# **Respresenting Focus in LTAG**

Kata Balogh Institut für Sprache und Information Heinrich-Heine Universität Düsseldorf Katalin.Balogh@hhu.de

### Abstract

The paper proposes an LTAG semantic analysis to derive semantic representations for different focus constructions in a uniform way. The proposal is shown via examples of different narrow focus constructions, multiple foci and focus in questions.

## 1 Introduction

This paper proposes an analysis in Lexicalized Tree-Adjoining Grammar (LTAG) (Joshi and Schabes, 1997), that calculates the semantic representations of various focused sentences based on their syntactic structure and intonation pattern. The paper presents a proposal of extending the focus analysis of Balogh (2009) with the LTAG syntax-semantic interface from Kallmeyer & Romero (2008). Balogh (2009) provides a context-based approach of focusing, that gives a logical-semantic analysis of (narrow) focus constructions within the framework of Inquisitive Semantics (Groenendijk, 2009). One of the central claims of the analysis is that focusing leads to a special theme/rheme division of the utterance, that further relates it to the underlying context, and as such it regulates the coherent dialogue flow. This approach investigates the interpretation of focus from a semantic/pragmatic perspective, providing an analysis of phenomena as questionanswer congruence and the exhaustive interpretation of answers. However, the analysis lacks an important part - the syntax-semantics interface -, that builds the semantic representations as theme-rheme structures of natural language sentences driven by their syntactic structure.

## 1.1 Aims

The main aim of the current paper is twofold. It wants to broaden the coverage of linguistics analyses in LTAG and as a primary aim it wants to fill this gap of Balogh (2009) by proposing an analysis of the syntax-semantics interface that provides the semantic representations of the different kinds of focus constructions. These representations can further be interpreted according to the desired semantic/pragmatic framework: Inquisitive Semantics (InqS) (Groenendijk, 2009). The choice for the logical-semantic system of InqS as opposed to, e.g., Alternative Semantics (Rooth, 1992)<sup>1</sup> has several motivations. One of the main aims in favor of InqS is, that its semantics and dialogue management system offers an elegant way to analyze various discourse-related phenomena involving focus such as: focusing in answers, questionanswer relations, contrast in denial and specification by focusing. The analysis in this paper concentrates on narrow focus constructions, however a proposal of extending it to broad focus constructions and focus projection is also given.

## 2 Frameworks

The analysis proposed in this paper offers a compositional way to calculate the semantic representations for different (narrow) focus constructions in a uniform way. The analysis of the syntaxsemantics interface as introduced here is provided within the framework of Lexicalized Tree-Adjoining Grammar, LTAG (Joshi and Schabes, 1997) with a semantic component as developed by Kallmeyer & Joshi (2003) and Kallmeyer & Romero (2008).

<sup>&</sup>lt;sup>1</sup>As in the proposal from Babko-Malaya (2004), that integrates LTAG with Alternative Semantics.

## 2.1 LTAG Semantics

For the semantic representation I adapt the LTAG semantics based on unification as introduced by Kallmeyer & Romero (2008). In this approach each elementary tree comes with a featurestructure and a (flat) semantic representation, each of them consisting of a set of labelled propositions and a set of scope constraints. These propositions and constraints contain meta-variables of individuals, propositions or situations, all of them given by boxed numbers. The feature structures are all linked to a semantic representation and by substitution and adjunction of the trees, feature structures get unified and the meta-variables get values. Also the semantic representation of the resulting tree is calculated by taking the union of the representations of the participating trees. For an illustration of LTAG semantics see Example 2.1, the derivation of the question Who walks? assigning the semantic representation as  $\exists x.walk(x)$ .

Example 2.1 Who walks?



The S-tree of 'walks' comes with a semantic representation consisting of two propositions:  $l_0$  contributes the question-operator applied to a proposition given here as the meta-variable [4]. The proposition  $l_1$  says, that the predicate walk is applied to the individual variable [1] that is contributed by the NP-tree substituted to the given position: given by I = 1 on the feature structure of the substitution node. Here, two special features are introduced: WHMAX and WHMIN. These features are inspired by the idea of a wh-scope window from Romero & Kallmeyer & Babko-Malaya (2004) and by the MAXS and MINS features from Kallmeyer & Romero (2008) that indicate the scope window of a given quantificational phrase. The features MAXS and MINS determine the maximum and minimum scope of quantificational NPs such as 'someone' or 'everyone', while WHMAX and WHMIN indicate the scope window for (wh-)questions and for focus. Separating these two different scope windows has the advantage to account for, e.g., quantifying into questions. Next to the propositions  $l_0$  and  $l_1$ , the scope constrains  $|4| \geq |3|$ ,  $|3| \geq |8|$  are defined that determine the scope relations between the given propositions. The scope constrains are defined between the propositional meta-variables and the propositional labels.

The NP-tree of the wh-phrase gets substituted into the S-tree of 'walks' resulting in the equations 1 = x, 4 = 6, 3 = 7 and since nothing is adjoined at the VP node<sup>2</sup>, we have  $8 = l_1$ . After these equations the combination of the semantic representations results in:

$\begin{array}{l} l_{0}: ?[4], l_{1}: walk(x), l_{2}: \exists x \texttt{5} \\ 4 \geq \texttt{3}, \texttt{3} \geq l_{1}, \texttt{4} \geq l_{2}, \texttt{5} \geq \texttt{3} \end{array}$
$4 \ge 3, 3 \ge l_1, 4 \ge l_2, 5 \ge 3$

Following these scope constraints, the possible plugging is:  $4 \mapsto l_2, 5 \mapsto l_1, 3 \mapsto l_1$ , resulting in the semantic representation as  $\exists x.walks(x)$ .

### 2.2 Inquisitive Semantics

In the semantic representation I follow the language of Inquisitive Semantics, serving the broader purpose to integrate the current analysis with a component of semantic-pragmatic interpretation and discourse modeling (e.g. modeling question-answer relations). The semantic representation  $?\exists x.walks(x)$  is the translation of the wh-question *Who walks?* according to the logical system of InqS.

As already introduced before, in my analysis I adapt several ideas of the system of Inquisitive Semantics (Groenendijk, 2009). The main aim behind this framework is to create a logical system that models the flow of a coherent dialogue. The principal goal is to provide a model of information exchange as a cooperative process of raising and resolving issues. In the semantic interpretation of utterances, the main source of inquisitiveness is *disjunction*. The disjunction of two propositions

<sup>&</sup>lt;sup>2</sup>To keep the examples easier, none of the following examples contain adjunction at the VP node, so in later examples I will skip the P features at the VP and S nodes.

is naturally interpreted as providing the information that one of the two propositions is true and also raising the issue *which one* of them is true. Hence the disjunction  $p \lor q$  provides two *possibilites*: either p is true or q is true, while eliminating the option that both of them are false.

Consider now the meaning of a question. According to the classical theories<sup>3</sup>, the meaning of a question is the set of its (true/complete) answers. Hence the meaning of the polar question Is it raining? (?p) is identified by the set of two propositions it is raining (p) and it is not raining  $(\neg p)$ and the questioner wants to know which one of the two holds. Since the questioner is interested whether p or  $\neg p$  is the case, the question ?p can be defined as the disjunction of its two possibilities:  $p \lor \neg p$ , hence in general questions can be defined in terms of disjunction:  $?\phi = \phi \lor \neg \phi$ . The main conclusion that can be drawn here is that like questions, disjunctions have the characteristic of introducing possibilities, and they both get an alternative interpretation.

The system of IngS is developed in such a way that sentences can provide data (informativeness) and raise issues (inquisitiveness). In terms of these two notions three meaningful sentence types can be defined: (i) assertions: informative and not inquisitive, (ii) questions: inquisitive and not informative, and (iii) hybrids: informative and inquisitive. Such a hybrid type is the proposition  $p \lor q$ , that provides the information that  $\neg p \land \neg q$ is not the case, while it raises the issue which one of p or q is true, thus it gives two possibilities. The question  $?(p \lor q)$  is not informative, it does not exclude anything, it only raises the issue whether p or q or  $\neg p \land \neg q$  is the case (three possibilities).<sup>4</sup> Similarly to  $p \lor q$  the predicate logical expression  $\exists x.\phi$  also provides the information that  $\neg \exists x. \phi$  is not the case and additionally it raises the issue which individuals are  $\phi$ . It leads to several possibilities depending on the number of individuals in the domain. Take, for example, the proposition  $\exists x. P(x)$  and a small domain of three individuals  $D = \{a, b, c\}$ . The existential expression  $\exists x. P(x)$  then excludes the option that none of a, b, c is P, and raises the issue which one is P. Relative to the given domain

D, this expression leads to the set of three possibilities: P(a), P(b), P(c). Following this line, I assume the standard logical translation of a constituent question to be of the form  $\exists x.\phi$ . A constituent question is interpreted as a set of possibilities, corresponding to its possible answers. I give a Hamblin-style interpretation of questions as sets of propositions, however with the crucial difference that in my analysis the set contains the proposition *nobody is* P as well. The wh-question *Who walks*? is translated as  $\exists x.walk(x)$  which is the same as the disjunction of the propositions (possibilities)  $walk(d_1) \lor walk(d_2) \lor ... \lor walk(d_n) \lor \neg \exists x.came(x)$  relative to the given domain of individuals.

In the logical language of Inquisitive Semantics all utterances are claimed to be divided into a theme and a rheme, where the rheme corresponds to the information content of the given utterance and the theme to the issue that the utterance addresses. Balogh (2009) proposes an analysis of focused sentences claiming that focusing leads to a special theme-rheme division. Next to the parallelisms with the distinction of new and old information in the generative view, an important difference is that in this analysis the sentences itself are not split into two parts, but the way is defined how to signal the inherent issue (theme) of the utterance and the data it provides (rheme). The theme of an utterance is a question, and as such it is inquisitive, introducing two or more possibilities. In order to derive the special theme and rheme of a focused sentence Balogh (2009) defines the Rule of Division by focusing.

## **Definition 2.2 Rule of Division**

Let  $\alpha$  be an utterance in natural language,  $\alpha'$  the translation of  $\alpha$  in the language of InqS and  $\natural$  the operation:  $\varphi^{\natural} = \psi$  if  $\varphi = ?\psi$ , otherwise  $\varphi^{\natural} = \varphi$ .

Every utterance  $\alpha$  is divided into a theme and rheme:  $TH(\alpha)$ ;  $RH(\alpha)$  where

 $TH(\alpha) = ?\exists \vec{x} (\alpha' [\vec{a_F}'/\vec{x}])^{\natural}; and$  $RH(\alpha) = \alpha'$ 

This definition correctly derives the themerheme division of various narrow focus constructions, that further get interpreted in the system of InqS. This proposal provides a contextbased analysis of focusing with special attention to question-answer congruence, exhaustivity, contrast in denials, and specification. However, the system of Balogh (2009) lacks the syntax-

<sup>&</sup>lt;sup>3</sup>e.g. (Hamblin, 1973; Karttunen, 1977)

<sup>&</sup>lt;sup>4</sup>Note, that  $?\phi$  is not a separate category in the syntax of the logical language, but it is defined in terms of disjunction as given above.

semantics interface. As it can be seen in Definition 2.2, focus marking of constituents get directly translated in the logical language as  $\phi_F$  not referring to the syntactic structure and the contribution of the focused constituent.

The analysis proposed in this paper wants to fill this gap defining the syntax-semantics interface, that provide the correct semantic representation (theme-rheme pair) on basis of the syntactic structure of the utterance.

## **3** Proposal

<

As a starting point, the current analysis suggests the semantic representations of utterances consisting of two components: one that represents the theme and one that represents the rheme. According to this, each S-tree comes with a semantic representation as the following:

In this two-component representation the above part is the representation of the theme, while the below one is the representation of the rheme. Defined in this way all S-trees come with a semantic representation, where the theme will lead to a question: the *issue behind*, and the rheme leads to a proposition: the *semantic content*.

### 3.1 Narrow focus constructions

The representation of focusing first of all has to provide different structures for the different (narrow) focus constructions. Consider first the basic cases of a sentence with a transitive verb: (i) none of the arguments is focused, (ii) the subject is focused or (iii) the object is focused. All these sentences lead to different theme-rheme divisions:

(1) Pim likes Sam.  $\sim$  TH: ?*like*(*p*, *s*) + RH: *like*(*p*, *s*) PIM<sub>*F*</sub> likes Sam.  $\sim$  TH: ? $\exists x.like(x, s)$  + RH: *like*(*p*, *s*) Pim likes SAM<sub>*F*</sub>.  $\sim$  TH: ? $\exists y.like(p, y)$  + RH: *like*(*p*, *s*)

Take first the sentence *Pim likes Sam* that is built of three elementary trees, the S-tree of the verb and two NP-trees of the two arguments.

Example 3.1 Pim likes Sam



By substituting the NP-tree in the S-tree the features on the nodes get unified (thus 1 = x, 2 = y) and the corresponding semantic representations are combined, resulting in the semantic representation of the sentence as:

(	$ \begin{matrix} l_0 : ? \textcircled{4}, l_1 : like(x, y), l_2 : pim(x), \\ l_3 : sam(y), \textcircled{4} \ge \textcircled{3}, \textcircled{3} \ge l_1 \end{matrix} $	$ \rangle$
\	$l_1: like(x, y), l_2: pim(x), l_3: sam(y)$	/

There is one way of plugging possible here:  $[4] \mapsto l_1, [3] \mapsto l_1$ , that derives the semantics representation of the given sentence as the following, where the theme corresponds to the polar question *Does Pim like Sam?* and the rheme corresponds to the proposition *Pim likes Sam*.



#### 3.1.1 Subject / object in focus

Sentences consisting of a transitive verb have the possibilities of narrow focus: either the subject or the object (or both) can be focused. First, look at the sentences in (1) with single focus. The analysis derives the rheme as the proposition *Pim likes Sam* for both, while the different focus structures lead to two different themes corresponding to the inherent questions: *Who likes Sam?* and *Whom does Pim like?* respectively.

In the analysis of  $PIM_F$  likes Sam with narrow focus on the subject, we take the S-tree of 'likes' as above and substitute two NP-trees: for the nonfocused object the tree for 'Sam' as in Example 3.1, while for the focused subject we take the tree for 'Pim' with its special semantics:



The semantic representation of the focused NP contributes a special theme as an existential expression. The substitutions of the two NPs carried out and the respective meta-variables unified: 1 = x, 2 = y, 8 = 4, 7 = 3, that leads to the semantics:

$$\left\langle \begin{array}{c} l_{0}:?[4], l_{1}: like(x, y), l_{2}: \exists x \exists 1 ], \\ l_{3}: sam(y), 4 \geq \exists, \exists \geq l_{1}, \\ 4 \geq l_{2}, \exists 1 \geq 3 \end{array} \right\rangle$$

$$\overline{l_{1}: like(x, y), l_{2}: pim(x), l_{3}: sam(y)}$$

Again, one way of plugging is possible here:  $4 \mapsto l_2, 11 \mapsto l_1, 3 \mapsto l_1$  providing the twofold semantic representation corresponding to the question *Who likes Sam*? as the theme and the proposition *Pim likes Sam* as the rheme of the utterance.

$$\left( \begin{array}{c} ?\exists x.like(x,y), sam(y) \\ \hline like(x,y), pim(x), sam(y) \end{array} \right)$$

The analysis of *Pim likes*  $SAM_F$  is similar, we take the sam S-tree for the non-focused subject we substitute the tree for 'Pim' as before and for the focused object we substitute the tree for 'Sam' as:



The two substitutions here lead to the semantic representations before and after plugging:

$$\left\langle \begin{array}{c} l_{0}:?[4], l_{1}: like(x, y), l_{2}: pim(x), \\ l_{3}: \exists y \boxed{12}, 4 \geq \boxed{3}, \boxed{3} \geq l_{1}, \\ \hline 4 \geq l_{3}, \boxed{12} \geq \boxed{3} \\ \hline l_{1}: like(x, y), l_{2}: pim(x), l_{3}: sam(y) \\ \\ \left\langle \begin{array}{c} ?\exists y. like(x, y), pim(x) \\ \hline like(x, y), pim(x), sam(y) \end{array} \right\rangle \\ \end{array} \right\rangle$$

Similarly to the previous example, this representation corresponds to the question *Whom does Pim like?* as the theme and to the proposition *Pim likes Sam* as the rheme.

## 3.2 Multiple focus

After showing the basic cases, let us now turn to more complex examples such as multiple focus. In sentences formed of a transitive verb, not only single focusing is possible, but also both arguments can be focused in the same time:  $PIM_F$ *likes*  $SAM_F$ . The rheme or information content of this sentence is again the proposition *Pim likes* Sam, while the theme or underlying issue is the multiple wh-question Who likes whom? The analysis derives the correct theme-rheme division straightforwardly, by substituting the NP-trees of the focused arguments (see in the previous section) into the S-tree of 'likes' (see Example 3.1). The substitutions of the focused subject and object lead to the semantic representation:

$$\left\langle \begin{array}{c} l_{0}:?[4], l_{1}: like(x, y), l_{2}: \exists x \boxed{11}, \\ l_{3}: \exists y \boxed{12}, 4 \ge \boxed{3}, \boxed{3} \ge l_{1}, 4 \ge l_{2}, \\ \boxed{11 \ge \boxed{3}, 4 \ge l_{3}, \boxed{12} \ge \boxed{3}} \\ \hline l_{1}: like(x, y), l_{2}: pim(x), l_{3}: sam(y) \end{array} \right\rangle$$

Here, two different pluggings are possible: (i)  $[4] \mapsto l_2, [1] \mapsto l_3, [12] \mapsto l_1, [3] \mapsto l_1$  and (ii)  $[4] \mapsto l_3, [12] \mapsto l_2, [11] \mapsto l_1, [3] \mapsto l_1$ , yielding two semantic representations, where the representations of the theme are slightly different: at plugging (i)  $?\exists x \exists y. like(x, y)$  and at (ii)  $?\exists y \exists x. like(x, y)$ . Since we have existential expressions, these two representations are equivalent, both corresponding to the question *Who likes whom*?

#### **3.3** Focus in questions

The analysis proposed above gives also a straightforward derivation of a special construction, when an argument is focused within a wh-question as, e.g. 'Who likes  $SAM_F$ ?' Such examples appear in, e.g., answering strategies, where the goal is to resolve a question, which can be reached via answering all of its (easier) sub-questions. As Roberts (1996) shows, the question 'Who likes whom?' can be resolved by the strategy of replacing the original question with its sub-questions, where the sub-question is only felicitous if it is appropriately focused. For an illustration, consider the answering strategy of the multiple whquestion 'Who likes whom?':

(2) Who likes whom?Who likes SAM<sub>F</sub>?Does Sam like Sam?Does Tom like Sam?

In the derivation of 'Who likes  $SAM_F$ ?' we take the S-tree of 'likes' (example 3.2) and substitute the elementary trees of the wh-phrase 'who' and the focused object 'Sam' (example 3.3):



By substitution of the two NP-trees we get the equations 1 = x, 2 = y, 4 = 6 = 10, 3 = 7 = 9 on the theme side and 1 = x, 2 = y, 4 = 7 = 10

6, 3 = 7 on the rheme side. This gives us the still underspecified semantic representation:

[	$ \begin{array}{l} l_0: ?[4], l_1: like(x,y), l_2: \exists x \underline{5}, l_3: \exists y \underline{12} \\ 4 \geq \underline{3}, \underline{3} \geq l_0, 4 \geq l_2, \underline{5} \geq \underline{3}, \\ 4 \geq l_3, \underline{12} \geq \underline{3} \end{array} $	
\	$\begin{array}{l} l_0: ?[4], l_1: like(x,y), l_2: \exists x \boxed{5}, l_3: sam(y) \\ 4 \geq \boxed{3}, \boxed{3} \geq l_0, 4 \geq l_2, \\ 5 \geq \boxed{3} \end{array}$	/

Similarily to the example with multiple foci, at the theme side of this example two pluggings are possible, that derive the representations  $\exists x \exists y.like(x, y)$  and  $\exists y \exists x.like(x, y)$  that are equivalent. On the rheme side only one plugging is possible, that derives the representation  $\exists x.like(x, y), pim(y)$ . Hence, the analysis correctly derives the theme and the rheme of 'Who likes PIM<sub>F</sub>?' as the multiple wh-question 'Who likes whom?' and the single wh-question 'Who likes Pim?' respectively.

## 4 Conclusion and further issues

The approach introduced here is a proposal towards an analysis of focus constructions using LTAG with a unification based semantics. The analysis derives the theme/rheme divisions of different (narrow) focus constructions including multiple foci and focusing in questions.

(3) Pim likes Sam.

 $\sim$  TH: ?like(x, y), pim(x), sam(y) $\sim$  RH: like(x, y), pim(x), sam(y)

PIM<sub>F</sub> likes Sam.

PIM<sub>F</sub> likes SAM<sub>F</sub>.  $\sim$  TH: ? $\exists x \exists y. like(x, y)$  $\sim$  RH: like(x, y), pim(x), sam(y)

The advantage of this analysis is that all four sentences bear the same propositional content (rheme), while the different focus structures lead to different inherent issues (theme) indicating that these sentences are felicitous in different contexts. Consequently, they relate to four different whquestions, which offers a straightforward way to deal with the basic cases of question-answer congruence. This analysis follows the core ideas of the context-based approach of Balogh (2009), that concentrates merely on the interpretation of different focus structures. The above analysis provides an extension to the syntax-semantics interface of Kallmeyer & Romero (2008). It determines the semantic representations as assumed in Balogh (2009) on basis of the syntactic structures of the sentences in a straightforward, intuitive and compositional way.

Since this paper is a report of a work in progress, several loose ends can be pointed out. First of all an analysis of the relation of accent placement and focus has to be given to deal with, among others, the phenomenon of Focus Projection (Selkirk, 1996). The second important issue to investigate is the relation of focusing and quantifier scope as one of the main reasons of choosing LTAG as the framework of the syntax-semantics interface. The semantic component of LTAG as introduced by Kallmeyer & Romero (2008) offers an analysis of scope ambiguities. In their analysis scope windows are introduced for quantificational NPs by the features MAXS and MINS signalling the maximal and minimal scope sides. Focus and questions also bear scope properties, different from the scope properties of quantificational NPs. To offer a uniform analysis of the similarities and differences of these scope sides, this paper introduces the features WHMAX and WH-MIN as the scope window for focus and questions (inspired by Romero & Kallmeyer & Babako-Malaya (2004)). The distinction of the two different scope windows gives the possibility to deal with the relation of focusing and quantifiers as well as quantifying into questions.

## 4.1 Focus marking and accenting

In section 2, the proposal of the analysis of narrow focus constructions was introduced, deriving a two-fold semantics of utterances representing the theme (underlying issue) and the rheme (propositional content). Focused constituents contribute a special semantics to the theme of the sentence meaning, yielding the corresponding whquestion. Each elementary tree of a focused constituent came with a different semantic representation as their non-focused counterpart. Focus marking can be signaled within the feature structure of the given elementary tree, introducing the feature FOC with possible values + and - for focused and non-focused occurences. In the previous examples all NP arguments are proper names with an elementary tree of a noun phrase without further inner structure and the focus feature can appear at the maximal projection. However, for an elegant account for focusing we need to be able to give an analysis of the placement of the pitch accent and the focus marking. Hence, we have to account for Focus Projection as well as focus marking within a complex NP. Following Selkirk's (1996) Focus Projection principle, the same accenting can receive different focus marking, hence different focus interpretation. As her focus marking principles suggest, pitch accent on the noun can lead to a narrow focus interpretation or to a broadad (VP) focus interpretation:

- (4) a. John rented [a BICYCLE]<sub>F</sub>.
  - b. John [rented a BICYCLE] $_F$ .

An important issue for the current approach is, how to analyze the relation between the placement of the pitch accent and the marking of the focused constituent. For this we need to introduce two features FOC and PITCH that stand for focus marking and accenting respectively. The placement of the pitch accent is given by the feature pitch = +coming from the lexicon together with the lexical anchor. The value of the pitch accent is then passed to the FOC feature that appears on some nodes of the elementary tree of the noun phrase.

$$\begin{array}{c} \operatorname{NP}_{foc=1} \\ \downarrow \\ \operatorname{N} \diamond_{foc=1, pitch=1}^{foc=1, pitch=1} \end{array}$$

As for the focus projection, the + value of the FOC feature can be optionally passed up from the rightmost NP argument to the higher VP node marking the possible focus projection (FPP). This is not possible from the subject position (or from the not right-most argument), the focused NP in that position gets narrow focus intrepretation.



However, the picture is more complex, since by focusing we have to deal with (at least two) different issues: (i) the information structure of the sentence: which part of the content is the

Focus / Topic / Background as defined in formal pragmatic terms; (ii) the coherent discourse: what is "given / retrievable" and "non-given / noretrievable" information.

Towards an account of these issues, first of all, we take *focus* as a pragmatic notion, defined as the part of the answer that corresponds to the whpart of a question. Following this definition, the FOC feature is passed to the maximal projection of the noun phrase, marking the whole NP as the focus of the sentence. This raises the issue how we can deal with complex NPs like 'a green bicycle' where either the noun or the adjective gets the accent. In case the noun is accented 'a green  $BICYCLE_F$ ', it passes the focus marking to its maximal projection, and the whole noun phrase will be in focus. In case, that the adjective is accented 'a GREEN<sub>F</sub> bicycle' we can still mark the whole NP as focus, however we have to deal with the notion of giveness as well. Consider the following examples:

(5) **context1**: John attended a conference where the participants rented vehicles to move around.

What did John rent?

- a. John rented [a BICYCLE] $_F$ .
- b. John rented [a green BICYCLE] $_F$ .
- (6) **context2**: John attended a conference where all participants rented different kinds of bicycles to move around. *What did John rent?* 
  - a. #John rented [a BICYCLE] $_F$ .
  - b. #John rented [a green BICYCLE] $_F$ .
  - c. John rented [a GREEN bicycle] $_F$ .
  - d. John rented [a TANDEM] $_F$ .

In context 2, the wh-question is approporiate although the information that the rented vehicles are all bicycles is given. The answer in (6c) is felicitous with the focus marking of the whole NP as the corresponding to the wh-phrase of the question. Focus marking of the whole NP is thus possible even if some part of the NP is already given. Consequently, we need to distinguish the notions of *focus* and *given information*. This example supports the notion of *focus* in pragmatic terms and the proposed analysis of focus marking.

## 4.2 Focus and quantifier scope

The paper introduces the special scope window for focus and questions by the new features WH-MAX and WHMIN. These features follow the idea of MAXS and MINS from Kallmeyer & Romero (2008), however, in the previous examples we only discussed cases having only the focus window. In case we have both a quantificational NP and a focused constituent in the sentence, the distinction of the two scope windows get relevant and important. Consider, for example, the sentence  $SOMEONE_F$  walks. as an answer of the wh-question Who walks? In this example a quantificational NP is in focus, and its theme (issue) refers to the focus/question-window by the features WHMAX/WHIMN, while its rheme (content) makes use of the scope window by MAXS/MINS.

### Example 4.1



By substitution of the NP-tree of 'someone' (with focus) into the S-tree of 'walk', we derive the semantic representation of the sentence  $SOMEONE_F$  walks as the following.



That correctly derives – after plugging – the theme of the sentence as the wh-question *Who walks?* while the rheme as the proposition *someone walks*.



## Acknowledgments

I would like to thank the anonymous reviewers for their comments on the abstract of this paper and would like to express my special thanks to Laura Kallmeyer for all her comments and help, especially on the LTAG formalism.

## References

- Olga Babko-Malaya. 2004. Ltag semantics of focus. In Chung-hye Han and Anoop Sarkar, editors, *Proceedings of TAG+7: Seventh International Workshop on Tree Adjoining Grammar and Related Formalisms*, Vancouver, BC, Canada.
- Kata Balogh. 2009. *Theme with Variations. A Context-based Analysis of Focus.* Ph.D. thesis, ILLC, University of Amsterdam, Amsterdam.
- Jeroen Groenendijk. 2009. Inquisitive semantics: Two possibilities for disjunction. In Peter Bosch, David Gabelaia, and Jérôme Lang, editors, *Logic*, *Language and Computation. 7th International Tbilisi Symposium on Logic, Language and Computation. Revised Selected Papers.*, Berlin-Heidelberg. Springer-Verlag.
- C. L. Hamblin. 1973. Questions in montague english. *Foundations of Language*, 10.
- Aravind K. Joshi and Yves Schabes. 1997. Tree-Adjoning Grammars. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, pages 69–123. Springer, Berlin.
- Laura Kallmeyer and Aravind K. Joshi. 2003. Factoring predicate argument and scope semantics: Underspecified semantics with Itag. *Research on Language and Computation*, 1(1-2).
- Laura Kallmeyer and Maribel Romero. 2008. Scope and situation binding in LTAG using semantic unification. *Research on Language and Computation*, 6(1):3–52.
- Lauri Karttunen. 1977. Syntax and semantics of questions. *Linguistics and Philosophy*, 1.
- Craige Roberts. 1996. Information structure: Towards an integrated formal theory of pragmatics. In Jae-Hak Yoon and Andreas Kathol, editors, *OSUWPL: Papers in Semantics*, volume 49. Ohio State University, Columbus.

- Maribel Romero, Laura Kallmeyer, and Olga Babko-Malaya. 2004. LTAG Semantics for Questions. In Chung-hye Han and Anoop Sarkar, editors, *Proceedings of TAG+7: Seventh International Workshop on Tree Adjoining Grammar and Related Formalisms*, Vancouver, BC, Canada.
- Mats Rooth. 1992. A theory of focus interpretation. *Natural Language Semantics*, 1.
- Elisabeth Selkirk. 1996. Sentence prosody: Intonation, stress, and phrasing. In John A. Glodsmith, editor, *The Handbook of Phonological Theory*. Blackwell Publishing, Cambridge, MA / Oxford, UK.