect dependency approach is compositional in the sense that each lexical item is fully interpreted in the position at which it occurs at LF. We can therefore follow independently established procedures in interpreting scope-marking structures. The result is a simple and straightforward semantic analysis for this seemingly complex structure. The direct dependency approach, on the other hand, requires that we hold off on interpretation of *wh*-expressions in Spec position until a higher point is reached, in order to provide the quantificational force of the scope marker. This violation of compositionality leads to the adoption of rather complicated and ad hoc measures in the interpretation of more complex cases.

3.4. *Empirical Predictions*

In addition to its theoretical advantages, the indirect dependency approach has a substantive empirical payoff. For example, it captures the generalization that in all languages that show scope marking, the scope marker is also the lexical item used in question-answer pairs like (24), i.e. the *wh*-expression that analogously to English *what* allows quantification over propositions. The direct dependency approach completely misses this generalization.

As pointed out to me by Ken Hale and Maria Bittner (personal communication), the indirect dependency approach further predicts that if a language uses different lexical items for questioning the object position of a verb like *eat*, which requires quantification over inanimate individuals, and the object position of a verb like *think*, which requires quantification over propositions, it is the latter that will be used in scope-marking structures. And indeed, this prediction is borne out in Warlpiri. (29a) quantifies over individuals and uses *nyiya* whereas (29b) quantifies over propositions and uses *nyarrapa*. (29c), a scope-marking structure, uses the latter:

(29) a. **nyiya** ka   nga-rni?
      what   PRS.3s.3s. eat-NPST

      'What is he eating?'

   b. **nyarrpa-rlu** O-ngku   yimi-ngarru-rnu?
      how-ERG   PRF-3s.2s. speech-tell-PST

      'What did he tell you?'

   c. **nyarrpa-rlu** O-ngku   yimi-ngarru-rnu kuja-pala
      how-ERG   PRF-3s.2s. speech-tell-PST COMP-3ds

      kurdu-jarra **nyarrpara-kurra**   ya-nu?
      child-DUAL where-ALLATIVE   go-PST

      'Where did he tell you the two children went?'

It seems quite clear, then, that there are sound empirical grounds as well for adopting the indirect dependency approach to scope-marking structures.

3.5. *The Truth Requirement for Propositions*

In the semantics for questions that I have adopted so far (Hamblin 1973), questions denote the set of propositions that count as possible answers.

This account was modified by Karttunen (1977) in order to explain the valid inference from (30a, b) to (30c):

(30) a. John told Sue who came.
    b. Bill came.
    c. John told Sue that Bill came.

Since *tell* does not entail the truth of its complement, Karttunen concluded that the validity of the inference must be due to the semantics of the embedded question: the denotation of a question must impose a truth requirement on the propositions included. However, Karttunen's modification gives the wrong results for scope-marking structures. To see this, consider the denotation of (3), under the analysis in (28a), with the truth requirement added:

(31) $\lambda p\exists q[\exists x[\check{}q \;\& \;q = \hat{}\,\text{will-talk}'(m, x)] \;\& \;\check{}p \;\& \;p = \hat{}\,\text{think}'(j, q)]]$

Take a situation in which Mary does not, in fact, talk to Bill, but John thinks she does. Intuitively, we want the proposition ^*John thinks that Mary will talk to Bill* to be allowed into the denotation of the question, but the truth condition in (31) rules it out. Here *that Mary will talk to Bill* will be assigned as the value of the variable q, bound by the scope marker. Now, there are two places where truth will be checked. The truth requirement of the matrix clause will be satisfied, since it is the truth of p, *that John thinks q*, that will be checked at that level. But the complement too will only include those propositions that are true, and so *that Mary will talk to Bill* will not be in the complement propositional set. Thus q will not satisfy the first condition and the proposition *John thinks Mary will talk to Bill* will be prevented from entering the denotation of the question, an incorrect result.

Given the soundness of Karttunen's motivation for including a truth requirement in the semantics of questions, on the one hand, and the results of the semantics for scope marking we have obtained, on the other, there is clearly a need to reconcile the two. I will briefly speculate on one possible way of doing so.

Let us take the meaning of a question to be the set of possible answers to it, as proposed by Hamblin. We can easily define the notion of the true answer to a question as a function from the set of possible answers denoted by the question to that subset of propositions in it that are true: $Ans(Q) = \lambda p[p \in Q \;\& \;\check{}p]$. It seems plausible enough to suggest that to wonder about a question is really to wonder about what its answer is. To tell a question is to tell the answer to that question. By the same token, for one question to depend on another is really for the answer to one question

to depend on the answer to the other; and so on, for all of the cases discussed by Karttunen. That is, we maintain all of Karttunen's results, but we do so by assuming that the shift from the set of possible answers to the set of true answers is licensed only in the scope of certain verbs. In particular, intensional verbs will not license this shift while extensional verbs will. Thus *John wondered who came* will translate as *wondered'(j, ^Ans(λp∃x[p = ^came(x)]))*, which will not entail that if Bill came, John wondered whether Bill came. *John told Sue who came*, on the other hand, will translate as *told'(j, s, ˇ^Ans(λp∃x[p = ^came(x)]))*. Since *tell* is extensional, the members of the answer set in the actual world will be accessible, accounting for the inference in (30).

Turning now to scope-marking structures, we can see that nothing special need be said about them. Since the actual complement is not embedded under the verb, it will always denote the set of all possible answers, not just the true ones.[11]

In summary, I suggest that one way to reconcile the Hamblin-style semantics required for scope-marking structures with the Karttunen-style semantics required for cases like (30) is to distinguish between the meaning of a question — the set of possible answers — and the meaning of an answer — the subset of true propositions — and to restrict the availability of the latter to embedded contexts. The unavailability of the set of true propositions in the case of scope marking simply follows from the syntactic structure adopted here.

Thus it seems to me that the semantics proposed above for the indirect dependency approach will stand up to scrutiny. It is time now to return to the negation facts discussed by Rizzi and see how they can be incorporated.

4. NEGATION IN QUESTIONS

4.1. *Weak Islands Revisited*

The tacit assumption in previous analyses of scope-marking structures has been that they are equivalent to extraction structures. This was the intuition behind the direct dependency approach; in fact, it also initially motivated the indirect dependency approach proposed above. I now want to show that this was actually a problematic assumption: the two structures

---

[11] I thank Irene Heim and Angelika Kratzer for pointing out that once Karttunen's truth requirement is built into the embedded position rather than into the question denotation, the problem it poses for scope marking is resolved automatically.

cannot be collapsed. This poses real problems for the direct dependency approach. The indirect dependency approach, by contrast, can be extended to incorporate the necessary distinctions.

Let us reconsider Rizzi's discussion of negation in scope marking. The relevant example is repeated below, with the translation it would receive in the present approach:

(14) b. *__Was__ glaubst du nicht, __mit wem__ Maria gesprochen hat?
what think you not with whom Maria spoken has

'Who don't you think Maria has spoken to?'
λp∃q[∃x[q = ^talk'(m, x)] & p = ^not-think(you, q)]

As it stands, the indirect dependency approach cannot account for the ungrammaticality of (14b), as the translation shows. This suggests that perhaps the negation facts are best handled in the syntactic component, as claimed by Rizzi. Recall that his explanation draws a parallel between scope marking and adjunct extraction. To recapitulate briefly, the idea is that scope markers (like adjuncts) do not bind the expression in the lower Spec through referential indices but must antecedent govern them. The ungrammaticality of (14b) is expected since negation interferes with antecedent government.

Although Rizzi's explanation of the scope-marking facts is based on the direct dependency approach, it is not incompatible with the alternative being explored here, as I argued in Dayal (1993). The scope marker is no doubt in object position, but that position is devoid of semantic content. Thus, it is not implausible to assume that it could not carry referential indices. The rest would follow as in Rizzi's account.

Of course, Rizzi's account of the negative island effect is not uncontroversial. A problem addressed by Rizzi himself is that it requires negation to be analyzed as an A-bar specifier, while many recent studies claim that it is a functional head.[12] Furthermore, the account has been challenged by Szabolcsi and Zwarts (1993), who argue that the negative island effect is semantic rather than syntactic. While it is not immediately

---

[12] In fact, Dwivedi (1991) claims that Hindi negation is a functional head. In Dayal (1993) I show that it may be the case that Hindi has two negations. The functional head is the one that takes a complement to its left and yields a contrastive reading for it, as claimed by Dwivedi. Sentential negation, which seems to be relevant to the scope-marking facts, may be a verbal modifier that modifies the element to its right. Some support for this distinction comes from the following paradigm:

(i)   anu-ne   ravi-ko   kitaab dii   nahiiN, balki pheNkii.
Anu-ERG Ravi-DAT book gave not   but   threw

obvious how their system would handle negation in scope marking, it may well be capable of doing so. I will not try to demonstrate that here, however, since I no longer believe that the impossibility of negation in scope marking can be reduced to the case of negative islands. Let us see why.

One implication of Rizzi's account of the phenomenon is that scope marking is expected to be ruled out in every case that disallows adjunct extraction, i.e. in all weak islands. This correlation does not hold up, though. Factive islands, for example, are weak in that they allow argument extraction but block adjunct extraction (Cinque 1991):

(32) a. Who do you know that Mary will hire?
b. *Why do you know that Mary will hire Sue?

Contrary to expectation, scope marking is quite possible with factive islands:[13]

(33) a. jaun **kyaa** jaantaa hai meri **kis-se** baat karegii?
John what knows Mary who-with will-talk

'Who does John know Mary will talk to?'

b. tum-ko **kyaa** pataa calaa meri **kyuuN** nahiiN aayegii?
you-Dat what discovered Mary why not will-come

'Why did you discover that Mary won't come?'

Thus I take it that the impossibility of negation in scope-marking structures cannot be reduced to the weak island phenomenon; the explanation must be sought elsewhere.

4.2. *Negative Questions and D-linking*

In order to understand what is going on, let us step back a bit from scope-marking structures and look at the difference between simple affirmative and negative questions like (34):

(ii)   *anu-ne   ravi-ko   kitaab nahiiN dii,   balki pheNkii.
Anu-ERG Ravi-DAT book   not   gave but   threw

'Anu didn't give the book to Ravi, she threw it.'

If this distinction is real, it may be possible to maintain an explanation for the scope marking facts along the lines suggested above.

[13] Bayer (personal communication) points out that German does not allow scope marking with factive verbs. Interestingly, Cardinaletti (1990, 85—91) notes that German does not allow finite complements to be adjoined hihger than VP. The semantics for scope marking given here assumes adjunction to CP (see also fn. 10), and it may be that scope marking with VP adjunction is uninterpretable.

(34) a. What did John buy?
b. What didn't John buy?

Consider the contexts in which these questions could be uttered. Suppose the questioner sees several purchased items, say a_1, a_2, a_3 . . . , but does not know which of the items were bought by John. She could ask either (34a) or (34b). Though the presuppositions behind the two questions may differ, both denote sets that list for each of the three items whether John bought it or not.

Now consider a context where the questioner only knows that John went shopping but does not see any of the items purchased nor has any notion what things were on the shopping list. That is, there is no pre-established domain of quantification. Here, the affirmative question (34a) is possible but the negative question (34b) is ruled out. Presumably the reason is that it would be impossible to give an exhaustive list of all the things John did not buy. There are several important issues connected to this, discussed most recently in Lahiri (1991) and Szabolcsi and Zwarts (1993), that I cannot go into. What is relevant for present purposes is the rather uncontroversial intuition that negative questions are possible only with D-linked domains.[14]

Turning now to questions which quantify over propositional variables, we see that the requirement of a negative question for a D-linked domain is, if anything, stronger in these cases. Compare (35a) and (35b):

(35) a. What does John know?
b. What doesn't John know?

The negative question in (35b), unlike the affirmative question in (35a), has either an echo, a rhetorical, or a D-linked interpretation. Crucially, it lacks the normal open-ended interpretation. To see this, consider the following contexts of use.

A context that supports the D-linked interpretation would be something like this. Suppose students in Linguistics 101 are being evaluated with respect to a set of core facts, namely *that language is innate, that language is systematic, that all dialects of a language are equally good.* While grading Bill's paper, one T.A. remarks to another that Bill knows the first fact but not the other two. When she comes to John's paper, her friend might easily ask *And what doesn't John know?*. However, take a minimally

---

[14] Note that negative questions like *Why didn't John fix the car?* are not restricted to D-linked/echo interpretations. The non-D-linked reading, however, is only allowed if *why* is interpreted outside the scope of negation, as in *Why is it the case that John did not fix the car?*.

different situation now where personnel managers are simply discussing the relative merits of job candidates without having a fixed set of facts against which to evaluate them. They might ask the affirmative question (35a) but not the negative question (35b).

My claim is that if we put together the generalization that negated questions require D-linked domains of quantification with the analysis of scope-marking structures presented here, we obtain a straightforward explanation of the negation facts. In a D-linked context, the restriction T on the propositional variable has its value determined by context. Since T is not free, the complement denotation cannot be substituted for it via functional application. We therefore end up with a violation of the principle of Full Interpretation.

To sum up, I have drawn a connection between the possibility of negation in questions and D-linking. I have also explained the correlation between the possibility of interpreting the complement in a scope-marking structure and the occurrence of a free variable in the main clause. Under the assumption that T can either refer to the D-linked set of propositions or be combined with the complement propositional set, the impossibility of negation in scope marking is predicted.[15]

This approach to the negation facts makes another prediction. Affirmative questions, we know, are compatible with D-linked as well as non-D-linked domains. In my terms, this means that the value of T may or may not be contextually determined. However, scope-marking structures will only be licit if T is free, i.e. when the domain is not D-linked. It seems to me that the prediction is borne out: scope-marking structures never have a D-linked interpretation.[16]

5. CONCLUSION

I started out by posing a question. Should semantic equivalence between two structures dictate syntactic parallelism at abstract levels of representation or should surface syntactic differences be maintained in the input to meaning? The proposed analysis suggests an answer to this question.

---

[15] Szabolcsi and Zwarts (1993) give some examples of scope-marking structures in Hungarian in which negation is at least marginally acceptable. At this point, I do not have a full understanding of what may be at issue in these cases and will settle for accepting the explanation in terms of expletive replacement that they provide.

[16] Another interesting correlation that is not explored here is the inability of scope-marking structures to have echo and rhetorical interpretations, readings that are available for negative questions. My sense is that both kinds of readings involve contextual anchoring of the variable T, which blocks the functional application required for combining the complement and the main clause denotations of a scope-marking structure.

We have seen that there is considerable overlap in meaning between scope-marking and extraction structures. Yet, there are some subtle but real differences between the two. I hope to have demonstrated that a semantic analysis which is sensitive to syntactic distinctions is better able to handle the facts. At least for the case at hand, then, it appears that the more fruitful approach is to maintain surface syntactic distinctions. This, of course, does not imply that semantic intuitions should not play a role in syntactic analyses; it only means that if apparent semantic equivalence leads to LF representations that diverge radically from surface representations, the premise of equivalence needs to be checked. The moral of the story, as I take it, is that surface syntax is a good guide to determining meaning.

REFERENCES

Bayer, Josef: 1990, *Directionality of Government and Logical Form: A Study of Focusing Particles and WH-Scope*, unpublished habilitation thesis, University of Konstanz.
Belnap, Nuel and Jonathan Steel: 1963, *The Logic of Questions and Answers*, Yale University Press, New Haven.
Bittner, Maria: in press, *Case, Scope and Binding*, Kluwer, Dordrecht.
Cardinaletti, Anna: 1990, *Impersonal Constructions and Sentential Arguments in German*, Unipress, Padua.
Chomsky, Noam: 1986a, *Knowledge of Language: Its Nature, Origin and Use*, Praeger, New York.
Chomsky, Noam: 1986b, *Barriers*, MIT Press, Cambridge, Mass.
Cinque, Guglielmo: 1991, *Types of A-Bar-Dependencies*, MIT Press, Cambridge, Mass.
Davison, Alice: 1984, 'Syntactic Constraints on WH in situ: WH-questions in Hindi-Urdu', paper presented at the LSA Annual Meeting.
Dayal, Veneeta Srivastav: 1993, 'Negation in Scope Marking Structures', paper presented at the LSA Annual Meeting, Los Angeles.
Dwivedi, Veena: 1991, 'Negation as a Functional Projection in Hindi', *Proceedings of WECOL*.
Frege, Gottlob: 1960, *Translations from the Philosophical Writings of Gottlob Frege*, P. Geach and M. Black (eds.), Blackwell, Oxford.
Hamblin, C. L.: 1973, 'Questions in Montague English', *Foundations of Language* **10**, 41—53.
Higginbotham, James and Robert May: 1981, 'Questions, Quantifiers and Crossing', *The Linguistic Review* **1**, 41—80.
Huang, C.-T. James: 1982, *Logical Relations in Chinese and the Theory of Grammar*, unpublished Ph.D. dissertation, MIT.
Karttunen, Lauri: 1977, 'Syntax and Semantics of Questions', *Linguistics and Philosophy* **1**, 3—44.
Lahiri, Utpal: 1991, *Embedded Interrogatives and Predicates that Embed Them*, unpublished Ph.D. dissertation, MIT.
Mahajan, Anoop: 1990, *The A/A-bar Distinction and Movement Theory*, unpublished Ph.D. dissertation, MIT.
McCloskey, James: 1992, 'Adjunction, Selection and Embedded Verb Second', Working Paper LRC-92-07, Linguistics Research Center, UC Santa Cruz.

McDaniel, Dana: 1989, 'Partial and Multiple Wh-Movement', *Natural Language and Linguistic Theory* 7, 565—604.

Riemsdijk, Henk van: 1983, 'Correspondence Effects and the Empty Category Principle', *Tilburg Papers in Language and Literature.*

Rizzi, Luigi: 1992, 'Argument/Adjunct (A)symmetries', *Proceedings of NELS* 22, 365—381.

Srivastav, Veneeta: 1989, 'Hindi WH and Pleonastic Operators', *Proceedings of NELS* 20, 443—57.

Srivastav, Veneeta: 1991a, *WH Dependencies in Hindi and the Theory of Grammar,* unpublished Ph.D. dissertation, Cornell University.

Srivastav, Veneeta: 1991b, 'Subjacency Effects at LF: The Case of Hindi WH', *Linguistic Inquiry* 22(4), 762—769.

Stowell, Timothy: 1981, *Origins of Phrase Structure,* unpublished Ph.D. dissertation, MIT.

Szabolcsi, Anna and Frans Zwarts: 1993, 'Weak Islands and an Algebraic Semantics for Scope-Taking', *Natural Language Semantics* 1(3), 1—50.

Wahba, Wafaa Abdel-Faheem Batran: 1991, 'LF Movement in Iraqi Arabic', in C.-T. James Huang and Robert May (eds.), *Logical Structure and Linguistic Structure,* Kluwer, Dordrecht.

Rutgers University
Department of Linguistics
18 Seminary Place
New Brunswick, NJ 08903, U.S.A.
srivastav@zodiac.rutgers.edu

Reversible Grammar in Natural Language Processing

edited by **Tomek Strzalkowski**

Reversible grammar allows computational models to be built that are equally well suited for the analysis and generation of natural language utterances. This task can be viewed from very different perspectives by theoretical and computational linguists, and computer scientists. The papers in this volume present a broad range of approaches to reversible, bi-directional, and non-directional grammar systems that have emerged in recent years. This is also the first collection entirely devoted to the problems of reversibility in natural language processing. Most papers collected in this volume are derived from presentations at a workshop held at the University of California at Berkeley in the summer of 1991 organised under the auspices of the Association for Computational Linguistics.
This book will be a valuable reference to researchers in linguistics and computer science with interests in computational linguistics, natural language processing, and machine translation, as well as in practical aspects of computability.

New Publication

Contents and Contributors

KLUWER ACADEMIC PUBLISHERS