A Frame-Based Semantics of Locative Alternation in LTAG

Yulia Zinova

Heinrich Heine University Düsseldorf
zinova@phil.uni-duesseldorf.de

Abstract. In this paper I present the analysis of locative alternation phenomena in Russian and English. This analysis follows the approach proposed in [13] and uses LTAG and Frame Semantics. The combination of a syntactic theory with an extended domain of locality and frames provides a powerful mechanism for argument linking. Metagrammar factorization allows to determine not only lexical, but also constructional meaning that is essential for locative alternation analysis.

1 Introduction

There is a number of formalisms that capture the idea that the meaning of a verb-based construction depends both on the lexical meaning of the verb and on the construction in which the verb is used ([8], [19]). The question is how exactly the components of the meaning are distributed and how they combine.
In [13] a combination of Lexicalized Tree Adjoining Grammars ([9]) and Frame Semantics is introduced. It is shown that the resulting framework is very flexible with respect to the factorisation and combination of lexical and constructional units on the syntax and semantics level and is also suitable for computational processing.
Though there already exist a number of different approaches to semantic construction using LTAG ([10], [5], [12]) and approaches that combine other syntactic formalisms with Frame Semantics ([3], [4]), the novel combination of an LTAG and Frame Semantics benefits from both extended domain of locality and underspecification allowed by frames.
In this paper I want to present the analysis of locative alternation that benefits of flexibility offered by the novel framework.

2 Tree Adjoining Grammar

Tree Adjoining Grammar (TAG, [9]) is a tree-rewriting grammar formalism. It consists of a finite set of *elementary trees* with labelled nodes with two operations on them: *substitution* and *adjunction*. All elementary trees are either *auxiliary trees* or *initial trees*. An *auxiliary tree* is a tree which has exactly one *foot node* - a leaf that is marked with an asterisk. Leaf nodes can be labelled with terminals and other nodes are labelled only with non-terminals. The derivation process starts from an initial tree and in the final *derived tree* all the leaves must be labelled by terminals.
Substitution allows to replace a non-terminal leaf with a new tree and *adjunction* is used for replacing an internal node with an auxiliary tree. Adjunction to the node

R.K. Rendsvig and S. Katrenko (Eds.): *ESSLLI 2012 Student Session Proceedings*, CEUR Workshop Proceedings vol.: see http://ceur-ws.org/, 2012, pp. 202–210.

labelled X is allowed if the root and foot nodes of the adjoining auxiliary tree have the same label X. It is also possible to indicate nodes where adjunction is obligatory or not allowed with *OA* and *NA* subscripts respectively.
Figure 1 shows an example of a derivation: the initial tree for *Mary* substitutes into the subject slot of the elementary tree for *laughs*, and the *sometimes* auxiliary tree for the VP modifier adjoins to the VP node. *Feature-structure based TAG*, or FTAG, is a

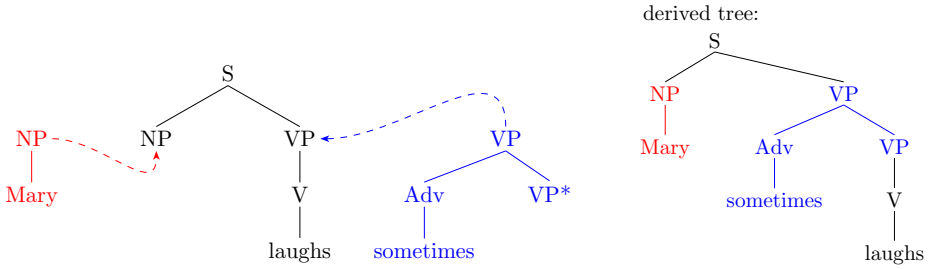

Fig. 1. Example of a TAG derivation

variant of TAG in which elementary trees are enriched with feature structures ([20]). In an FTAG each node has a top feature structure and all the nodes except substitution nodes have a bottom feature structure. Feature unification happens during the derivation process when adjunction and substitution take place. Due to the extended domain of locality, nodes within one elementary tree can share features, allowing to express constraints among dependent nodes easily.
For natural languages a specific version of TAG called *lexicalized TAG,* or LTAG is used. In an LTAG, each elementary tree must have at least one non-empty lexical item, called *lexical anchor*. The second important principle for a natural language TAG is that every elementary tree where the lexical anchor is a predicate must contain slots (leaves with non-terminal labels) for all arguments of this predicate, including the subject, and for nothing else (*theta-criterion for TAG*, [6]).
The facts that LTAGs have extended domains of locality and that elementary trees are lexicalized and contain slots for all the predicate's arguments, make them good candidates for combination with frame-based compositional semantics ([13]). In the approach proposed in [13], a single semantic representation (a semantic frame in this case) is linked to the entire elementary tree. When coupling an elementary tree with a semantic frame, syntactic arguments can be directly linked to their counterpart in the semantics. Described approach is similar to ones in [7] and [14], but uses different kind of semantic representation. Semantic composition is then modeled by unification which is a result of performing adjunction and substitution. Figure 2 provides a simple illustration of syntactic and semantic composition. In this example, substitutions trigger unifications between 1 and 3 and 2 and 4 which leads to correct insertion of argument frames into the frame of *loves*.
Linguistic generalizations in TAGs are captured by a metagrammar. There are two steps of factorization, which are important for this paper:

- *unanchored elementary trees* are specified separately from lexical anchors;

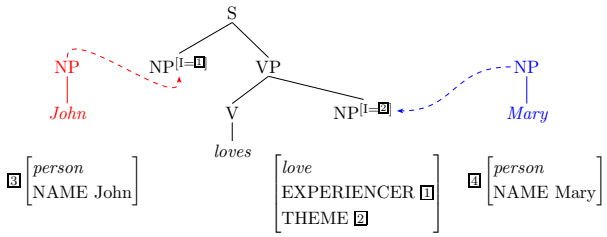

Fig. 2. Syntactic and semantic composition for *John loves Mary*

- trees are organized into *tree families* which represent different realizations of one subcategorization frame.

This allows to define a meaning for sets of unanchored elementary trees, i.e., a meaning of constructions.

3 The Data

3.1 Previous approaches

(1) and (2) show basic examples of locative alternation in English and Russian. Despite the fact that in English both constructions have a PP and it can be omitted without losing the specific construction meaning, let us call the first variant ((1a), (2a)) *prepositional phrase construction,* or *PPC* and the second variant ((1b), (2b)) - *instrumental case construction*, or *ICC* for convenience of referring to them.

(1) a. John$_{\boxed{1}}$ loaded the hay$_{\boxed{2}}$ into the truck$_{\boxed{3}}$. (PPC)
b. John$_{\boxed{1}}$ loaded the truck$_{\boxed{3}}$ with hay$_{\boxed{2}}$. (ICC)

(2) a. Ivan$_{\boxed{1}}$ zagruzil seno$_{\boxed{2}}$ v vagon$_{\boxed{3}}$.
Ivan loaded hay$_{acc,def}$ in wagon$_{gen,indef/def}$.
Ivan loaded the hay into a/the wagon.
b. Ivan$_{\boxed{1}}$ zagruzil vagon$_{\boxed{3}}$ senom$_{\boxed{2}}$.
Ivan loaded wagon$_{acc,def}$ hay$_{instr,indef}$.
Ivan loaded the wagon with hay.

PPCs are traditionally analyzed as having a change of location meaning and ICCs - as having a change of state meaning ([11], [16], [8]). An analysis for (1) following [11] is provided in (3). It demonstrates that there is a difference between the two constructions, but only the difference in the perspective is shown.

(3) a. X CAUSE [BECOME [hay BE ON truck]]
b. X CAUSE [BECOME [truck$_{\boxed{z}}$ BE [WITH [hay BE ON z]]]]

The analysis proposed in [16], which can be found under (4), provides more detailed information about the difference between PPCs and ICCs. (4a) tells us that the hay changes its location as a result of the loading event, while (4b) describes that the result is a change in the state of the wagon. One can notice that in (3) there is no explicit

reference to the verb itself and the only component that is taken from the verb meaning is that the result of the loading is that the THEME is on the LOCATION in the end.

(4) a. [[x ACT] CAUSE [y BECOME P_{loc} z] $[LOAD]_{MANNER}$]
b. [[x ACT] CAUSE [z BECOME $[]_{STATE}$ WITH-RESPECT-TO y] $[LOAD]_{MANNER}$]

Using frame semantics, one can assign two frames in 3 to the two different constructions. For the PPC, one has to remove the concrete verb *load* and replace it with *change_of_location* effect. So the first frame tells us that the activity of the Actor (X) causes the Theme (Y) to change its location to the Goal (Z). For the ICC's frame in order to introduce the Manner one can simply embed the caused change of location frame under the MANNER atribute. The second frame in this case would mean that the activity of the Actor (X) causes the Theme (Z) to change its state by means of changing the location of the third argument (Y) to Z.

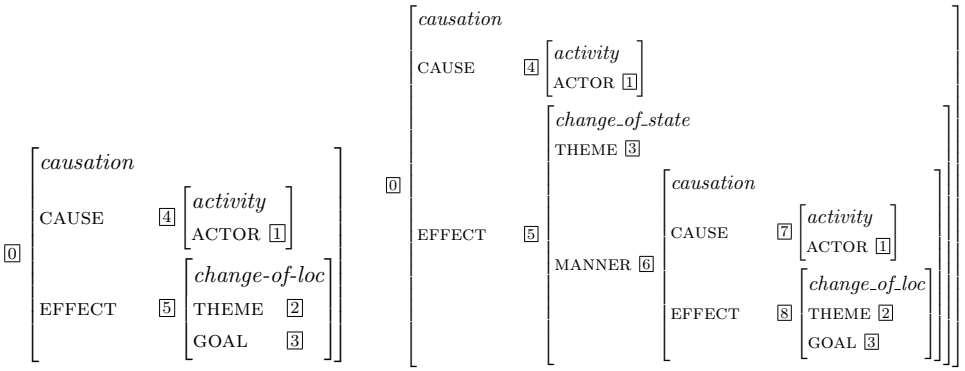

Fig. 3. Frame semantics for PPC and ICC

3.2 Detailed Russian and English Data

The question that arises if one looks carefully at what the sentences in (1) and (2) mean is whether it is really the case that there is no change of state in PPC examples? In fact, any loading activity leads to both a change of location of the content and some change of state of the container (if it is specified), and the difference between two constructions is that

- different components of the effect become more salient;
- in the case of ICC initial and result states of the container are specified.

In order to understand how the meaning of verbs and constructions should be represented let us look at the whole range of the verbs allowing locative alternation that one can find in English and Russian. [18] provides the following classification for English:

Content-oriented classes:

(a) simultaneous forceful contact and motion of a mass against a surface (brush, drape, spread, etc.);
(b) vertical arrangement on a horizontal surface (heap, pile, stack);
(c) force is imparted to a mass, causing ballistic motion in a specified spatial distribution along a trajectory (inject, splash, spray, etc.);
(d) mass is caused to move in a widespread or non directed distribution (scatter, seed, sow, etc.).

Container-oriented classes:

(e) a mass is forced into a container against the limit of its capacity (crowd, jam, stuff, etc.);
(f) a mass of size, shape, or type defined by the intended use of a contained is put into the container, enabling it to accomplish its function (load, pack, stock).

From the description of verb classes that allow locative alternation in English one can see that the result state of the container in case of ICC is such that the action cannot be performed any longer. There is no result state common for all the cases, so it depends on the verb, i.e. on how the change of location happens. The easiest way to solve this would be to assume different construction meanings for different verb classes (e.g. one with the Effect of the Theme being full and the other one with the Effect of the Theme being covered), but let us first look at some Russian data.
In Russian a lot of verbs allow only one of the constructions, i.e. a change of construction requires a change of verb prefix (a list can be found in [1]). However, some of the verbs from the list remain the same in both prepositional and instrumental constructions. Such verbs can be organised in three groups: the first one is similar to the (f) group in English (see example (2)), the second one is similar to group (a) in English, like in (5), and the third class is like a combination of the first and the second: a mass is put into a container, enabling it to accomplish its function, or on a container, covering its surface (6).
With the vebs from the third group an interesting effect can be observed: while in the case of PPC example (6a) there is a preposition which tells us that the content goes *in* the container, in the case of ICC example (6b) two different readings are possible: the content can be put in the container or the content can cover the container. In both cases there is a clear result state: either the container is full or the container's surface is fully covered with content. This means that the verb *zasypat'* ('to fill/to cover') does not provide information about how the THEME is positioned at the GOAL. In case of PPC this information comes from the preposition used (both *v* ('in') and *na* ('on') are possible) and in ICC the ambiguity can be resolved only using world knowledge. So (6) demonstrates conclusively that there should be one construction accounting for different result states of the theme and allowing to get different interpretations of one verb due to underspecification of how the change of location process goes.

(5) a. On namazal maslo na hleb.
He distributed butter$_{acc,def}$ on bread$_{acc,indef}$
He distributed butter over a piece of bread.
b. On namazal hleb maslom.
He covered bread$_{acc,def}$ butter$_{instr,indef}$
He covered a piece of bread with butter.

(6) a. On zasypal sahar v banku.
He put suggar$_{acc,def}$ in can$_{acc,indef/def}$

He put sugar in a/the tin.

b. On zasypal banku saharom.
He covered/filled $\text{tin}_{acc,def}$ $\text{sugar}_{instr,indef}$
He covered/filled the tin with sugar.

4 Locative Alternation: The Analysis

4.1 Syntactic representation

In the previous section we were looking only at "full" examples, where both container and content are present. However, the constructions that are being discussed can be used when only the direct object of the verb is present; in this case, they will have the same difference in semantics. Using LTAG and metagrammar decomposition one can obtain the tree family in 4 for the PPC and tree family in 5 for the ICC (the second NP_{INSTR} stands for both NP in instrumental case in Russian and PP with preposition "with" in English).

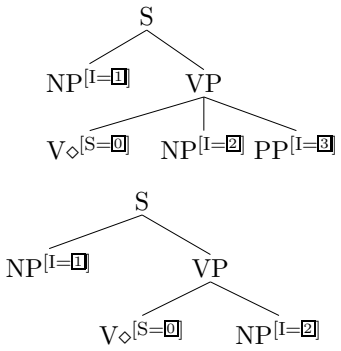

Fig. 4. Unanchored trees for the PPC

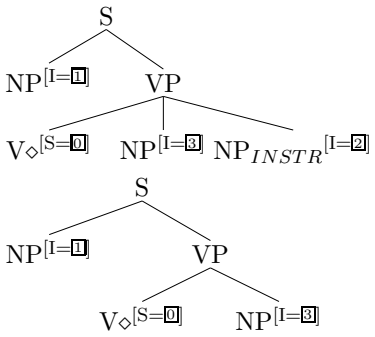

Fig. 5. Unanchored trees for the ICC

4.2 Scales and Proposed Frame Semantics

In the case of the PPC, the semantics of the whole phrase can be compositionaly derived from the semantics of the verb and its arguments, while in the case of the ICC there is a part of the meaning, that comes from the construction itself. The goal now is to provide such semantics for the ICC and verbs allowing locative alternation such that in combination they form the desired frame representation of the semantics of a sentence.

Following ideas in [17] where one can find a discussion of the representation of attributes, events, and results while implementing Fillmore's Frame Semantics ([2]) I introduce attributes of initial and result state and a scale which is determined by its type, start and end points. The change of state is either a decrease or an increase of the value on an ordered scale (discussion of analysis of scalar change can be found in [15]). The direction is given by the values of atributes ENDP and STARTP (end and start

points), which replaces the LESSER atribute of ordering proposed in [17]. Some of the verbs specify a concrete initial or result state (INIT and RESULT respectively), but *load* does not have any initial or result state specified within its semantics, so it just determines the scale and two values on it. Summarizing the ideas, one obtains the following scheme:

- in the verb both change of location and change of state effects are specified;
- MANNER attribute is not needed because it is already described in the change of location subframe;
- change of state is described by scale, initial state, and result state;
- SCALE atribute can have a type such as "degree of fullness" that is a subtype of the type "scale" and thus replaces it during the unification;
- initial and result states are values on the scale;
- the ICC specifies that initial and result states are equal to the start and end points of the scale, respectively.

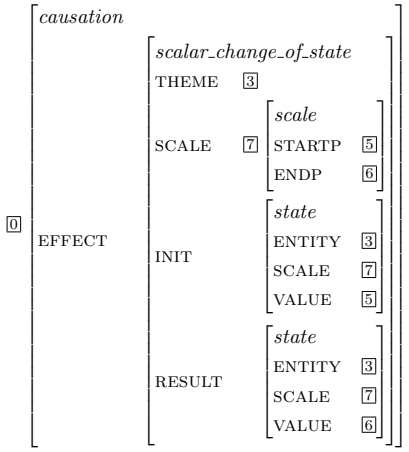

Fig. 6. Frame for the ICC

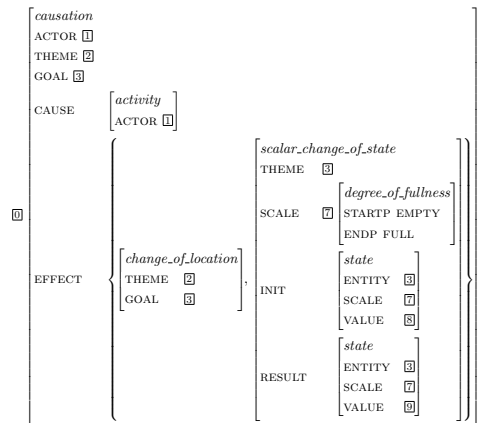

Fig. 7. Frame representation of the verb *load*

Figure 7 shows a lexical frame for the verb *load*. As one can see, when all the arguments are filled, the right meaning for the whole PPC follows automatically. The semantics of the ICC is a caused change of state meaning that gets further constrained when a specific lexical anchor is inserted. Figure 6 shows how the unanchored tree for the ICC is linked to its semantic frame. The correct argument linking happens because I features in the syntactic tree and the thematic roles in the semantic frame are identical. This is done in a local way (within the domain of an elementary tree) because of LTAG's extanded domain of locality. The S feature of the V node serves for unification of the lexical frame for the verb and the constructional frame. When a lexical anchor is inserted, this feature unifies with the S feature of the lexical item. The result of this unification for the ICC with a lexical anchor *load* is shown in Fig. 8.

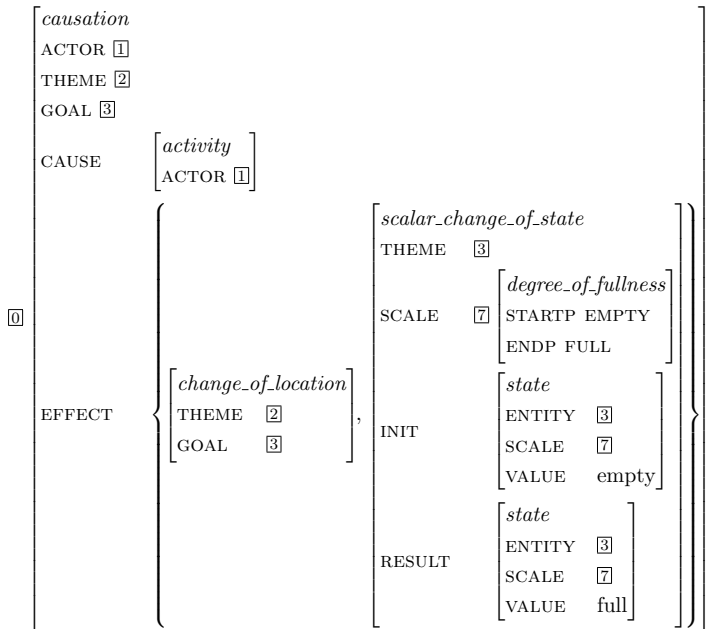

Fig. 8. Resulting frame for the ICC and the verb *load*

5 Conclusion

This is a case study for combining an LTAG with Frame Semantics, in which I have described a model for locative alternation in English and Russian. This analysis uses LTAG's mechanism of separation between unanchored elementary trees and lexical anchors to separate the contribution of the lexical meaning from the contribution of construction and follows the ideas expressed in [13]. Another advantage of combination of an LTAG with Frame Semantics is that LTAG's extended domain of locality allows direct linking of thematic roles of the arguments with corresponding syntatctic slots. As this framework is a new one, there are a lot of open questions and a wider range of semantic phenomena should be examined.

References

1. Dudchuk, P.: Locativnaja alternacija i struktura glagol'noj gruppy [Locative Alternation and the structure of VP]. diploma thesis, Moscow State University (2006)
2. Fillmore, C.J.: Frame semantics. In: of Korea, T.L.S. (ed.) Linguistics in the Morning Calm, pp. 111–137. Hanshin Publishing Co., Seoul (1982)
3. Frank, A.: Generalizations over corpus-induced frame assignment rules. In: Proceedings of the LREC 2004 Workshop on (2004)
4. Frank, A.: Question answering from structured knowledge sources. Journal of Applied Logic 5(1), 20 – 48 (2007)

5. Frank, A., van Genabith, J.: GlueTag. Linear logic based semantics for LTAG – and what it teaches us about LFG and LTAG. In: Butt, M., King, T.H. (eds.) Proceedings of the LFG01 Conference. Hong Kong (2001)
6. Frank, R.: Syntactic Locality and Tree Adjoining Grammar: Grammatical, Acquisition and Processing Perspectives. Ph.D. thesis, University of Pennsylvania (1992)
7. Gardent, C., Kallmeyer, L.: Semantic Construction in FTAG. In: Proceedings of EACL 2003. pp. 123–130. Budapest (2003)
8. Goldberg, A.E.: Constructions. A Construction Grammar Approach to Argument Structure. Cognitive Theory of Language and Culture, The University of Chicago Press, Chicago and London (1995)
9. Joshi, A.K., Schabes, Y.: Tree-Adjoning Grammars. In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages, pp. 69–123. Springer, Berlin (1997)
10. Joshi, A.K., Vijay-Shanker, K.: Compositional semantics with lexicalized Tree-Adjoining Grammar (LTAG): How much underspecification is necessary? In: Blunt, H.C., Thijsse, E.G.C. (eds.) Proceedings of the Third International Workshop on Computational Semantics (IWCS-3). pp. 131–145. Tilburg (1999)
11. Kageyama, T.: Denominal verbs and relative salience in lexical conceptual structure. In: Kageyama, T. (ed.) Verb semantics and syntactic structure, pp. 45–96. Kurioso Publishers, Tokyo (1997)
12. Kallmeyer, L., Joshi, A.K.: Factoring Predicate Argument and Scope Semantics: Underspecified Semantics with LTAG. Research on Language and Computation 1(1–2), 3–58 (2003)
13. Kallmeyer, L., Osswald, R.: A frame-based semantics of the dative alternation in lexicalized tree adjoining grammars (2012), submitted to Empirical Issues in Syntax and Semantics 9
14. Kallmeyer, L., Romero, M.: Scope and situation binding in LTAG using semantic unification. Research on Language and Computation 6(1), 3–52 (2008)
15. Kennedy, C., Levin, B.: Measure of change: The adjectival core of degree achievements. In: McNally, L., Kennedy, C. (eds.) Adjectives and adverbs. syntax, semantics, and discourse. Oxford University Press, Oxford (2008)
16. Levin, B., Rappaport Hovav, M.: Morphology and lexical semantics. In: Spencer, A., Zwicky, A.M. (eds.) Handbook of morphology, pp. 248–271. Blackwell Publishers, Oxford (1998)
17. Osswald, R., Van Valin, Jr., R.: Framenet, frame structure, and the syntax-semantics interface (2012), `http://www.phil-fak.uni-duesseldorf.de/fileadmin/Redaktion/Institute/Allgemeine_Sprachwissenschaft/Dokumente/Rainer_Osswald/FrameStructure-Osswald-VanValin-2012-02-20.pdf`, manuskript Heinrich-Heine Universität
18. Pinker, S.: Learnability and cognition: The aquisition of argument structure. MIT Press, Cambridge, MA (1989)
19. Van Valin, R.D., LaPolla, R.J.: Syntax. Cambridge University Press, Cambridge (1997)
20. Vijay-Shanker, K., Joshi, A.K.: Feature structures based tree adjoining grammar. In: Proceedings of COLING. pp. 714–719. Budapest (1988)