NPI Licensing, Intervention and Discourse Representation Structures in HPSG

Manfred Sailer
Universität Göttingen

Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar
Stanford Department of Linguistics and CSLI’s LinGO Lab
Stefan Müller (Editor)
2007
CSLI Publications
pages 214–234

Abstract

Negative Polarity Items (NPI) are expressions such as English ever and lift a finger that only occur in sentences that are somehow “negative”. NPIs have puzzled linguists working in syntax, semantics and pragmatics, but no final conclusion as to which module of the grammar should be responsible for the licensing has been reached. Within HPSG interest in NPI has developed only relatively recently and is mainly inspired by the entailment-based approach of Ladusaw (1980) and Zwarts (1997). Since HPSG’s CONTENT value is a semantic representation, the integration of such a denotational theory cannot be done directly. Adopting Discourse Representation Theory (DRT, Kamp and Reyle (1993); von Genabith et al. (2004)) I show that it is possible to formulate a theory of NPI licensing that uses purely representational notions. In contrast to most other frameworks in semantics, DRT attributes theoretical significance to the representation of meaning, i.e. to a “logical form”, and not only to the denotation itself. This makes DRT particularly well-suited to my purpose.

1 Introduction

Negative Polarity Items (NPI) are expressions that only occur in sentences that are somehow “negative” (or “affective”, Klima (1964)). The typical examples for NPIs are English any and ever. NPIs have puzzled linguists working in syntax, semantics and pragmatics, but no final conclusion as to which module of the grammar should be responsible for the licensing has been reached. Within HPSG interest in NPI has developed only relatively recently with Tonhauser (2001) and Richter and Soehn (2006). While superficially very different, the two papers agree in many respects. In particular they both attempt to rebuild notions that stem from entailment-based theories of NPI licensing such as Ladusaw (1980) and Zwarts (1997). This theory is based on the denotation of the licensing contexts. Since HPSG’s CONTENT value is a semantic representation, the integration of such a denotational theory cannot be done directly. In the present paper I build on the earlier HPSG studies, but I show that it is possible to formulate a theory of NPI licensing that uses purely representational notions. For this enterprise, I adopt the framework of Discourse Representation Theory (DRT, Kamp and Reyle (1993); von Genabith et al. (2004)). In contrast to most other frameworks in semantics, DRT attributes theoretical significance to the representation of meaning, i.e. to a “logical form”, and not only to the denotation itself. This makes DRT particularly well-suited to my purpose.

† I presented parts of this paper at a workshop of the research network Constraint-based Grammar of the Deutsche Forschungsgemeinschaft in April 2007. I am grateful to the audience there and at the HPSG conference for their comments and discussion. I also thank Regine Eckardt, Janina Radó, Frank Richter, and Gert Webelhuth for their remarks and pointers concerning data and literature, and Garrett Hubing for proofreading this paper.
2 Data

I only consider two basic facts about NPIs that are commonly acknowledged in the literature, leaving many other aspects aside: the distinction between weak and strong NPIs and so-called intervention effects in NPI licensing. I also limit myself to NPIs in declarative clauses.

There are at least two types of NPIs, strong and weak NPIs (Zwarts, 1997). Within the considered contexts, strong NPIs can only occur in the scope of a clause-mate negation, as expressed in English with a negated auxiliary or an n-word (such as nobody), and in the restrictor of a universal quantifier. Weak NPIs are furthermore possible in the scope of expressions such as few N. If no such licenser is present in a sentence, both weak and strong NPIs are ungrammatical. Prototypical data are shown in (1) and (2).

(1) Distribution of a strong NPI:
   a. Pat won’t lift a finger to help me.
   b. Nobody will lift a finger to help me.
   c. Every student who lifts a finger will pass the exam.
   d. *Few students lifted a finger to help me.
   e. *Pat will lift a finger to help me.

(2) Distribution of a weak NPI:
   a. Pat didn’t budge during the experiment.
   b. Nobody bugged during the experiment.
   c. Every student who bugged during the experiment was excluded from further participation.
   d. Few students bugged during the experiment.
   e. *Pat bugged during the experiment.

The second observation that I discuss are so-called intervention effects: Consider the minimal pair in (3) for illustration. In a sentence with a negation, an NPI and a universal quantifier the universal quantifier may not take scope between the negation and the NPI. In (4) I sketch the three potential readings of the sentences in (3). The second formula expresses the unavailable ‘intervention’ reading. Given the word order in (3-b) this would be the most natural scope reading for this sentence. Since this reading is unavailable in some papers, for example Jackson (1995), sentences such as (3-b) are claimed to be ungrammatical.

(3) a. Kim didn’t give any apple to every teacher.
   b. ??Kim didn’t give every teacher any apple.

(i) a. I doubt that he will lift a finger to help me.
   b. Who will lift a finger to help clean up after the party?

1 There are more contexts that allow for strong NPIs, such as the complement clause of adversative predicates or rhetorical questions, to mention just two of these contexts.
(4)  a. $\neg \exists x [\text{apple}(x) \land \forall y [\text{teacher}(y) \rightarrow \text{give(kim, x, y) \}]]$

b. $\# \neg \forall y [\text{teacher}(y) \rightarrow \exists x [\text{apple}(y) \land \text{give(kim, x, y) \}]]$

c. $\forall y [\text{teacher}(y) \rightarrow \neg \exists x [\text{apple}(x) \land \text{give(kim, x, y) \}]]$

3 Previous Approaches

In this section I limit myself to a discussion of the entailment-based approach to NPI licensing and to three proposals that have been made within HPSG. I leave aside purely syntactic approaches such as Klima (1964) and Progovac (1994) as well as pragmatic approaches such as Krifka (1994) or Chierchia (2004) and the mixed approach of Linebarger (1980).

3.1 The Entailment-based Approach

The most influential theory in NPI research, Ladusaw (1980) and Zwarts (1997), states that NPIs must occur in a downward-entailing context, i.e. a context that allows inference from supersets to subsets. For strong NPIs this context must even be anti-additive, i.e. display an entailment behavior that is even closer to that of negation than simple downward-monotonicity. This theory captures the basic data in (1) and (2) correctly: Affirmative sentences are not downward entailing, thus (1-e) and (2-e) are predicted to be excluded. In all other sentences in (1) and (2) the NPI is in a downward-entailing context. The different types of downward-entailingness are needed to differentiate between strong and weak NPIs. The scope of few $N$ is downward entailing, but not anti-additive. Consequently, only weak NPIs are allowed here. It is the particular attractiveness of this account that it allows one to group the restrictor of the universal quantifier together with negation.

However, over the years the entailment-based account has faced a number of problems. One problem is that trivially downward-entailing contexts such as the one constituted by zero or more $N$ do not license NPIs. A problem that will be central to our discussion is that the theory does not account for the intervention effect illustrated in (3). Even in the unavailable reading (4-b) the context of the NPI is downward-entailing, notwithstanding the intervening universal quantifier (Jackson, 1995). In this case the entailment-based theory lacks the means to limit the domain of the licensing.

3.2 Scope Constraints: Tonhauser (2001)

Tonhauser (2001) attempts to encode an entailment-based theory of NPI licensing using a version of Minimal Recursion Semantics (MRS) in which potential licensors indicate the licensing strength of their scopal arguments. Tonhauser’s theory depends on properties of the MRS version of Copestake et al. (1997) which are not part of the published version (Copestake et al., 2005).
anti-additive. The lexical specification of an NPI includes a scoping constraint in its HANDLE-CONSTRAINTS-list that there must be some operator of the right strength that has scope over the NPI. Tonhauser’s theory shows the paradox of an HPSG rendering of entailment-based notions: When we look at the denotation of an expression, it is natural to talk about the entailments of that expression. In a representational framework such as HPSG, however, the entailment behavior has to be explicitly encoded in the structure. In the case of Tonhauser, this is done with otherwise unmotivated diacritic marking.

3.3 Collocations: Richter and Soehn (2006)

There are a number of collocational approaches within HPSG. The most recent and explicit one among them is Richter and Soehn (2006). I will focus on this approach in the present discussion and briefly address a second one, Sailer and Richter (2002) in the next subsection.

Richter and Soehn (2006) use Lexical Resource Semantics (LRS, Richter and Sailer (2004)) as their theory of semantic combinatorics. Just like MRS, LRS stands in the tradition of formalisms of underspecified semantics. In contrast to Tonhauser, however, Richter and Soehn do not include the NPI requirement in the constraints on semantic combinatorics but they treat them as collocational requirements, assuming a theory of collocation as employed in Soehn (2004) for idioms. This collocational treatment of NPIs has been put forward for example in van der Wouden (1997). Richter and Soehn use a feature COLL whose value indicates the lexically specified collocational restrictions of a sign. The COLL-list contains objects of sort barrier, which specify the syntactic domain within which the requirement must be fulfilled, such as the sentence or the utterance that contains the NPI. To give a concrete example, their lexical entry for the German NPI scheren (care) is given in (5). This NPI is licensed by any kind of licenser, as long as the licensing occurs in the same clause.

(5) Lexical entry of the German NPI scheren (care), adapted from Richter and Soehn (2006):

```
PHON ⟨scheren⟩
SYNS LOCAL CAT HEAD verb scheren CONT MAIN 1
COLL ⟨complete-clause LF-LIC EX-CONT⟩ & imp-op(2,1)
```

This lexical entry has a COLL specification that contains a complete-clause object. A general constraint in the grammar insures that the LF-LIC value of this object is identical to the logical form (i.e. the LF value) of the smallest complete clause that dominates the word.\(^4\)

\(^4\)In LRS a difference is made between the CONTENT value and the LF value. The former includes only the index and the main semantic constant contributed by a word, i.e. the information needed for
The lexical entry imposes a constraint on the logical form of this syntactic domain. It requires that its logical form contain the semantics of the NPI (the constant *scheren*) in the scope of a licensing operator. This is achieved by a number of relations for different types of semantic licensing domains. For illustration I give the definition of the relation downward-entailing-strength-operator in (6). Since downward-entailing contexts are a subgroup of anti-additive contexts, the relation de-str-op also holds if the stronger relation aa-str-op (anti-additive-strength-operator) holds.

(6) The definition of the relation de-str-op from Richter and Soehn (2006):

\[ \text{de-str-op}(\text{lf}, 1) \leftrightarrow \exists 2 \exists 3 \left( \begin{array}{l}
\exists 1 \prec 3 \text{ and } 2 \prec \text{lf} \text{ and } \\
2 \text{FEW}(\ldots, 3) \prec \text{lf} \text{ or } \\
2 \text{AT-MOST-n}(\ldots, 3) \prec \text{lf} \text{ or } \\
2 \text{HARDLY}(3) \prec \text{lf} \text{ or } \\
\ldots
\end{array} \right) \]

\[ \text{or aa-str-op}(\text{lf}, 1) \]

For the NPI *scheren* the semantic requirements are even weaker since this NPI may also occur in imperatives and in interrogative clauses. The example of the definition in (6) shows that while Richter and Soehn (2006) capture the mutual inclusion of the licensing contexts, every licensing operator is mentioned explicitly in the relation. The authors state that this explicit listing is made only for the sake of concreteness in the current absence of a better semantic generalization. In this sense the present paper can be seen as a step towards such a generalization for a core class of licensing contexts.

The strength of the approach in Richter and Soehn (2006) is its formal explicitness and the fact that it discusses a wide range different types of NPIs. On the other hand it fails to capture the unifying property of NPIs in an intuitive way: the fact that they are licensed under negation.

### 3.4 Decomposition: Sailer and Richter (2002)

The main concern of Sailer and Richter (2002) is to show that NPI-hood goes hand in hand with other collocational properties. In the paper we assume a collocational module similar to the one in Richter and Soehn (2006). It is only towards the end of the paper that we address the question of what kinds of semantic representations license NPIs. We speculate that all NPI-licensing contexts can be decomposed into logical forms that contain a negation. In the case of strong NPIs the NPI’s semantics must be in the immediate scope of the negation, in the case of weak NPIs, semantic operators may intervene. The discussion does not go beyond the sketchy representational reformulation of the entailment-properties of prototypical NPI-licensing contexts given in (7).
Representation of NPI-licensing contexts (Sailer and Richter, 2002):

<table>
<thead>
<tr>
<th>Entailment Classification</th>
<th>Example</th>
<th>If Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antimorphic</td>
<td>nicht (not)</td>
<td>¬[... φ ...]</td>
</tr>
<tr>
<td>Anti-additive</td>
<td>niemand (nobody)</td>
<td>¬∃[... φ ...]</td>
</tr>
<tr>
<td>Downward entailing</td>
<td>wenige (few)</td>
<td>¬many[... φ ...]</td>
</tr>
</tbody>
</table>

This approach is certainly very close to the ideas developed in the present paper. However, due to the choice of Ty2 (Gallin, 1975) as the underlying semantic representation language, the contexts did not cluster naturally. In the following section I will show that using DRT the intuitions behind the approach in Sailer and Richter (2002) can be expressed in a transparent way.

4 Discourse Representation Structures in HPSG

I assume that the content value of a sign is a Discourse Representation Structure (DRS, Kamp and Reyle (1993); von Genabith et al. (2004)). The use of DRT semantics within HPSG is not wide-spread, but has a number of predecessors, such as Frank and Reyle (1995), Eberle (1997), Holler (2003), Arnold (2004), Marshall and Sáfár (2004) to name just a few. My analysis does not depend on a particular choice of how to encode DRSs in HPSG. It is also independent of which combinatorial mechanism is used to arrive at the logical form of a complex sign.\(^5\)

I use the standard definitions for DRT, as they can be found in Kamp and Reyle (1993) or von Genabith et al. (2004). In this paper, space restrictions force me to use the so-called linear notation instead of the more traditional box notation. To give an example, the semantic representation that I assume for the sentence in (8-a) is given in box notation in (8-b) and in linear notation in (8-c).

(8) a. Kim gave an apple to a student.

\begin{align*}
\text{b. Box notation:} \\
&| e, x, y \\
&\text{apple}(x) \\
&\text{student}(y) \\
&\text{give}(e, \text{kim}, x, y)
\end{align*}

\begin{align*}
\text{c. Linear notation:} & [\{e, x, y | \text{apple}(x), \text{student}(y), \text{give}(e, \text{kim}, x, y)\}]
\end{align*}

A DRS consists of two parts: a set of variables, written at the top of the DRS, and a set of conditions. A variable that is introduced in a universe is interpreted as being existentially bound. The example illustrates that I use eventuality variables \((e)\) in addition to the variables for individuals.

In addition to simple DRSs, DRT allows for DRSs as part of conditions, i.e. in-

\(^5\)However, I have argued elsewhere that a combinatorial semantics that uses techniques of underspecification as used in LRS can lead to an elegant account of NPI licensing in negative raising constructions (Sailer, 2006). At the end of Section 7 I mention another potential advantage of LRS.
side the body of a DRS. Those are used for negation ($\neg K$), implication ($K \Rightarrow K'$) and disjunction ($K \text{ or } K'$). It should be noted that DRT treats negation, implication and universal quantification alike: A condition of the form $\neg K$ is equivalent to a condition of the form $K \Rightarrow \text{false}$. Expressions of the form every $N$ VP are translated into $[x|\phi] \Rightarrow \psi$, where $\phi$ and $\psi$ correspond to the translation of $N$ and VP respectively. In this paper, I will assume that the notation $\neg K$ is in fact just an abbreviation of $K \Rightarrow \text{false}$.

The basic DRT-language naturally extends to generalized quantifiers, using conditions of the form $Qx K_1 K_2$, for some determiner $Q$, some variable $x$ and DRSs $K_1$ and $K_2$ (von Genabith et al., 2004). This kind of representation is called a duplex condition, i.e. there are two DRSs, a restrictor and a scope, that are part of the condition. It has been emphasized in Partee (1988) that duplex conditions should be used for the representation of proportional determiners, whereas cardinal determiners will be treated as indefinites, introducing just a single DRS. From this point of view, conditions of the form $K \Rightarrow K'$ should also be considered duplex conditions. This is in line with Partee’s approach since the universal determiner is also proportional.

DRT uses a traditional notion of subexpression or component. In addition there is a notion of accessibility: A DRS $K$ is accessible from an expression $\phi$ iff (i) $\phi$ is a subexpression of $K$, or (ii) $K$ is the first DRS in a duplex condition and $\phi$ is the second DRS in this condition (i.e. there is a condition of the form $K \Rightarrow \phi$ or $QxK\phi$), (iii) or $\phi$ and $K$ are in the transitive closure of (i) and (ii). Accessibility is a central concept in DRT. To account for so-called donkey sentences as in (9), the interpretation of a variable occurrence is determined by the closest accessible DRS that contains this variable in its universe. In (9) the occurrence of the variable $x$ in the consequent of the implication is determined by the antecedent, since the DRS is accessible from the consequent and contains the variable in its universe.

(9) a. If a man called, he left a message.
   b. $[x\text{man}(x)] \Rightarrow [e\text{leave-message}(e, x)]$

As is common practice in DRT, I assume lexical decomposition. In particular, I decompose downward-entailing operators into a combination of a negation and an upward-entailing operator which is in the scope of the negation. This was proposed for example in Krahmer and Muskens (1995) for negative verbs such as lack and forget. Applying this to determiners, the negative indefinite determiner no is represented as a negation and an indefinite ($\neg[x|\ldots]$). Downward-entailing proportional quantifiers such as few are represented as $\neg[\emptyset|\text{many}xK_1K_2]$.

For a cardinal determiner $\text{Det}$ the truth of $\text{Det}(A)(B)$ only depends on the size of the set $A \cap B$. For a proportional determiner, it also depends on the size of the set $A$. Some is cardinal, since $\text{some}(A)(B)$ is true iff $|A \cap B| \neq 0$. Every is proportional, since $\text{every}(A)(B)$ is true iff $|A \cap B| = |A|$.

The determiners many and few have both a proportional and a cardinal reading. In the cardinal reading of the intersection of the restrictor and the scope must be above (resp. below) a contextually specified minimal value. In the proportional reading it must be above (below) a contextually specified minimal value.
5 A DRT-based Account of NPI Licensing

In (10) I sketch the DRSs for the sentences in (2). I use eventuality variables $e$ and $e'$, and I only mention the relevant conditions.

(10) Simplified DRSs for the sentences in (2):
   a. $[\emptyset \neg [e \text{budge}(e, \text{pat})]]$
   b. $[\emptyset \neg [x, e \text{budge}(e, x)]]$
   c. $[\emptyset [x, e \text{budge}(e, x)] \Rightarrow [e' : \text{be-excluded}(e', x)]]$
   d. $[\emptyset \neg [\forall x [x \text{student}(x)][e : \text{budge}(e, x)]]$}
   e. $*[e \text{budge}(e, \text{pat})]$

In (a) and (b) the semantics of the NPI is a condition in a sub-DRS of the negation. In (c) it occurs in the antecedent of a conditional DRS. I assume that these contexts are NPI-licensing DRSs. This notion is defined in (12).

(11) NPI-licensing DRS (first attempt):
    A DRS $K$ is an NPI-licensing DRS in a larger DRS $K'$ iff $K$ occurs in $K'$ as part of a condition of the form $\neg K$ or $K \Rightarrow K''$.

As mentioned above, negation in DRT is nothing but an implication with a contradiction in the consequent ($K \Rightarrow \text{false}$). This allows us to simplify the definition of an NPI-licensing DRS.

(12) NPI-licensing DRS (final version):
    A DRS $K$ is an NPI-licensing DRS in a larger DRS $K'$ iff $K$ occurs in $K'$ as part of a condition of the form $K \Rightarrow K''$.

I use this notion to express a necessary condition for the occurrence of NPIs: An NPI must always occur inside an NPI-licensing DRS. This condition is expressed in (13).

(13) General structural constraint on NPI licensing:
    The logical form of an NPI must be a subexpression of an NPI-licensing DRS.

   The sentences in (1-e) and (2-e) violate this constraint. The semantic representation of (2-e) is sketched in (10-e) above. Since it does not contain an NPI-licensing DRS, the NPI cannot be a subexpression of an NPI-licensing DRS.

---

percentage of the restrictor set. In this paper I focus on the proportional reading. These two readings of few lead to distinct entailment behavior in the restrictor. For both readings the scope is downward entailing, it is, however, only in the cardinal reading that the restrictor position is also downward entailing. Thus, the following entailment is not valid:

(i) Few (= a small percentage) people know Latin.
   $\models \neg \text{Few (= a small percentage) classicists know Latin.}$
Similarly, the NPI is not licensed in the scope of a universal quantifier, as shown in (14). While the DRS \( [x|\text{student}(x)] \) is an NPI-licensing DRS in this sentence, the DRS that contains the NPI is not a sub-DRS of this DRS.

(14) a. *Every student gives a damn about syntax.
   
   b. \( [\emptyset|x|\text{student}(x)] \Rightarrow [e|\text{give-a-damn}(e, x)] \)

In addition to this general structural constraint we also need special constraints for the different kinds of NPIs. If we compare the semantic representation of a sentence that contains an n-word, (10-b), with that of a sentence that contains a downward-entailing quantifier such as \textit{few}, (10-d), the latter contains an additional DRS with a non-empty universe that is accessible from the NPI. I will refer to accessible DRSs with a non-empty universe that intervene between an NPI and its licensing DRS as \textit{potential interveners}. This notion plays a central role in characterizing the difference between strong and weak NPIs. It is defined in (15).

(15) A DRS \( K \) is a potential intervener for an NPI \( \phi \) in a DRS \( K' \) iff

1. \( K \neq \phi \),
2. \( K' \) is an NPI licensing DRS that contains \( \phi \) and \( K \),
3. \( K \) is accessible from \( \phi \), and
4. \( K \) has a non-empty universe

I use this notion to express the different occurrence requirements of strong and weak NPIs.

(16) Special constraints:

a. Strong NPI: There is no potential intervener for the NPI in its NPI-licensing DRS.

b. Weak NPI: There is at most one potential intervener for the NPI in its NPI-licensing DRS.

To see the effect of these two special constraints we have to look at a context in which a weak NPI can occur but a strong NPI is excluded. The scope of \textit{few} was shown to be such a context. The DRSs for (1-d) and (2-d) are given in (17).

(17) a. *Few students lifted a finger to help me.
   
   \[ \emptyset|\neg[\emptyset|\text{many}x|x|\text{student}(x)] [e|\text{lift-a-finger}(e, x)] \]

b. Few students budged during the experiment.
   
   \[ \emptyset|\neg[\emptyset|\text{many}x|x|\text{student}(x)] [e|\text{budge}(e, x)] \] (\( = (10-d) \))

In both cases the structural constraint in (13) is met. The DRSs contain an NPI-licensing DRS: the scope of the negation (\( \neg[\emptyset|\text{many}x \ldots] \)). Inside this NPI-licensing DRS we find the DRS that contains the NPI (\([e|\text{lift-a-finger}(e, x)]\)) in

---

8To be precise: \( \phi \) is the smallest DRS that contains the semantic contribution of the NPI.
In both cases, the DRS that contains the NPI is the second DRS in the duplex condition introduced by the determiner many. The restrictor of the quantifier, the DRS \([x|\text{student}(x)]\), is a potential intervener. We can check the conditions in (15) to prove this. Condition 1: the restrictor and the scope of many are not the same DRS. Condition 2: the NPI-licensing DRS contains both the restrictor and the scope of many. Condition 3: the restrictor is accessible from the scope by the definition of accessibility. Condition 4: the restrictor has a non-empty universe since it contains the variable bound by the quantifier.

The special constraint on strong NPIs in (16-a) forbids the occurrence of a potential intervener. Consequently the semantic representation in (17-a) is not compatible with this restriction on strong NPIs. For weak NPIs, one potential intervener is allowed. Since the restrictor of many is the only potential intervener, the DRS does not violate the occurrence constraint on weak NPIs in (16-b).

While only weak NPIs are licensed in the scope of few, both weak and strong NPIs are licensed in negated sentences, in the scope of n-words and in the restrictor of every. DRSs of these types of sentences are given in (10-a), (10-b) and (10-c) respectively. In the first two cases the DRS that contains the NPI is the scope of the negation. Consequently, the DRS that contains the NPI is itself an NPI-licensing DRS. For this reason there cannot be an intervener between the DRS that contains the NPI and its licenser. This explains why both weak and strong NPIs can occur in these constructions.

We should also consider the case of an NPI in the restrictor of a universal quantifier as in (10-c). Structurally this case is identical to the one with a negation or an n-word. The DRS that contains the NPI is the antecedent of a conditional. Therefore it is itself an NPI-licensing DRS according to the definition in (12). This, again, explains why both weak and strong NPIs are possible in this position.

In this section I outlined the basics of a representational theory of NPI licensing. It was always considered one of the major achievements of the entailment-based theory that it accounted for the occurrence of NPIs under negation and in the restrictor of universal quantifiers by uniform means. The approach developed here shows that these contexts also form a natural class from a DRT perspective: the antecedent of an implication. In the next section I show how my account can capture intervention effects — which are problematic for entailment-based approaches.

6 Intervention Effects

Intervention effects seem to be the strongest argument for a structural, i.e. representational, theory of NPI licensing. For this reason they are also a test case of the present approach. First I illustrate that intervention effects with two quantifiers are immediately accounted for by the theory. Second I look at the case of an intervention effect with verbal negation and a quantifier such as in example (3). This latter case will require us to reconsider the DRSs for negated sentences.
The present approach elegantly captures intervention effects with two quantifiers. Sentence (18-a) cannot have a reading in which few takes scope over most which, in turn, scopes over any. I sketch the DRS of this unavailable reading in (18-b). I provide the box notation for better readability. I abbreviate the semantic contribution of the nouns and the verb with the corresponding upper case letters.

(18) a. Few students gave every teacher any apple.
    b. Hypothetical DRS of the unavailable reading:

The DRS \([z, e|A(z), G(e, x, z, y)]\) is the DRS that contains the NPI in this semantic representation. The NPI-licensing DRS is the scope of the negation. The NPI is contained in the scope of the universal which, in turn, is part of the scope of many. Since both the universal determiner and many contribute a restrictor with a non-empty universe, their restrictors are potential interveners for the NPI in the sense of (15). Consequently there are two potential interveners and the occurrence restriction of the weak NPI is violated. Strong NPIs are of course also excluded since they wouldn’t even allow for a single potential intervener.

This example shows that the DRT-based approach directly accounts for the intervention effects induced by the presence of two proportional determiners. Let us now turn to a more subtle type of intervention effect, induced by a verbal negation and a quantifier such as example (3). If we look at this example naively, it seems that the theory fails to prevent the weak NPI from occurring. In (19) I restate the unavailable reading of sentence (3-b) in DRT terms.

(19) Naive DRS for the excluded reading of (3-b):
    ?? Kim didn’t give every teacher any apple.
    \([\emptyset|\neg[\emptyset[y|\text{teacher}(y)] \Rightarrow [x, e|\text{apple}(x), \text{give}(e, \text{kim}, x, y)]]]\)

The NPI contributes the variable \(x\) and the DRS that contains the NPI is the scope of the universal quantifier \([x, e|\text{apple}(x), \text{give}(e, \text{kim}, x, y)]\). The NPI-licensing DRS is the scope of the negation \((\neg[\emptyset|\ldots])\). There is exactly one potential intervener between these two DRSs: the restrictor of the universal quantifier \([y|\text{teacher}(y)]\). In fact this constellation is analogous to the one of licensing in the scope of few (see (17-a)). For this reason we would expect that the weak NPI
is licensed in this reading and, consequently, fail to predict the intervention effect.

I will show that there is independent evidence that the DRS in (19) is not correct. I demonstrate that sentences of this form do in fact contain a second potential intervener. This additional potential intervener will be the NPI-licensing DRS itself. I will show that it contains a non-empty universe in cases of verbal negation. I concentrate on the following pair of sentences for my argument.

(20)  a. Not every visitor got a/any present.
    b. Every visitor didn’t get a/any present.

In their most natural readings both sentences in (20) are interpreted with the universal in the scope of the negation. For (20-a) this is the only possible reading, and the weak NPI is licensed. Matters are different, however, for sentence (20-b). While the reading with a wide scope of the negation, the so-called inverse scope reading is the preferred reading of this sentence if there is no NPI, this reading is unavailable if the NPI is in the sentence. In the presence of any, sentence (20-b) can only have the surface-scope reading, i.e. wide scope of the universal quantifier with respect to the negation.

The data in (20) illustrate that there is an intervention effect for the inverse scope reading of (20-b), but no intervention effect for (20-a). If the representational theory outlined in the previous section is on the right track, we should find evidence for an additional potential intervener for the inverse scope reading of (20-b). Such evidence comes from reference to abstract objects. I will show that there is an abstract discourse referent which is introduced between the negation and the universal quantifier.

Discourse referents introduced in the scope of negation are normally not accessible as antecedents for pronouns outside the scope of this negation (Kamp and Reyle, 1993), see (21-a). Such a pronominal reference is possible if there is a continuation with a modal or hypothetical context, as in (21-b). This modal subordination allows us to “skip” the outermost negation and gives access to discourse referents in its scope.

(21)  a. Pedro doesn’t own [a donkey],. He calls it, Emma.
    b. Pedro doesn’t own [a donkey]. He would call it, Emma.

To apply the same test to the data with universally quantified subjects, I use appositive which relative clauses.\(^9\) There, the relative pronoun typically refers to abstract entities from the main clause. With a continuation in the indicative, (22), there is no difference between the two antecedent clauses: which refers to the situation in which some visitors did not get presents.

(22) Every visitor didn’t get a present/ Not every visitor got a present,
    a. #which was very expensive. (which = every visitor got a present)
    b. which was a bit unfair. (which = some visitors didn’t get a present)

\(^9\)See Holler (2003) for a discussion of the corresponding type of sentences in German.
An irrealis continuation allows for modal subordination as in (21-b). For the sentence with a universally quantified subject and a verbal negation as (20-b), the best continuation refers to a situation in which every visitor actually received a present, i.e. (23-a). This continuation is unavailable for the sentence which has a subject of the form *not every* $N$ as shown in (24).

(23) Every visitor didn’t get a present, ...  
   a. which would have been very expensive.  
      (which = every visitor got a present)  
   b. ??which would have been a bit unfair.  
      (which = some visitors didn’t get a present)  

(24) Not every visitor got a present, ...  
   a. #which would have been very expensive.  
      (which = every visitor got a present.  
   b. ??which would have been a bit unfair.  
      (which = some visitors didn’t get a present)  

This contrast follows if we assume the presence of an abstract discourse referent, written as $p$, which can serve as the antecedent in (23). I refrain from committing myself to the concrete nature of $p$. It would be a state in classical DRT, a proposition in SDRT, or a situation in other variants. This referent is not present in (24). The resulting DRSs are shown in (25), where I abbreviate the semantic contribution of the nouns and verbs with capital letters.

(25) a. DRS for (23): $\neg\exists[p : [\emptyset] [x] V(x)] \supset [y, e | P(y), G(e, x, y)]]$  
   b. DRS for (24): $\neg\exists[p : [\emptyset] [x] V(x)] \supset [y, e | P(y), G(e, x, y)]]$

If modal subordination allows us to ignore the highest negation, the DRS in (25-a) provides an antecedent for *which*, but the DRS in (25-b) does not.

After this discussion we can come back to the original problem that sentence (20-b) cannot have an inverse-scope reading if the NPI is present. The DRS for the hypothetical inverse scope reading of sentence (20-b) is identical to the one in (25-a). In this DRS there are, now, two potential interveners for the NPI. First, the restrictor of the universal quantifier is an intervener, as we have seen above. In addition, the NPI-licensing DRS itself is a second intervener. To verify that this DRS is a potential intervener we have to go through the four conditions in the definition in (15). First, it is a DRS that is distinct from the DRS that contains the NPI. Second, it is also contained in the NPI-licensing DRS, since every DRS is contained within itself. Third, it is accessible from the DRS that contains the NPI, because the NPI is contained in a sub-DRS of the intervener. Fourth, the DRS has a non-empty universe: its universe contains the abstract discourse referent $p$.

This shows that our original theory from Section 5 accounts for the contrast in (20) under the independently motivated DRSs for negated sentences. We can now adapt the hypothetical DRS for the intervention reading from (19) accordingly.
This results in the following semantic structure.

(26) DRS for the excluded reading of (3-b):
?? Kim didn’t give every teacher any apple.
\[ \emptyset \neg [p | [y | \text{teacher}(y)] \Rightarrow [x, e | \text{apple}(x), \text{give}(e, \text{kim}, x, y)] \] ]

This DRS is analogous to the one in (25-a), i.e. the NPI is contained in a DRS for which there are two potential interveners in the overall semantic representation of the sentence: the restrictor of the universal and the scope of the negation.

In Section 5 I provided the basic definitions of a DRT-based theory of NPI licensing. This was enough to account for the basic data. In the present section I demonstrated that this theory is able to capture intervention effects directly.

7 HPSG Encoding of the Analysis

To integrate my analysis into HPSG, I follow Richter and Soehn (2006) in adopting a collocational approach to NPI licensing. I will focus exclusively on the NPI properties of the lexical items, leaving aside other collocational requirements they may have. My improvement over Richter and Soehn’s account lies in the uniform characterization of the licensers and in the fact that the intervention effects follow directly from the licensing conditions of the different types of NPIs.

I adopt the \textsc{coll} feature from Richter and Soehn (2006) as sketched in Section 3.3. If I ignore for the moment the syntactic domain within which particular NPIs need to be licensed (such as within the same clause as in (5)), there is a general principle of the grammar — the \textit{Licensing Principle} in Richter and Soehn (2006) — which guarantees that the \textsc{lf-lic} values on a word’s \textsc{coll} list are identical to the logical form of some sign that dominates the word.

(27) Licensing Principle:
In every unembedded sign \(s\), and for each lexical sign \(w\) in \(s\):

- every object on \(w\)’s \textsc{coll} value has an \textsc{lf-lic} value that is identical to the \textsc{cont} value of some sign \(s’\) that dominates \(w\) in \(s\).

We need relations that correspond to the notions \textit{NPI-licensing DRS} and \textit{potential intervener} as defined in (12) and (15) above. I assume RSRL (\textit{Relational SpeciMate Re-Entrant Language}, Richter et al. (1999) and Richter (2004)) as the underlying formalism of HPSG grammars. RSRL provides the use of relations and quantification over subcomponents of feature structures. This very expressive language allows us to define the necessary relations and to formulate the collocational constraints. The concrete definition of the relations depend on details of the HPSG encoding of DRSs. For this reason I will not provide these definitions here but give formal specification of the relations instead.

The HPSG encoding of DRSs comes along with a specification of the relations \textsc{component} (written as “\(\leq\)”), which holds of a pair \((k, k’)\) iff \(k\) is contained in
the DRS $k'$. The notion of accessibility used in DRT must also be translated into
the HPSG encoding. Here I assume a three-place relation accessible which
holds of a triple $⟨k, k', k''⟩$ iff $k$ is accessible from $k'$ within a larger DRS $k''$.

After these general relations that are needed for any integration of DRT into
HPSG I turn to the relations that are specific to the present theory. I start with the
relation $\text{npi-lic}$. It holds of a pair $⟨k, k'⟩$ iff $k$ is an NPI-licensing DRS in $k'$.

(28) Specification of the relation $\text{npi-lic}$:

The relation $\text{npi-lic}$ holds of a pair $⟨k, k'⟩$ iff there is some $k''$ such that $k' = k \Rightarrow k''$.

The second relation that is fundamental to my approach is a relation that iden-
tifies potential interveners. The relation $\text{pot(ential)-inter(vener)}$ holds
of a triple $⟨k, p, k'⟩$ iff $k$ is a potential intervener for the logical form of an NPI $p$
in a larger structure $k'$. The definition follows the conditions in (15).

(29) Specification of the relation $\text{pot(ential)-inter(vener)}$:

The relation $\text{pot-inter}$ holds of a triple $⟨k, p, k'⟩$ iff there is a DRS $k_p$
($k_p \leq k'$) which is the smallest DRS that contains $p$ and
1. $k \neq k_p$,
2. $k \leq k'$ and $k_p \leq k'$,
3. $⟨k, k_p, k'⟩ \in \text{accessible}$, and
4. $k$ has a non-empty universe.

Note that all these notions are defined purely in terms of the semantic repre-
sentation and do not refer the denotation.

With the help of these relations, we can formalize the lexical specifications
of a weak and a strong NPI schematically in (30) and (31). In both cases, $\Box$ is
the semantics of the NPI and $\Box$ is the semantics of a larger sign that contains the
NPI-licensing DRS $\Box$ for the NPI. The general structural constraint is expressed by
the line "$\text{npi-lic}(\Box) \& \Box \leq \Box$". The condition below this line expresses the
special constraint for weak NPIs in (30). Correspondingly, in (31) the line below
the general structural constraint is a direct rendering of the interpretive constraint
of strong NPIs.

(30) Schematic lexical specification of a weak NPI:

$$\begin{align*}
\text{SYNS LOC} \left[\text{CONT} \Box \right] & \& \exists \Box \left( \text{npi-lic}(\Box) \& \Box \leq \Box \right) \\
\text{COLL} \langle \text{LF-LIC} \ \Box \ drs \rangle & \& \exists \Box \left( \text{npi-lic}(\Box) \& \Box \leq \Box \right)
\end{align*}$$
Schematic lexical specification of a strong NPI:

\[
\left[ \text{SYNS LOC} \begin{bmatrix} \text{CONT} \end{bmatrix} \right] \& \exists \exists \left( \text{npi-lic}(\exists \exists) & \exists \leq \exists \right.
\left. \& \neg \exists \exists \left( \text{pot-inter}(\exists \exists, \exists \exists) \right) \right)
\]

The specifications in (30) and (31) are necessarily very schematic. It is known that NPIs show variation with respect to their licensing contexts. Since the theory developed in this paper encodes the licensing requirement as a lexical property of an NPI, it allows further restrictions on individual NPIs or a loosening of the restrictions for more permissive NPIs. At the same time, the schematic specifications exemplify the distinctions that are generally acknowledged to play a role in NPI licensing beyond finer idiosyncratic variation.

Let me address the issue of the syntactic domain of the NPI licensing. So far, I followed Richter and Soehn (2006) in this respect. A simpler theory would assume that it is enough to state that the collocational conditions must be met by the semantic representation of some sign that dominates the NPI. This simplification would still account for almost all the data reported by Richter and Soehn. The only exception are NPI verbs such as scheren (care) (see (5)). This verb is a weak NPI that requires a clausemate licenser. Here, the simplified theory would overgenerate.

Depending on how the syntax-semantics interface is handled, the simplification might be possible nonetheless. Klooster (1993) argues that weak clausebounded NPIs are typically verbs — he calls them Negative Polar Heads. In LRS, which is also the semantic framework assumed in Richter and Soehn (2006), verbs have a semantic specification that is identical to the logical form of the clause in which they occur. This is the EXTERNAL-CONTENT value. To account for the clause-boundedness effect of verbs it is enough to require that the LF-LIC value must be identical to their EX-CONT. This ensures that the NPI is licensed within its own clause. A schematic lexical entry of a Negative Polar Head is shown in (32).

Schematic lexical specification of a weak clausebounded NPI:

\[
\left[ \text{SYNS LOC} \begin{bmatrix} \text{CONT} \end{bmatrix} \right] \& \exists \exists \left( \text{npi-lic}(\exists \exists) & \exists \leq \exists \right.
\left. \& \neg \exists \exists \left( \text{pot-inter}(\exists \exists, \exists \exists) \right) \right)
\]

It seems that the syntactic component from Richter and Soehn (2006) can be removed if we combine the DRT-based theory with an LRS combinatorics. This also leads to a more restrictive theory: It predicts that verbal NPIs, but not nominal NPIs, can be weak and yet clause-bounded — because in LRS verbs, but not nouns, are assumed to have the same EX-CONT value as the clause in which they occur.

I showed how the DRT-based theory of NPI licensing can be formalized in HPSG using the COLL module. I pointed to some differences between my proposal and the one in Richter and Soehn (2006). It should be noted that the elimination of the syntactic domains relies on a particular framework of semantic combinatorics.
Consequently, it is only a side-remark in the present paper.

8 Conclusion

The integration of a theory of NPI licensing has to face two problems: first, how to characterize the licensing domain and second, how to encode the context requirement of an NPI inside its lexical entry. This paper attempts to make an original contribution to the first of these two questions, while building on an earlier HPSG analysis within a collocational framework for the second question.

DRT allows for a purely representational formulation of the contexts in which NPIs can occur. Instead of listing all NPI licensers individually or marking them explicitly as licensers, the decomposed semantic representation of the licensers is sufficient. Since licensers such as few introduce a negation and a quantifier, the occurrence constraints of NPIs immediately account for the fact that only weak NPIs are possible in such constellations. The constraints also capture the attested intervention effects. Future work has to show whether reasonable logical forms can be given for non-declarative sentences which allow for a natural extension of the present theory to NPI-licensing contexts such as interrogatives and imperatives.

Another extension would be to generalize the notion of an NPI-licensing DRS from the antecedent of an implicational condition \((K \Rightarrow K')\) to all first DRSs in a duplex condition. This would still capture the NPI licensing in the discussed contexts, but it would at the same time generalize to contexts which have been identified as highly problematic for entailment-based theories of NPI licensing. Israel (1995, 2004) shows that an NPI is licensed in the restrictor of a proportional determiner independent of its monotonicity properties if a rule-like interpretation of the sentence is possible. The following variation of an example with a universal quantifier from Heim (1984) can be used for illustration.

\[(33) \quad \text{[Most restaurants that charge as much as a dime for iceberg lettuce]}\]
\[\text{a. should be shut down.}\]
\[\text{b. *happen to have four stars in the handbook.}\]

In this example the strong NPI as much as is licensed in the restrictor of most, even though this position is not downward-entailing. In a DRT-based approach this context patterns naturally with the other NPI-licensing contexts: It is the first DRS in a duplex condition. While this certainly is a straightforward and promising extension of the present analysis, further research is needed to capture the contrast between the two different continuations in (33), i.e. the question of why the strong NPI is only felicitous in a rule-like statement but not in a more episodical statement.

The combination of DRT and HPSG has proven fruitful in a number of other papers quoted at the beginning of Section 4. The present discussion has shown that the independently motivated semantic representations assumed in DRT provide exactly the right structures and distinctions for a representational theory of
NPI licensing. Since HPSG’s linguistic objects contain semantic representations, DRT is a natural choice as a semantic formalism. Finally, the research on collocations carried out within HPSG can be put to use for an explicit encoding of NPI properties as distributional idiosyncrasies of individual lexical items.

References


