

# Capturing Word Order Asymmetries in English

Chan Chung  
Dongseo University

Jong-Bok Kim  
Kyung Hee University

Proceedings of the 10th International Conference on  
Head-Driven Phrase Structure Grammar

Michigan State University

Stefan Müller (Editor)

2003

pages 68–87

Chung, Chan, & Kim, Jong-Bok. 2003. Capturing Word Order Asymmetries in English. In Müller, Stefan (Ed.), *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar, Michigan State University*, 68–87. DOI: 10.21248/hpsg.2003.7.



## Abstract

Even though the word order in English is rather straightforward, the distributional possibilities of left-peripheral elements like topic phrases, *wh*-phrases, and negative operators (introducing an SAI) are quite intriguing and complex. In particular, there seems to exist no straightforward way of capturing the linear order asymmetries of these elements in the main and embedded clauses. The prevailing analyses have resorted to movement processes with multiple functional projections. The goal of this paper is to explore an alternative analysis to such movement-based analyses. In particular, this paper adopts the notion of topological fields (DOMAIN) proposed by Kathol (2000, 2001) within the framework of HPSG. The paper shows that within this DOMAIN system, the distributional possibilities as well as the asymmetries we find in English left peripheral constructions can easily follow from the two traditional views: (i) a topic precedes a focus element, and (ii) in English a *wh*-phrase and a complementizer competes with each other for the same position.

Key words: left-peripheral elements, domain, HPSG, *wh*-phrases, topic, negative-operator, focus

## 1 Basic Facts

### 1.1 Left Peripheral Elements in the Main Clause

Most of the English speakers we consulted do not allow two consecutive topics or *wh*-phrases in finite root clauses:

- (1) a. \*?[On the desk], [this book], John put.  
b. \*[To whom], [what] should Bill give?.  
c. \*What when will you do?

However, when two different kinds of these phrases occur in the left-peripheral positions, the topic needs to precede the *wh*-phrase (Hooper and Thompson 1973, Langendoen 1979, Watanabe 1993, Haegeman 2000):

- (2) a. [This book], [to whom] should Bill give?  
b. [These prices], [what] can anyone do about?  
c. [During the vacation], [for what kind of jobs] would you go into the office?
- (3) a. \*And [to whom], [a book like this], would you give?  
b. \*[What], [these prices] can anyone do about?  
c. \*[For what kinds of jobs] [during the vacation] would you go into the office? (Baltin 1982)

When a topic occurs with a negative operator that combines with an inverted sentence, the topic also needs to precede the operator (Haegeman 2000):<sup>1</sup>

- (4) a. [To John], [nothing] would we give.  
       b. [These books], [only with great difficulty] can she carry.  
       c. [During the vacation], [on no account] would I go into the office.
- (5) a. \*[Nothing], [to John], would we give.  
       b. \*[Only with great difficulty], [these books], can she carry.  
       c. \*[On no account] [during the vacation], would I go into the office.

However, the situation is different with a *wh*-phrase. In root clauses, the *wh*-phrase cannot appear together with a negative SAI operator, regardless of its sequential ordering relation with the operator.

- (6) a. \*[On which table] [only with great difficulty] would she put the big rock?  
       b. \*[Only with great difficulty] [on which table] would she put the big rock?

As in (6), neither a *wh*-phrase nor a negative operator can precede the other.

## 1.2 Left-Peripheral Elements in the Embedded Clause

With respect to the order of left peripheral elements, English embedded clauses differ from root clauses in several respects. One contrast we can find is the ordering relations between *wh* phrases and topic phrases. For example, unlike in root clauses, the *wh*-phrase in embedded clause must precede the topic phrase:<sup>2</sup>

- (7) a. the man [to whom], [liberty], we could never grant  
       b. ?I wonder [to whom] [this book], Bill should give.  
       c. I was wondering [for which jobs], [during the vacation], I should go into the office.
- (8) a. \*the man [liberty], [to whom] we could never grant  
       b. \*I wonder [this book], [to whom] Bill should give. (Petsetsky 1989)  
       c. \*I was wondering, [during the vacation], [for which jobs] I should go into the office.

---

<sup>1</sup>We assign the term ‘negative operator’ for the negative expressions as well as expressions like *only* that combine with an SAI (subject-auxiliary inversion) sentence.

<sup>2</sup>An outside reviewer of the original abstract points out that there could be cases that appear to violate this ordering restriction:

- (i) a. (??) I’ve always had this sort of attitude that Joe<sub>i</sub>, how much<sub>j</sub> can you say \_\_<sub>j</sub> about \_\_<sub>i</sub>?  
       b. (??) On the other had, you’re always kind of thinking that Joe<sub>i</sub>, what a lot of nice things<sub>j</sub>, there are to say \_\_<sub>i</sub> about \_\_<sub>j</sub>.

In the examples, the topic precedes the *wh*-phrase. Most of the speakers we consulted did not accept these. The embedded clauses here seem to function as exclamative clauses that cannot appear in canonical root clauses. We believe such examples are allowed in highly limited, colloquial contexts with proper phonological prominence.

The positional possibilities of a negative operator are also slightly different: As in root clauses, the negative operator should follow the topic clause as shown in (9) and (10).<sup>3</sup>

- (9) a. Becky said that [these books], [only with great difficulty], can she carry.  
 b. He said that [beans], [never in his life], had he been able to stand.
- (10) a. \*He said that [never in his life], [beans], had he been able to stand.  
 b. \* I promised that [on no account] [during the holidays] will I write a paper. (Haegeman 2000)

However, in embedded clauses, the negative operator can appear with a *wh*-phrase when the operator follows it:

- (11) a. I wonder [on which table] [only with great difficulty] would she put the big rock.  
 b. \*I wonder [only with great difficulty] [on which table] would she put the big rock.

The embedded clause ordering in (11)b is thus not allowed in the root clause.

The table (12) summarizes what we have observed so far with respect to the ordering relations among left peripheral elements in English.

Sequence		Root clause	Embedded clause
(12)	Topic-ph and Wh-ph	OK	*/??
	Topic-ph and Neg-Op	OK	OK
	Wh-ph and Topic-ph	*	OK
	Wh-ph and Negative-op	*	OK
	Neg-op and Topic-ph	*	*
	Neg-op and Wh-ph	*	*

Though not all, most English speakers exhibit clear contrasts between root and embedded clauses with respect to the positional possibilities of left-peripheral elements. As in (12), when a topic and a *wh*-phrase can cooccur in a root clause, the topic precedes the *wh*-phrase. However, when they cooccur in an embedded clause, the order gets reversed. Another asymmetry observed is that in root clauses, a topic can appear together with a negative operator when the first precedes the latter, whereas a *wh*-phrase and a negative phrase cannot cooccur together at all. They can appear together only in embedded clauses with the sequence of *wh*-phrase and negative operator.

<sup>3</sup>One thing to note at this point is that an SAI with the negative operator does not occur with a topic phrase in either order:

- (i) a. \*Becky said that [these books], [only with great difficulty] she can carry.  
 b. \*He said that [beans], [never in his life] he had been able to stand.

Numerous attempts have been made to account for the asymmetries we have observed here, but most within the interactions between functional projections and movement processes (e.g, Rizzi 1997, Haegeman 2000, etc). In what follows, we will provide a nonconfigurational analysis that relies on the notion of topological fields developed by Kathol (2000, 2002).

## 2 Analysis with the Notion of Topological Fields

The theoretical framework that we adopt to account for such asymmetries in English left-peripheral constructions is HPSG (Head-driven Phrase Structure Grammar). In particular, we adopt the word order domain theory developed by Reape (1994) and Kathol (2000, 2001, 2002) for the clausal structures of German. One prominent example Kathol (2000) cites is the complementary distribution of German complementizers and finite verbs:

- (13) a. ob        Hans die Zeitung    liest  
               whether Hans the newspaper reads.  
               ‘whether Hans read the newspaper.’  
       b. \*ob        liest Hans die Zeitung  
               whether reads Hans the newspaper

As noted here, the presence of the complementizer requires the verb to be in the sentence final position even if in other environments finite verbs can occur clause-initially. Central to Kathol’s analysis is the level of DOMAIN consisting of an ordered list of elements that contain phonological and categorial information. The order domain of the mother category is computed from the information provided by the daughter constituents at each syntactic combination. Each element of a clausal domain is uniquely marked for the region that it belongs to. For example, within the DOMAIN given in the following German complementizers are occur only in [2] whereas finite verbs occur in either in [2] or [4].

- (14) Distributional Constraints (German) (Kathol 2000):  
       [1], [2], [3], ....[3], [4], ...[4], ....

With the general constraint that limits the number of elements that instantiate as [2] to one, Kathol’s analysis takes the complementarity between finite verbs and complementizers as straightforward constraint satisfaction.<sup>4</sup>

Adopting this idea, we here provide an analysis of English left peripheral elements. The table in (15) is the clausal domain we assume for English:

---

<sup>4</sup>Kathol (2002) also applies this idea into English. For example, he takes English inverted finite auxiliary verbs and complementizers to have the identical index number, capturing the complementarity effects between the two. See Kathol (2002) for details.

	marker field 1	topic field 2	focus field 3
(15) main-cl:		topic	foc ( <i>wh</i> & neg-op)
embedded-cl:	comp <i>wh</i>	topic	neg-op

The table in (15) reflects English word order generalizations: a topic phrase precedes a focused element (*wh* & negative operator). The negative operator gets the function of a ‘focus’ operator, triggering the following sentence to be an inverted one.<sup>5</sup>

The only difference from root and embedded clauses is that a *wh*-phrase in embedded clauses competes with a complementizer for the first position.<sup>6</sup>

As in Kathol (2001), we assume that the different topological fields emerge by virtue of the topological number index (from 1 to 3 for the scope of this paper) borne by a domain element. The assignment of the index numbers can be either lexical or constructional. For example, English lexical complementizers including *that* are always assigned to the positional class 1:

$$(16) \left[ \begin{array}{c} \text{DOM} \left\langle \begin{array}{c} 1 \\ \text{PHON} \langle \text{that} \rangle \end{array} \right\rangle \\ \text{HEAD } \textit{comp} \end{array} \right]$$

Meanwhile, constructional constraints will impose appropriate index numbers to a topic, a focus, or a *wh*-phrase as the following:<sup>7</sup>

(17) *topic-cl*:

$$[] \rightarrow \left[ \begin{array}{c} \text{DOM} \langle [2] \rangle \\ \text{TOPIC} + \end{array} \right], \quad \mathbf{H} \left[ \begin{array}{c} \text{VFORM } \textit{fin} \\ \text{IC} + \\ \text{INV} - \end{array} \right]$$

(18) *foc-cl*:

$$[] \rightarrow \left[ \begin{array}{c} \text{DOM} \langle [3] \rangle \\ \text{WH} +/\text{NEG} + \end{array} \right], \quad \mathbf{H} [\text{IC} +]$$

<sup>5</sup>There exist several commonalities between *wh*-phrases and negative operators, as noted in Rizzi (1999) and Haegeman (2000). They both occupy *A'* positions and combine with inverted sentences. These two, identically licensing negative polarity items, also are in a sense quantificational. Adopting this line of observation and Haegeman (2000), we call these two as focus markers though there remain finer distinctions.

<sup>6</sup>This partly reflects Chomsky & Lasnik’s (1977) Doubly-Filled COMP Filter (DFCF) constraint and can be found in German too. In standard German, a front *wh*-phrase never cooccurs with an overt complementizer. See Kathol (2001: 38).

<sup>7</sup>We assume the existence of *foc(us)*-clause whose subtypes include (direct and indirect) interrogative *wh*-clauses and negative SAI clauses (cf. Ginzburg and Sag 2000). Thus, a *wh*-element will be focus marked either in the inverted or in the first position of non-inverted embedded clauses. The NEG operator construction itself requires its sentence to be inverted as its constructional constraint.

(19) *embed-wh-cl*:

$$\left[ \text{IC} - \right] \rightarrow \left[ \begin{array}{c} \text{DOM} \langle [1] \rangle \\ \text{WH} + \end{array} \right], \quad \mathbf{H}[\text{INV} +]$$

The clausal constructions here are independently motivated for the proper descriptions of English (Ginzburg and Sag 2000). For example, topic clauses in (17) can be built from independent finite clauses ([IC +]), blocking examples like the following:<sup>8</sup>

- (20) a. \*John persuaded Bill [the project]<sub>i</sub> to finish \_\_<sub>i</sub>.  
 b. \*John tried [the man]<sub>i</sub> to kill.

The *foc-cl* in (18) is also required for the obligatory inversion with the negative operator:

- (21) a. At no time would Leslie run for any public office.  
 b. \*At no time, Leslie would run for any public office.
- (22) a. \*In no time would Leslie run for any public office.  
 b. In no time, Leslie would run for any public office.

When the expression functioning as a negative operator occurs in the sentence initial position, it should combine with an SAI sentence.<sup>9</sup>

Finally the existence of *embed-wh-cl* in (19) as a subtype of *wh-cl* (cf. Ginzburg and Sag 2000) allows a *wh*-phrase to combine with a noninverted sentence when occurring in embedded clauses. This is one main difference between clauses with a negative operator and embedded *wh*-clauses.<sup>10</sup>

With such independently required constructional constraints, the present analysis just introduces a topological index number to each syntactic constituent. The constraint in (17) specifies that a topichood phrase bears the index number 2 whereas the one in (18) tells us that a *wh*-phrase or a negative-operator focused phrase gets the domain index number 3 only if this combines with an SAI sentence. However, when a *wh*-phrase combines with a noninverted sentence, the phrase is assigned to the index number 1. With these quite general and independently motivated constraints on clauses from (17) to (19), the domain indices impose linear sequence constraints on the position classes by means of the linear precedence constraint in (23):

- (23) Topological Linear Precedence Constraint (cf. Kathol 2001):  
 $1 < 2 < 3$

<sup>8</sup>The feature IC means independent clause in Ginzburg and Sag 2000.

<sup>9</sup>We thus in a sense assume that the negative operator construction is a subtype of *foc-cl* with this SAI construction.

<sup>10</sup>We assume that *wh*-clauses are partitioned into *root-wh-cl* and *embed-wh-cl*, both of which are again classified into *sub-wh-cl* and *non-subj-cl*. See Ginzburg and Sag 2000.

The constraint in (23) will make it possible to directly impose an ordering restriction on signs in an order domain, not just their phonology values.

One could observe that this simple system could provide a straightforward account of the main asymmetries between root and embedded clauses we have observed in section 1. As noted in (2) and (3), we have observed that in root clauses a topic must precede a *wh*-element or a negative operator, but not the other way around. The clausal domain in (15) and the LP (Linear Precedence) constraint in (23) together can easily capture this contrast. For example, the sentence (2)a, repeated here in (24)a, would have the domain order in (24)b:

- (24) a. [This book], [to whom] should Bill give?  
 b.  $\left[ \text{DOM} \left\langle \left[ \begin{array}{c} 2 \\ \text{PHON } \langle \text{this book} \rangle \end{array} \right] \left[ \begin{array}{c} 3 \\ \text{PHON } \langle \text{to whom} \rangle \end{array} \right] \dots \right\rangle \right]$

(24)b observes all the relevant constraints. However, examples like (3)a, repeated here in (25)a, are simply not licensed since the *wh*-phrase precedes the topic. The domain order in (25)b illustrates this point:

- (25) a. \*To whom a book like this would you give?  
 b.  $\left[ \text{DOM} \left\langle \left[ \begin{array}{c} 3 \\ \text{PHON } \langle \text{to whom} \rangle \end{array} \right] \left[ \begin{array}{c} 2 \\ \text{PHON } \langle \text{a book like this} \rangle \end{array} \right] \dots \right\rangle \right]$

We can also predict that in root clauses, the topic phrase needs to precede the negative operator as we have seen in (4)a, given in (26)a again. The ordering domain of this sentence given in (25)b proves this clearly:

- (26) a. [To John] [nothing] would we give.  
 b.  $\left[ \text{DOM} \left\langle \left[ \begin{array}{c} 2 \\ \text{PHON } \langle \text{To John} \rangle \end{array} \right] \left[ \begin{array}{c} 3 \\ \text{PHON } \langle \text{nothing} \rangle \end{array} \right] \dots \right\rangle \right]$

When this ordering is reversed as in (5) (e.g., \**Nothing, to John, would we give.*) we obtain an undesirable ordering since the topic phrase with the domain index 2 does not precede the negator operator with the index number 3.

The word order of left-peripheral elements in embedded clauses can also be explained straightforwardly. The main difference between root and embedded clauses comes from the fact that the domain index value 1 is assigned to a *wh*-phrase combining with a noninverted sentence as well as to English complementizers like *that*. This reflects the well known competition between complementizers and *wh*-phrases. There is therefore nothing wrong to have the sequence of *wh*-phrase and topic phrase as in the examples (7). We repeat the example here in (27)a and represent its domain order in (27)b (i.e.,  $1 < 2$ ):



(27) a. the man to whom, liberty, we could never grant..

$$b. \left[ \text{DOM} \left\langle \left[ \begin{array}{c} 1 \\ \text{PHON } \langle \text{to whom} \rangle \end{array} \right] \left[ \begin{array}{c} 2 \\ \text{PHON } \langle \text{liberty} \rangle \end{array} \right] , \dots \right\rangle \right]$$

However, the examples in (8), one of which is given in (28)a, again are all ruled out since the topic 2 precedes the *wh*-phrase 1. This domain order in (28)b verifies this:

(28) a. \*the man, liberty, to whom, we could never grant..

$$b. \left[ \text{DOM} \left\langle \left[ \begin{array}{c} 2 \\ \text{PHON } \langle \text{liberty} \rangle \end{array} \right] \left[ \begin{array}{c} 1 \\ \text{PHON } \langle \text{to whom} \rangle \end{array} \right] , \dots \right\rangle \right]$$

In the same spirit, the present analysis allows examples like (10)a but not those like (10)b, repeated here in (29)a and b, respectively:

(29) a. Becky said that [these books], [only with great difficulty], can she carry.

b. \*He said that [never in his life], [beans], had he been able to stand.

In both root and embedded clauses, the topic field must precede the negative operator since the former's index value is 2 whereas the latter bears the index value 3 in both clauses. (29)b is thus unacceptable since it violates the LP constraint in (23).

Since in embedded clauses, the *wh*-phrase, combining with a non-inverted sentence, is designated as bearing the index number 1, we can expect cases like (11)a where the *wh*-phrase precedes the negative operator. The sentence (11)a is repeated in (30)a and part of its domain value is given in (30)b. As noticed, the domain index number 1 precedes the domain number 3.

(30) a. I wonder on which table only with great difficulty would she put the big rock.

$$b. \left[ \text{DOM} \left\langle \left[ \begin{array}{c} 1 \\ \text{PHON } \langle \text{on which table} \rangle \end{array} \right] \left[ \begin{array}{c} 3 \\ \text{PHON } \langle \text{only with great difficulty} \rangle \end{array} \right] , \dots \right\rangle \right]$$

To capture the competition for one identical position between a *wh*-phrase and a complementizer in an embedded clause, we adopt the Uniqueness Condition in Kathol (2001) given in (31):<sup>11</sup>

(31) Uniqueness Condition (cf. Kathol 2001):

$$i < i \quad (\text{where } i = \{1, 2, 3\})$$

<sup>11</sup>As noted in Kathol (2001), this condition seems to be rather counterintuitive. However, since no element can precede itself, there couldn't be two elements with the same number index within one DOMAIN.

This condition can easily capture the complementary distribution between a *wh*-phrase and a complementizer in English embedded clauses. English does not allow cases like (32)a simply due to the fact that the two elements *to whom* and *that* here both have the identical topological domain number, as represented in (32)b:

- (32) a. \*I don't know [to whom] [that] Bill gave the book.  
 b. 
$$* \left[ \text{DOM} \left\langle \left[ \begin{array}{c} 1 \\ \text{PHON} \langle \text{to whom} \rangle \end{array} \right] \left[ \begin{array}{c} 1 \\ \text{PHON} \langle \text{that} \rangle \end{array} \right] \dots \right\rangle \right]$$

The constraint in (31) also can rule out examples like (6) where in a root clause we have the sequence of *wh*-phrase and negative operator as repeated here:

- (33) \*On which table only with great difficulty would she put the big rock?

Such examples are not allowed simply because by the constructional constraints of the clause types given in (17) and (18), both the *wh*-phrase and the negative operator in the main clause combine with inverted sentences and both are assigned the domain index value 3. This would then violate the Uniqueness Condition.<sup>12</sup>

In addition, the present analysis allows neither two consecutive topics nor *wh*-phrases, whose examples we have already seen in (1). For example, the badness of (1)a, repeated here in (34)a, can easily be seen from the DOMAIN field given in (34)b:

- (34) a. \*? On the desk, this book, John put.  
 b. 
$$* \left[ \text{DOM} \left\langle \left[ \begin{array}{c} 2 \\ \text{PHON} \langle \text{on the desk} \rangle \end{array} \right] \left[ \begin{array}{c} 2 \\ \text{PHON} \langle \text{this book} \rangle \end{array} \right] \dots \right\rangle \right]$$

A crucial assumption of our analysis is that in embedded clauses the *wh*-phrase behaves just like a complementizer in terms of the distributional possibilities (basically similar to traditional observations such as that of Chomsky and Lasnik

<sup>12</sup>As an anonymous reviewer points out, in Old English a *wh*-phrase and a complementizer can occur together.

- (i) a. I wonder which dish that they picked.  
 b. They wonder what had John done.  
 c. Who did he hope would he see?

In the present framework this implies that in Old English the two elements have different domain index values with the *wh*-phrase preceding the complementizer.

Another point the anonymous reviewer made is about the fact that the complementizer also competes with an inverted auxiliary verb:

- (ii) a. \*I wonder which dish that did they pick. b. \*Who did he hope that would he see?

The present system follows Kathol's (2002) analysis in which the complementizer and inverted auxiliary verbs are assigned the identical index number. A parallel fact can be found in German too, which basically motivated a DOMAIN analysis for German clausal structures.

(1977)). This assumption can bring us another desirable consequence. In English, topicalization or negative SAI inversion is possible within an embedded clause only when a complementizer like *that* exists (cf. Ginzburg and Sag 2000 and others):

- (35) a. She subtly suggested \*(that) [problems of this sort, our analysis would never account for].  
 b. They believed \*(that) [never again would we have to do housework].

As noted by Ginzburg and Sag (2000), verbs like *suggest* select a nonindependent clause ([IC –]) whose value is originated from the complementizer *that*. Meanwhile, as given in (36)b, the complementizer *that* itself selects a finite S with this IC feature unspecified and as the result generates a noindependent clause:

- (36) a. 
$$\left[ \begin{array}{l} \langle \text{suggest} \rangle \\ \text{HEAD} \left[ \begin{array}{l} \text{verb} \\ \text{VFORM fin} \end{array} \right] \\ \text{ARG-ST} \left\langle \left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} \text{verb} \\ \text{IC –} \end{array} \right] \right] \right\rangle \\ \text{SUBJ} \langle \quad \rangle \end{array} \right\rangle \end{array} \right]$$
  
 b. 
$$\left[ \begin{array}{l} \langle \text{that} \rangle \\ \text{HEAD} \left[ \begin{array}{l} \text{comp} \\ \text{IC –} \\ \text{VFORM fin} \end{array} \right] \\ \text{ARG-ST} \left\langle \text{S} \left[ \begin{array}{l} \text{VFORM fin} \\ \text{SUBJ} \langle \quad \rangle \end{array} \right] \right\rangle \end{array} \right]$$

The absence of *that* in (35) thus means the violation of the verb's subcategorization requirements: With the absence of *that*, the verb *suggest* would combine with an inverted independent clause ([IC +]) even though it lexically selects a dependent clause ([IC –]) as given in its lexical specification (36)a.

The negative SAI construction is possible in an embedded clause only when either the complementizer *that* or a *wh*-phrase occupies the initial marker field:

- (37) a. I wonder [on which table] [only with great difficulty] would she put the big rock.  
 b. \*I wonder [[only with great difficulty] [would she put the big rock on the table]].

Just like the verb *suggest*, we simply need to say that *wonder* select an [IC –] indirect question. (37)b is thus unacceptable since the verb combines with an [IC +] clause: it violates the subcategorization requirements.

One thing to notice here is that, as noted in Culicover (1996), with particular prosodic satisfactions, multiple topics could be possible too:

- (38) a. (??) To that man, liberty, we would never grant.  
 b. (??) They told me that to that man, liberty, we would never grant.

To allow such examples, a solution we could adopt is to assume that a context would allow not to apply the Uniqueness Condition in (31) to the topic index '2', hence making the condition apply only to the index values 1 and 3.

There remains one contrast we need to account for: In embedded clauses, two *wh*-phrases can occur when the second one functions as the subject:

- (39) a. We have to figure out which hotels which clients should be assigned to \_\_  
 .  
 b. Which man knows where which woman will live? (Pesetsky (1987))

However, in root clauses, two consecutive *wh*-phrases are not possible even if the second one functions as the subject as in (40):

- (40) a. \*To whom, who gave the book \_\_ ?  
 b. \*Where which woman will live?

It appears that root and embedded clauses behave differently with respect to the distribution of *wh*-phrases.<sup>13</sup> In the present system, this asymmetry again follows easily. Since the two *wh*-phrases are both assigned the domain index value '3', the sentences violate the Uniqueness Condition in (31). However, the situation is slightly different in embedded clauses in which the first *wh*-phrase can function as a kind of complementizer with the domain index '1' as we have noticed. This in turn means that *which hotels* in (39)a bears the domain index '1' whereas *which clients* carries the index number '3'. Thus there is no violation of the Uniqueness Condition in embedded clauses.

### 3 Extraction from an Embedded Clause

Another striking asymmetry is involved with the extraction from a negative SAI sentences. As shown in (41) and (42), the extraction out of the embedded clause is possible only in the negative SAI (Culicover 1991):

- (41) a. \*These are the books which Lee says that [only with great difficulty], she can carry \_\_ .  
 b. These are the books which Lee says that [only with great difficulty], can she carry \_\_ .  
 (42) a. \*Which books did Lee say that [only with great difficulty], she can carry \_\_ ?  
 b. Which books did Lee say that [only with great difficulty] can she carry \_\_ ?

One possible way of explaining this asymmetry is to assume that *only with great difficulty* in (41)b and (42)b is a base-generated adjunct modifying the inverted clause, whereas the phrase in (41)a and (42)a is a topicalized phrase. This then can

<sup>13</sup>As an anonymous reviewer pointed out, one could ascribe the ungrammaticality of these sentences to superiority effects.

attribute the unacceptability of (41)a and (42)a to the constraint that nothing can move out of a topic clause. Such a line of approach could get strong support from the scope ambiguity in data like (43) (Rochemont 1986):

- (43) a. [With no job [would John be happy]]. (adjunction)  
 b. [[With no job] [John would be happy \_]]. (topicalization)

Example (43)a means that there is no job that would make John happy, while (43)b means that John would be happy even without having a job. The scope difference could be captured with the assumption that the PP in (43)a is a base-generated adjunct (having the wide scope) while the one in (43)b is extracted with the narrow scope reading.

However, one difficulty such a base-generated adjunction approach may have is examples like (44) where the operator *not only these books* is a complement:

- (44) a. \*[On which table] did Lee say that [these books] [she will put \_ ]?  
 b. [On which table] did Lee say that [not only these books [would she put \_ ]]?

The base-generated adjunction approach then has to assume that the complement *not only these books* in (44)b is also a phrase adjoined to the SAI, even though it is undoubtedly the complement of the verb *put*. This is quite untraditional wisdom.

We have seen that only the negative SAI operator (combining with an SAI sentence) functions as focus and gets the index number 3, whereas a negative expression combining with a noninverted S is still a topic with the index number 2. The generalizations we can draw from (41) and (42) can be represented as the following:

- (45) a. \*XP<sub>i</sub> CP[that Topic \_ i...]  
 b. XP<sub>i</sub> CP[that Focus \_ i...]

As noted in Rochemont (1989), Culicover (1991), Browning (1996) and others, English seems to observe topic islands when the embedded clause is headed not by a *wh*-phrase rather but by the complementizer *that*. Further data like (46) attest this:

- (46) a. \*Which books<sub>i</sub> did Lee say [that [on the table]<sub>j</sub> she will put \_ i \_ j]?  
 b. \*These are the books which<sub>i</sub> Lee says [that [to Robin]<sub>j</sub>, she will give \_ j \_ i].

One possible way to tackle such a fact seems to resort to the peculiarities of the complementizer *that* as in *that*-trace effect. In the present context, we interpret this as a constructional constraint on a CP headed by a complementizer as in (47):<sup>14</sup>

<sup>14</sup>The phrase *cp-topic-cl* is a subtype of *cp-cl* assumed in Ginzburg and Sag (2000).

(47) *cp-top-cl*:

$$\text{CP}[\text{GAP} \langle \quad \rangle] \rightarrow \mathbf{H}[\text{HEAD } \textit{comp}], \quad [\textit{topic-cl}]$$

The constraint in (47) simply tells us that when a *topic-cl* serves as the complement of a complementizer, the resulting CP contains no empty element. This simple constraint is enough to account for the extraction asymmetries provided in (41) and (42). For example, (42)a is unacceptable since the CP clause *[[that][only with great difficulty she can carry \_ ]]* has a nonempty GAP value though as indicated in (46) its GAP value should be empty.

## 4 Other Alternatives

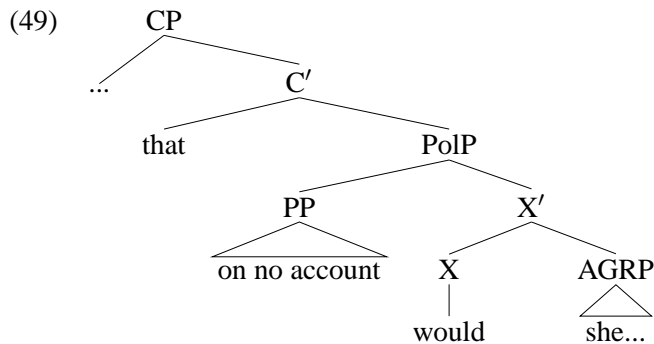
### 4.1 Brief Comparison with Configurational Approaches

The prevailing accounts of the ordering restrictions among left peripheral elements have been provided in terms of purely configurational perspectives or constructional properties. The dominant perspective is to posit hierarchical functional projections with the mechanism of movement as set forth by Culicover (1996), Rizzi (1997), Haegeman (2000), among others.

Literature have observed that the possibility of inverting a negative operator in embedded clauses as in (48) makes a single CP analysis unsatisfactory. For example, consider the examples in (48):

- (48) a. She said that on no account would she go there.  
 b. \*She asked me under what circumstances would I go there.

Since *that* in (48)a occupies the C position, an additional position is required to accommodate the sentence-peripheral constituent *on no account*. In addition (48)b indicates that negative and interrogative inversion both behave differently. One option that Culicover (1991, 1996) has taken is to introduce an independent functional projection PolP as roughly represented in (49):



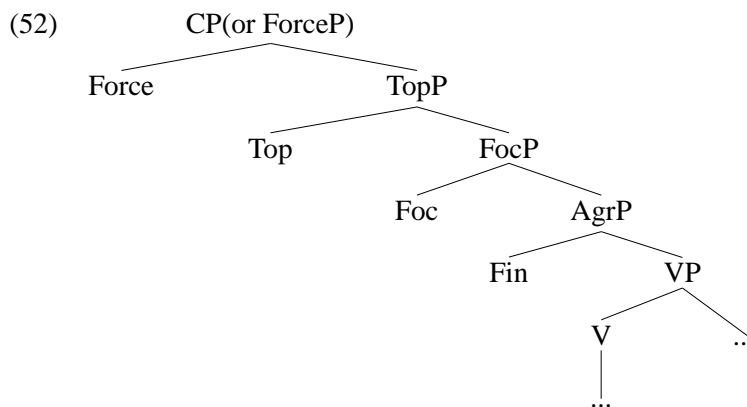
However, as noted earlier, since we have cases with more than one peripheral elements, the postulation of one additional functional projection is not enough:

- (50) a. I promise that [during the holidays] [on no account] will I write a paper.  
 b. \*I promise that [on no account] [during the holidays] will I write a paper.

The existence of examples like this has led to the accounts that treat embedded topicalization as an operation of XP adjunction to IP (cf. Baltin 1982, Lasnik and Saito 1992, and Rochemont 1986, among others). However, as noted in Haegeman (2000), if topicalization is a simple adjunction to a maximal projection, this adjunction approach would not block examples in (50)b as well as those like (51) at all.<sup>15</sup>

- (51) \*I promise that on no account will during the holidays I write a paper.

As a way of opening more positions, Rizzi (1997) and Haegeman (2000) have in a sense introduced the ‘Split CP hypothesis’ as roughly represented in (52):



Such a configurational ordering could capture certain ordering constraints between topic and focus together with the head movement triggering criterion given in (53):

- (53) The Wh & Neg Criterion: (Rizzi 1997, Haegeman 2000, Haegeman 2002)  
 A *wh*-operator/a neg operator must be in a spec-head configuration with a [+Wh/+neg]-X<sup>0</sup>

Though such a movement-based analysis sketched so far is appealing, it seems to require additional assumptions to account for the positional possibilities among left-peripheral elements as well as the contrast between embedded and root clause asymmetries. The key assumptions that an analysis like that of Haegeman and Guéron (1999) or that of Haegeman (2000) adopts could be summarized as follow (also see Rizzi 1997):

<sup>15</sup>If following Chomsky (1986) in which Aux moves to I to C then, forming a CP projection for the negative SAI sentence, a topic phrase like *during the holidays* in (50)a has to be attached not to an IP but to a CP. This then could not account for the presence of *that* here.

- Assumptions for Simple Cases:
  1. Root clause: The *wh*-phrase moves to the SpecCP. The head *wh*-feature is base generated within IP and movement of the tensed Aux to C will create the necessary spec-head configuration in generating sentences like ‘*What will you do about it?*’
  2. Embedded clause: The *wh* feature is base-generated on a head within CP (selected by the matrix predicate). Fronting the *wh*-phrase to the specifier of the relevant head creates the appropriate spec-head relation. This then would generate examples like *I wonder what you will do about it.*<sup>16</sup>
- Assumptions for Neg-operator SAI cases
  1. Root clause: The Neg-operator moves to the CP domain and the Aux *did* moves to C, creating the required spec-head relation. This would generate examples like *Not a single paper did he read.*
  2. Embedded clause: Since the embedded negative clauses are not selected by a matrix predicate, the NEG-feature is based generated on T whereas the *wh* feature is generated on C. This will allow cases like *I wonder [on which table] [only with great difficulty] would she put the big rock.*
- Assumptions for complex cases:
  1. Root clauses: Topic moves to SpecTopicP, *wh-ph* moves to SpecFoc, and the Aux to Foc. This would then generate the Topic-Wh phrase sequence as in (54):
 

(54) a. During the vacation, for what kind of jobs would you go into the office?  
       b. During the vacation, on no account would I go into the office.

In addition, a Neg operator and a *wh*-phrase target the same projection, blocking examples like the following:

(55) a. \*In no way, why would Robin volunteer?  
       b. \*Why in no way would robin volunteer.

  2. Embedded clauses: The highest head of the CP (Force) is associated with *wh*, generating the sequence of Wh-Topic as in *I wonder to whom this book Bill should give.* Unlike the *wh* phrase, Neg sentences are not selected: this makes the *wh* phrase targets CP whereas the negative SAI targets FocP. Thus the grammar generates cases like *I wonder on which table only with great difficulty would she put the big rock.*

---

<sup>16</sup>There exists still a need to block Aux from moving to C in embedded clauses.



Even though the assumptions give above, together with the hierarchically organized functional projections and movement operations, can state various distributional properties of left peripheral elements, further elaborations seem to be required if we look into further data. For example, let us consider the following examples:

- (56) a. \*To whom would [a book like this] you send?  
 b. \*Never in my life will, [beans], I eat.  
 c. \*I stress that on no account will, [during the vacation], I go into the office.  
 d. \*For which jobs would [during the vacation] you go into the office?  
 (Haegeman 2000)

All these examples would observe the Neg or Wh-criterion since the auxiliary verb immediately follows either a *wh*-phrase or a neg-operator. For example, in (56)d, one option would be to assume that the topic *during the vacation* somehow blocks the movement of *would*. However, this is quite unusual in derivational perspectives in which only a lexical element can block movement of a head lexical element. Or one could claim that no feature triggers the auxiliary verb to move. It appears that when more elements are involved in the left-peripheral constructions, additional assumptions are unavoidable: for example, Haegeman and Guéron (1999) introduces the feature TOPIC to the head of TopicP and assume that this feature blocks the head movement of Aux to a higher phrase.

As we have seen, the analysis presented here requires no additional machinery for such cases. Cases like (56) are simply predicted from the independently motivated constructional properties of inverted clauses. Adopting idea of Ginzburg and Sag (2000), we accept the view that English has the construction of *sai-ph* whose constraints are given in (57):

(57) *sai-ph*:

$$\left[ \text{SUBJ} \langle \quad \rangle \right] \rightarrow \mathbf{H} \begin{bmatrix} \text{INV} + \\ \text{AUX} + \\ \text{SUBJ} \langle \boxed{\text{I}} \rangle \\ \text{COMPS} \boxed{\text{A}} \end{bmatrix}, \boxed{\text{I}}, \boxed{\text{A}}$$

Given this, we can easily see why the examples in (56) are ill-formed: the filler, topic is between the head auxiliary and the subject, which is not licensed by this constraint at all.

We could not do all justice to the derivational analyses here, but it appears to be clear that such analyses require a series of hierarchically-ordered functional projections as well as constraints on the movement operations to generate the acceptable ordering relations among left peripheral elements. Meanwhile, our analysis relying on the notion of topological fields is much simpler in that it just assigns the

domain index numbers to the relevant elements based on not an arbitrary, but common assumptions that (i) topic precedes focus and (ii) in root clauses *wh*-phrases and complementizers positionally compete with each other.

## 4.2 Lexical and Constructional Constraints

A similar question arises whether lexical properties can tell us all the ordering restrictions among peripheral elements. For example, the *wh* initial position in the embedded clause is independently required since it is selected by the higher verb as represented in the following:

- (58) a. I wonder who John met last night.  
b. \*I wonder that John met Bill last night.

However, it seems that constructional constraints are also imposing restrictions on the ordering restrictions. For example, nothing lexically seems to restrict the sequence between a relative *wh*-phrase and a topic phrase as in (59):

- (59) a. the man [to whom] [liberty] we could never grant  
b. \*the man [liberty] [to whom] we could never grant.

Even if the relative *wh*-phrase *to whom* is not selected by a head, it should be in the initial position. We could not simply rule out the sequence of Topic and Wh phrase since this is what we find in root clauses as in (60):

- (60) This book, to whom should Bill give?

One could still question that the asymmetries we have discussed here could be followed directly from structural properties of the constructions involved. If we look into the main contrasts between root and embedded clauses in detail which we summarized in (61), we could see that structural properties do not give us all the answers:

	Root clause	Embedded clause
(61) a.	*Wh-ph – Topic – Neg-Op	Wh-ph – Topic – Neg-Op
b.	Topic – Wh-ph	*Topic – Wh-ph
c.	*Wh-ph – Neg-Op	Wh-ph – Neg-Op

One option to capture the contrasts would be to resort to constructional constraints in each case. For example, to capture (61)a, one could assume that in root clauses a *wh*-clause cannot combine with a topic clause. However, this restriction cannot be applied in embedded clauses since there is nothing wrong with this ordering in an embedded clause. We could simply say that in a root clause a *wh-ph* cannot combine with a Neg-operator whereas this is a possible ordering in an embedded

clause. It appears that no structural properties give us reason why English allows the Topic and Wh-phrase only in main clauses whereas the Neg-Op sequence only in embedded clauses, even if part of the distributional properties could follow from the relevant structural properties.

It seems at least to us that such a purely constructional analysis requires a more complicated grammar than the present analysis with the notion of DOMAIN. The point this paper tries to make is that the clear contrasts in the ordering relations of left-peripheral elements between root and embedded clauses are closely related with the tight interactions among various grammatical components including the domain order. That is, lexical, constructional, and topological fields all play their own roles in determining the ordering possibilities of left peripheral elements.

## 5 Conclusion

The main goal of this paper has been to explore how to analyze the word order asymmetries in the left peripheral constructions under the framework of HPSG. In particular, the distributional possibilities of left-peripheral elements like topic phrases, *wh*-phrases, and negative operators (introducing an SAI) could not be spelled out in a simple manner. This paper has explored a simple way of capturing such intriguing properties by introducing the notion of topological fields (DOMAIN) proposed by Kathol (2000, 2001). This DOMAIN-based approach just needs two traditional views: (i) a topic precedes a focus element, and (ii) in English a *wh*-element and a complementizer competes with each other for a position. The analysis provided here appears to serve as a plausible alternative to movement approaches that rely on movement processes together with multiple functional projections.

## References

- Authier, M. 1992. Iterated CPs and embedded topicalization. *Linguistic Inquiry* 23, 329–336.
- Baltin, M. 1982. A landing site for movement rules. *Linguistic Inquiry*, 13: 1–38.
- Browning, Margaret. 1996. CP recursion and *that*-t effects. *Linguistic Inquiry*, 27: 237–256.
- Chomsky, N. 1977. On *wh*-movement, In Adrian Akmajian et al. (eds.), *Formal Syntax*. New York: Academic Press, 71–132.
- Chomsky, N. 1986. *Knowledge of Language*. New York: Praeger.
- Chomsky, N. and Lasnik, H. 1977. Filters and control. *Linguistic Inquiry* 8: 425–504.
- Culicover, P. 1991. Topicalization, inversion, and complementizers in English. OTS Working papers.

- Culicover, Peter. 1993. Evidence against ECP accounts of the *that*-t effect. *Linguistic Inquiry* 24: 557–561.
- Culicover, Peter. 1996. On distinguishing A' movements. *Linguistic Inquiry* 27, 445–463.
- Ginzburg, Jonathan and Ivan Sag. 2000. *Interrogative Investigations*. CSLI Publications, Stanford.
- Green, Georgia. 1976. Main clause phenomena in Subordinate Clauses. *Language*, 52: 382–397.
- Haegeman, L. and J. Guéron. 1999. *English Grammar: A Generative Approach*. Blackwell Publishing.
- Haegeman, L. 2000. Inversion, Non-adjacent inversion and adjuncts in CP. *Transactions of the Philological Society*, 98:1, 121–160.
- Haegeman, L. 2002. Negative inversion, the NEG-criterion and the structure of CP. In Larry Horn and Yasuhiko Kato (eds.), *Studies in the Syntax and Semantics of Negation*. Oxford: Oxford Univ. Prss.
- Hooper, J. and S. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4: 459–497.
- Kathol, A. 2000. *Linear Syntax*. Oxford: Oxford University Press.
- Kathol, A. 2001. Positional effects in a monostratal grammar of German. *Journal of Linguistics*, 37: 35–66.
- Kathol, A. 2002. A Linearization-based approach to inversion and verb-second phenomena in English. In *Proceedings of the 2002 LSK International Summer Conference Volume II: Workshops on Complex Predicates, Inversion, and OT Phonology*. 223–234.
- Langendoen, T. 1979. More on locative-inversion and structure preserving hypothesis. *Linguistic Analysis* 5. 421–437.
- Lasnik, H. and M. Saito. 1992. *Move  $\alpha$* . Cambridge: MIT Press.
- Pesetsky, D. 1987. Language-particular processes and the earliness principle. Ms. <http://web.mit.edu/linguistics/www/pesetsky.home.htm>.
- Reape, M. 1994. Domain union and word order variation in German. German in Head-Driven Phrase Structure Grammar. CSLI Publications.
- Rizzi, L. 1997. The fine structure of the left periphery. In Haegeman, L. (ed.), *Elements of Syntax: A Handbook in Generate Grammar*. pp. 281–337. Dordrecht: Kluwer Publisher.
- Rochmont, M. 1986. *Focus in Generative Grammar*. John Benjamin.
- Rochmont, M. 1989. Topic islands and the Subjacency parameter. *Canadian Journal of Linguistics*, 34: 145–170.
- Watanabe, A. 1993. AGR-Based Case Theory and Its Interaction with the A-Bar System. *MIT Working Papers in Linguistics*. MIT: Cambridge.