Modeling Quantification with Polysemous Nouns

Laura Kallmeyer Heinrich-Heine-Universität Düsseldorf kallmeyer@phil.hhu.de Rainer Osswald Heinrich-Heine-Universität Düsseldorf osswald@phil.hhu.de

1 Introduction

Babonnaud et al. (2016) introduce a framework for modeling systematic polysemy that combines Lexicalized Tree Adjoining Grammar (LTAG) with frame semantics and Hybrid Logic (HL). The components of the syntax-semantics interface are elementary LTAG trees paired with frame descriptions that are expressed by possibly underspecified HL formulas. A brief review of this framework will be given in Section 2 below. In Babonnaud et al. (2016), an inherently polysemous noun such as 'book', which provides referential access to both a physical and an informational object, is assumed to refer to entities of type *phys(ical)-obj(ect)* which have an attribute CONTENT whose value is of type *information*. That is, the semantic type of 'book' is a subtype of the complex type *phys-obj* \land (CONTENT)*information* or *info(rmation)-carrier* for short (cf. Section 3 below). The physical aspect and the informational aspect are then addressed in different ways in contexts such as 'read the book', 'carry the book', 'master the book' or 'a heavy book on magic'. The verb 'read', for instance, is analyzed as denoting a complex event that consists of two components, the perception of the physical aspect and the comprehension of the informational aspect (following Pustejovsky 1998), and thus combines with both aspects of 'book'. By comparison, the verb 'master' calls for a grammatical object that denotes an entity of type *information* or *info-carrier* and semantically selects only the informational aspect in the latter case.

Treating books as physical information carriers in this way gives rise to the following "quantification puzzle" as noted in Asher and Pustejovsky (2005, 2006); see also Asher (2011). While (1a) poses no problem since the domain of quantification consists of physical entities, it not obvious how to cope with (1b), which is naturally interpreted as quantifying over all contents of the books in the library.

- (1) a. John carried off every book in the library.
 - b. John read every book in the library.

Since the library may own more than one copy of a book, there is not necessarily a one-to-one correspondence between the physical books in the library and the book contents. It is not even necessary that John used a copy from the library at all. The solution proposed in Section 4 retains the idea that a book basically denotes a physical object which has informational content associated to it. To this end, this basic representation will be embedded in an underspecified representation which allows the referential index of the NP to refer to the physical or to the informational component of the basic structure.

2 LTAG, Frames and Hybrid Logic

Babonnaud et al. (2016) follow Kallmeyer et al. (2015, 2016) in modeling the syntax-semantics interface by combining LTAG with frame semantics. More concretely, every elementary syntactic tree is paired with a frame description formulated in Hybrid Logic (Areces and ten Cate, 2007). Frames are semantic graphs with labeled nodes and edges, as in Fig. 1, where nodes correspond to entities (individuals, events, ...) and edges to (functional or non-functional) relations between these entities. In Fig. 1 all relations except *part-of* are functional. Frames can be formalized as extended typed feature structures



Figure 1: Frame for 'the man walked to the house' (adapted from Kallmeyer and Osswald (2013))

and specified as models of a suitable logical language (Kallmeyer and Osswald, 2013). In order to enable quantification over entities or events, Kallmeyer et al. (2016) propose to use Hybrid Logic (HL), an extension of modal logic. HL formulas have the following syntax: Let $Rel = Func \cup PropRel$ be a set of functional and non-functional *relation symbols*, Type a set of *type symbols*, Nom a set of *nominals* (node names), and Nvar a set of *node variables*, with Node = Nom \cup Nvar. Then *formulas* are defined as Forms ::= $\top \mid p \mid n \mid \neg \phi \mid \phi_1 \land \phi_2 \mid \langle R \rangle \phi \mid \exists \phi \mid @_n \phi \mid \downarrow x.\phi \mid \exists x.\phi, \text{ with } p \in \mathsf{Type},$ $n \in \mathsf{Node}, x \in \mathsf{Nvar}, R \in \mathsf{Rel} \text{ and } \phi, \phi_1, \phi_2 \in \mathsf{Forms.}$ Moreover, $\forall \phi \equiv \neg \exists (\neg \phi), \phi \rightarrow \psi \equiv \neg (\phi \land \neg \psi)$ and $\phi \lor \psi \equiv \neg(\neg \phi \land \neg \psi)$, as usual. A model is a triple $\langle M, (R^M)_{R \in \mathsf{Rel}}, V \rangle$ where M is a nonempty set (of "nodes"), each R^M is a binary relation on M, and V is a function (the *valuation*) from Type \cup Nom to $\wp(M)$ such that V(i) is a singleton for every nominal *i*. An assignment g is a function from Nvar to M; the assignment g_v^x differs from g at most on x, and $g_v^x(x) = v$. The following examples illustrate the notions of satisfaction and truth with reference to the frame structure in Fig. 1; see Kallmeyer et al. (2015, 2016) for more details and a formal definition. The formula motion, a type symbol, is true at the node denoted by the nominal n_0 in the structure of Fig. 1. Formula $\langle R \rangle \phi$ is true at a node if the relation R^M holds between that node and a node at which ϕ is true. For example, $(AGENT)man \land (MANNER)walking \land (PATH) \land (ENDP) \top$ is true at n_0 . (Notice that since HL does not distinguish between functional and non-functional edge labels, functionality has to be enforced by additional constraints.) There are different ways to express existential quantification in HL, namely $\exists \phi$ and $\exists x.\phi$. Formula $\exists \phi$ is true at node v if there exists a node v' at which ϕ holds. For instance, $\exists house$ is true at all nodes of the example structure. Formula $\exists x.\phi$ is true at v if there is a v' such that ϕ is true at v under an assignment of x to v'. E.g., $\exists x. \langle PATH \rangle \langle ENDP \rangle \langle part-of \rangle (x \land region) \land \exists (house \land \langle AT-REGION \rangle x)$ is true at n_0 . Besides quantification, HL also allows us to use nominals or variables to refer to nodes via the @ operator: $@_n \phi$ is true at a node if ϕ is true at n. The \downarrow operator allows us to assign the current node to a variable: $\downarrow x.\phi$ is true at v if ϕ is true at v under the assignment g_v^x . E.g., $\langle PATH \rangle \langle ENDP \rangle \langle part-of \rangle (\downarrow x.region \land \exists (house \land \langle AT-REGION \rangle x))$ is true at n_0 .

A Lexicalized Tree Adjoining Grammar (LTAG; Joshi and Schabes 1997; Abeillé and Rambow 2000) consists of a finite set of *elementary trees*. Larger trees are derived via *substitution* (replacing a leaf with a tree) and *adjunction* (replacing an internal node with a tree). An adjoining tree has a unique *foot node* (marked with an asterisk), which is a non-terminal leaf labeled with the same category as the root of the tree. When adjoining such a tree to some node n of another tree, in the resulting tree, the subtree with root n from the original tree is attached at the foot node of the adjoining tree. Non-terminal nodes in LTAG are usually enriched with feature structures (Vijay-Shanker and Joshi, 1988): Each node has a top and a bottom feature structure (except substitution nodes, which have only a top). Features can be shared within elementary trees. In a substitution step, the top of the root of the new tree unifies with the top of the adjouncion site and the bottom of the foot of the adjoining tree unifies with the top of the adjunction site. Furthermore, in the final derived tree, top and bottom unify in all nodes.

In line with previous proposals of how to add semantics to LTAG, we pair each elementary tree with a semantic representation that consists of a set of HL formulas which can contain holes and which can be labeled. In other words, *hole semantics* (Bos, 1995) is applied to HL and these underspecified formulas are linked to the elementary trees. Composition is triggered by the syntactic unifications using interface features on the syntactic trees (cf. Gardent and Kallmeyer 2003; Kallmeyer and Romero 2008). As an example consider the derivation in Fig. 2 where the two NPs are substituted into the two argument slots



Figure 2: Derivation of 'John ate pizza'

in the 'ate' tree. The interface features I on the NP nodes make sure that the contributions of the two arguments feed into the AGENT and THEME nodes of the frame structure. Furthermore, an interface feature MINS is used for providing the label of the \exists (*eating*...) formula as minimal scope to a possible quantifier. The feature unifications on the syntactic tree lead to 1 = i, 2 = x and $4 = l_1$. As a result, when collecting all formulas, we obtain the underspecified representation (2a).

(2) a. $@_i(person \land \langle NAME \rangle John), \ l_1: \exists (eating \land \langle AGENT \rangle i \land \langle THEME \rangle x), \ \exists (\downarrow x.pizza \land \exists), \ \exists \lhd^* l_1$ b. $@_i(person \land \langle NAME \rangle John), \ \exists (\downarrow x.pizza \land \exists (eating \land \langle AGENT \rangle i \land \langle THEME \rangle x))$

The relation \triangleleft^* links holes to labels: $h \triangleleft^* l$ signifies that the formula labeled l is a subformula of h. In (2a), the \exists (*eating*...) formula, labeled l_1 , has to be part of the nuclear scope of the quantifier (hole \exists). Disambiguating such underspecified representations consists of "plugging" the labeled formulas into the holes while respecting the given constraints. (2a) has a unique disambiguation, namely $\exists \rightarrow l_1$. This leads to (2b), which is then interpreted conjunctively.

3 Basic analysis of polysemous nouns

We will now sketch the analysis of Babonnaud et al. (2016) of inherently polysemous nouns such as 'book' and their composition with verbs such as *read*. The noun 'book' carries two meaning aspects in that it can denote a physical object ('the book is heavy') or an informational content ('the book is interesting'). Babonnaud et al. (2016) assume that the semantic frame associated with 'book' contains two nodes of type *information* and *phys-obj*, respectively, with an explicit relation between them: The physical aspect of the frame is taken to be its referential node, which is linked via a CONTENT attribute to the information it carries. The type *book* is a subtype of *info-carrier*, which in turn is a subtype of *phys-obj* and which has a $\langle CONTENT \rangle$ attribute whose value is of type *information*.

The verb 'read' allows for the direct selection of the dot object *book* as complement (cf. Pustejovsky, 1998) but also enables coercion of its complement from type *information* ('read the story') as well as from type *phys-obj* ('read the blackboard'). Babonnaud et al. (2016) assume an event node of type *reading* with attributes PERC(EPTUAL)-COMP(ONENT) and MENT(AL)-COMP(ONENT), whose values are respectively of type *perception* and *comprehension*. These nodes represent the decomposition of the activity of reading into two subevents, the action of looking at a physical object (the perception) and the action of processing the provided information (the comprehension). Moreover, the *perception* node has an attribute STIMULUS of type *info-carrier*, and the *comprehension* node has an attribute CONTENT whose value is the information being read, which coincides with the value of CONTENT attribute of the *info-carrier* node. The fact that the argument contributed by the object can be either the stimulus of the perception (*phys-obj*) or its content is captured by a disjunction.

A sample analysis is given in Fig. 3. The label of the HL formula coming with 'read' is provided as potential minimal scope for quantifiers at the NP slots. The determiner 'the' is treated as an existential quantifier, disregarding the presuppositions it carries. The label of the formula associated with 'book' is made available via an interface feature P (for "proposition"). Due to the two scope constraints, this proposition will be part of the restriction of the quantifier (i.e., part of the subformula at $\boxed{4}$) while the 'read' formula will be part of the nuclear scope, i.e., part of the subformula at $\boxed{5}$. Substitutions and



Figure 3: Derivation for 'John read the book'

adjunctions lead to the unifications $\overline{0} = i$, $\overline{1} = z$, $\overline{2} = l_0$ and $\overline{3} = l_1$ on the interface features. The result is the underspecified representation in (3), which has to be disambiguated by the mapping $\overline{4} \mapsto l_1$, $\overline{5} \mapsto l_0$. Since *information* and *phys-obj* are incompatible, it follows that $z \leftrightarrow x$ and $\neg(z \leftrightarrow y)$.

(3) $\begin{array}{l} @_{i}(person \land \langle \mathsf{NAME} \rangle John), \ \exists (\downarrow z. \underline{4} \land \underline{5}), \\ l_{0}: \exists x. \exists y. \exists (reading \land \langle \mathsf{AGENT} \rangle i \\ \land \langle \mathsf{PERC}\text{-}\mathsf{COMP} \rangle \langle \mathsf{STIMULUS} \rangle x \land \langle \mathsf{MENT}\text{-}\mathsf{COMP} \rangle \langle \mathsf{CONTENT} \rangle y \\ \land @_{x}(phys \text{-}obj \land \langle \mathsf{CONTENT} \rangle (information \land y)) \land (z \leftrightarrow x \lor z \leftrightarrow y)), \\ l_{1}: book \land \langle \mathsf{CONTENT} \rangle information, \\ \underline{4} \lhd^{*} l_{1}, \ \underline{5} \lhd^{*} l_{0} \end{array}$

4 Polysemous nouns and quantification

A problem of the approach described above concerns the decision to provide the *phys-obj* node of the book frame as an argument in a predicate argument structure. As explained in the introduction, this is problematic in examples like (1b) ('John read every book in the library') where the noun interacts with a quantifier in such a way that the quantifier's restriction concerns specific pairs of physical objects and associated information content but the predicate the quantifier scopes over applies only to the latter. The natural reading of (1b) is that for every book in the library, John read a (possibly different) copy of that book, i.e., John read some book with the content of that book. Simply changing the nodes provided for predicate-argument composition to the information content does not work either, as shown in (1a) ('John carried off every book in the library') where 'carried off' applies to the physical objects.

As a solution, we revise the entry of 'book' such that it explicitly provides an underspecified I feature at the syntax-semantics interface (I = \mathbb{S} in Fig. 4a)) and the value of this feature can either be a variable referring to the *phys-obj* node or a variable referring to the *information* node (expressed by the disjunction $\mathbb{S} \leftrightarrow u \vee \mathbb{S} \leftrightarrow v$). The restriction of the adjoining quantifier (variable \mathbb{A}) has to embed the entire formula contributed by the 'book' tree (i.e., the formula labeled l_1), while the formula that characterizes the *book* frame (label l_4) has to be provided as well for further modification, for instance by 'in the library' (interface feature P). This latter formula is embedded under the formula labeled l_1 (constraint $\mathbb{II} \triangleleft^* l_4$). The derivation of (1b) is given in Fig. 4a). The result of combining the trees for the complex NP 'every book in the library' is shown in Fig. 4b). It is underspecified with respect to whether the variable z introduced by the quantifier refers to the *book* node or to the *information* node. The frame structure contributed by the NP is, however, still the *phys-obj* node with a CONTENT attribute of type *information*, as in the basic analysis presented in Section 3. As a result of the entire derivation, we obtain the frame description in (4), in particular $\mathbb{I} = z$. Disambiguating the scope constraints leads to the plugging $\mathbb{A} \to l_1$, $\mathbb{B} \to l_0$. But the disjunction $z \leftrightarrow x \lor z \leftrightarrow y$ cannot be resolved, i.e., we are not



Figure 4: (a) Derivation for (1b); (b) result of deriving 'every book in the library'

able to disambiguate between quantification over the physical objects (option $z \leftrightarrow x$ which would imply $z \leftrightarrow u$) on the one hand and information content ($z \leftrightarrow y$ and $z \leftrightarrow v$) on the other hand, since 'read' allows for both argument types. In a case like (1a), a disambiguation towards the first option takes place.

$$\begin{aligned} & (4) \quad \forall (\downarrow z. \underline{4} \to \underline{5}), \\ & l_1: \exists u. \exists v. @_u(book \land \langle \text{CONTENT} \rangle (information \land v) \land \langle \text{LOCATION} \rangle library) \\ & \land (z \leftrightarrow u \lor z \leftrightarrow v), \\ & l_0: \exists x. \exists y. \underline{3} (reading \land \langle \text{AGENT} \rangle i \land \langle \text{PERC-COMP} \rangle \langle \text{STIMULUS} \rangle x \land \langle \text{MENT-COMP} \rangle \langle \text{CONTENT} \rangle y \\ & \land @_x(phys-obj \land \langle \text{CONTENT} \rangle (information \land y)) \\ & \land (z \leftrightarrow x \lor z \leftrightarrow y)), \\ & @_i(person \land \langle \text{NAME} \rangle John), \\ & \underline{4} \lhd^* l_1, \ \underline{5} \lhd^* l_0 \end{aligned}$$

5 Copredication and quantification

A crucial property of the proposed analysis is that we do not assume separate nodes for dot objects. In a quantificational NP involving *book*, we therefore have to quantify either over the physical objects or over the information contents. This might pose problems for cases of copredication as in (5).

(5) a. John destroyed every book in the library that Mary had mastered



Figure 5: Derivation for every book in the library that _ had mastered

- b. John mastered and then destroyed every book in the library.
- c. Mary mastered every book $_i$ in the library before John destroyed it $_i$.

Note that the analyses of coordination as in (5b) and also of anaphoric reference as in (5c) are still open questions in LTAG. Concerning (5a), the predicate *destroy* clearly is a predication over the physical objects while the predicate *master* takes the informational content as argument. In the following, we will detail the analysis of (5a) in order to show that disambiguating towards one of the two options in the context of one predication does not necessarily exclude predicating in a different place over the other component.

Fig. 5 gives the derivation of *every book in the library that _ had mastered*. As in the examples in the preceding sections, we have features P and TOP on the NP nodes. The value of P is the label of the frame description contributed by the (complex) noun without considering the case of quantification. The value of TOP gives the label of the formula that characterizes the possible "entry points" into this frame (in our case the physical object node and the information node). Once these formulas are introduced via the noun, any further predication can access the frame itself via the P feature and extend its description. This is the case for *in the library* (as already in Fig. 4), and the predication *that Mary had mastered* works similarly (see the tree-frame pair on the right of Fig. 5). In other words, the relative clause headed by *mastered* adjoins to the NP above the *in the library* PP and simply adds another conjunct to the frame description that holds at the physical object node of the *book* frame.

A crucial feature of the frame description associated with *master* is that its THEME argument has to be of type *knowledge* (which is a subtype of *information*) and it is either the frame modified by the relative clause or the value of the CONTENT attribute of this frame (see also Babonnaud et al., 2016, for similar examples of copredication). This disjunction accounts for the coercion effect in (5a) since it allows *master* to predicate over the value of the CONTENT attribute of the modified *book* frame. Because of this, (5c) and probably also (5b) are not problematic. Even though the predicate *destroy* requires that the variable of the quantification is the physical object, the predicate *master* still has access to the informational component via the CONTENT attribute.

6 Futher issues

As Asher (2011) points out, the relation from books as physical objects to books as informational units is not necessarily functional because of omnibus editions. Likewise, the content of a single novel can be spread over several physical volumes. For this reason, Asher posits dot objects as entities of their own right which are related to their physical and informational aspects by appropriate elaboration relations. However, the problem just mentioned is more complicated since it can happen with multi-volume editions of collected works that one novel is distributed over two volumes and that the second volume contains another novel in addition to the final part of the first novel. This means that if we want to quantify over books understood as novels, which may have once been published as single volumes, the only way to cope with situations like the one just described is to quantify over the elements of an appropriate segmentation of a (mereological) sum of the CONTENT values of the (physical) books in the library.

A related topic concerns the interaction of copredication and counting as illustrated in the examples in (6), which are taken from Gotham (2017).

(6) Three informative books are heavy.

In order to be true, sentence (6) seems to require three different copies of books with different informational contents. That is, an appropriate representation of (6) must be able to encode the constraints that the three books have different informational content and that they are different physical objects. We leave the analysis of such examples within our framework for future work.

Another interesting issue arises in examples like (7), which make explicit that reading a book does not necessarily require to stick to a single copy. In general, it is possible to switch between different physical copies including instantiations on electronic devices such as ebook readers and the like.

(7) Mary read the heavy book on magic. She read part of it on her ebook reader for convenience.

If we want to build the variability of the physical carrier into our approach while preserving the idea that reading consists of a perception and a comprehension component, we need to zoom into the course of the reading event and describe it as consisting of different subevents, each of which is bound to a certain physical information carrier. To this end, we need to extend our model by some version of the homomorphic event-to-object relation that comes with incremental themes.

Our analysis of copredication described in Section 5 assumes that copredication is flexible in the sense that picking an attribute as argument (for instance the CONTENT) does not mean that the original root node is no longer accessible. In many cases, this makes the right predictions. However, there seem to be cases where a predication over one aspect blocks the other meaning aspects for further access; compare, for instance, the examples in (8) taken from Retoré (2014):

- (8) a. Liverpool is spread out and voted (last Sunday).
 - b. # Liverpool voted and won (last Sunday).

Retoré (2014) models coercion within the framework of the Montagovian Generative Lexicon (Mery et al., 2007), which is a many-sorted higher order predicate calculus where the lexical entries associated with words are finite sets of λ -terms, one of them being the principal λ -term. Coercion is triggered by type mismatch, and one of the optional λ -terms is used in this case, which changes the type of either the predicate or the argument. Some of these type shifts are specified as rigid in the lexicon, which means that they block coercion to other types. Our system shares with the Montagovian Generative Lexicon the assumption of a rich type system and of functional relations between the different instances of these types present in the meaning of a word. We will look more closely into examples of copredication and its interaction with quantification in future work, in particular into the question of flexibility/rigidity of coercive shifts.

References

- Abeillé, A. and O. Rambow (2000). Tree Adjoining Grammar: An Overview. In A. Abeillé and O. Rambow (Eds.), *Tree Adjoining Grammars: Formalisms, Linguistic Analysis and Processing*, pp. 1–68. CSLI.
- Areces, C. and B. ten Cate (2007). Hybrid logics. In P. Blackburn, J. V. Benthem, and F. Wolter (Eds.), *Handbook of Modal Logic*, Volume 3 of *Studies in Logic and Practical Reasoning*, Chapter 14, pp. 821–868. Elsevier.
- Asher, N. (2011). Lexical Meaning in Context. A Web of Words. Cambridge: Cambridge University Press.
- Asher, N. and J. Pustejovsky (2005). Word meaning and commonsense metaphysics. Manuscript, available at semanticsarchive.net/Archive/TgxMDNkM/.
- Asher, N. and J. Pustejovsky (2006). A type composition logic for generative lexicon. *Journal of Cognitive Science* 6, 1–38.
- Babonnaud, W., L. Kallmeyer, and R. Osswald (2016). Polysemy and coercion a frame-based approach using LTAG and Hybrid Logic. In M. Amblard, P. de Groote, S. Pogodalla, and C. Retoré (Eds.), *Logical Aspects of Computational Linguistics, 9th International Conference*, Lecture Notes in Artificial Intelligence 10054, pp. 18–33. Berlin: Springer.
- Bos, J. (1995). Predicate logic unplugged. In P. Dekker and M. Stokhof (Eds.), *Proceedings of the 10th Amsterdam Colloquium*, pp. 133–142.
- Gardent, C. and L. Kallmeyer (2003). Semantic Construction in FTAG. In *Proceedings of EACL 2003*, Budapest, pp. 123–130.
- Gotham, M. (2017). Composing criteria of individuation in copredication. *Journal of Semantics 34*, 333–371.
- Joshi, A. K. and Y. Schabes (1997). Tree-Adjoning Grammars. In G. Rozenberg and A. Salomaa (Eds.), *Handbook of Formal Languages*, pp. 69–123. Berlin: Springer.
- Kallmeyer, L., T. Lichte, R. Osswald, S. Pogodalla, and C. Wurm (2015, August). Quantification in Frame Semantics with Hybrid Logic. In R. Cooper and C. Retoré (Eds.), *Type Theory and Lexical Semantics*, ESSLLI 2015, Barcelona, Spain.
- Kallmeyer, L. and R. Osswald (2013). Syntax-driven semantic frame composition in Lexicalized Tree Adjoining Grammars. *Journal of Language Modelling 1*(2), 267–330.
- Kallmeyer, L., R. Osswald, and S. Pogodalla (2016). *For*-adverbials and aspectual interpretation: An LTAG analysis using hybrid logic and frame semantics. In C. Piñón (Ed.), *Empirical Issues in Syntax and Semantics EISS*, Volume 11.
- Kallmeyer, L. and M. Romero (2008). Scope and situation binding in LTAG using semantic unification. *Research on Language and Computation* 6(1), 3–52.
- Mery, B., C. Bassac, and C. Retoré (2007). A montagovian generative lexicon. In L. Kallmeyer, P. Monachesi, G. Penn, and G. Satta (Eds.), 12th conference on Formal Grammar (FG 2007), Dublin, Ireland. CSLI Publications.
- Pustejovsky, J. (1998). The semantics of lexical underspecification. Folia Linguistica 32(3-4), 323-348.

- Retoré, C. (2014). The Montagovian Generative Lexicon $\Lambda T y_n$: a type theoretical framework for natural language semantics. In *TYPES: International Workshop on Types and Proofs for Programs, April 2013, Toulouse, France*, pp. 202–229.
- Vijay-Shanker, K. and A. K. Joshi (1988). Feature structures based tree adjoining grammar. In *Proceedings of COLING*, Budapest, pp. 714–719.