LiLT volume 16, issue (3) February 2019

Complex predicates: Structure, potential structure and underspecification

Stefan Müller, *Humboldt-Universität zu Berlin*

Complex predicates: Structure, potential structure and underspecification

Stefan Müller, *Humboldt-Universität zu Berlin*

16th February 2019

Abstract

This paper compares a recent TAG-based analysis of complex predicates in Hindi/Urdu with its HPSG analog. It points out that TAG combines actual structure while HPSG (and Categorial Grammar and other valence-based frameworks) specify valence of lexical items and hence potential structure. This makes it possible to have light verbs decide which arguments of embedded heads get realized, somthing that is not possible in TAG. TAG has to retreat to disjunctions instead. While this allows straight-forward analyses of active/passive alternations based on the light verb in valence-based frameworks, such an option does not exist for TAG and it has to be assumed that preverbs come with different sets of arguments.

1 Introduction

While comparing current syntactic theories (Hagemann and Staffeldt, 2014, Müller, 2010a, 2018, Kertész et al., 2019), one may get the impression that the frameworks are rather similar and easily translatable into each other. One reason for this is that we all deal with the same data and provided that we made similar categorial distinctions in the raw analysis of the data it would be a surprise if the analyses were radically different. However, it is not all the same. It matters which formalism is chosen and some are able to express intuitions rather directly while others are not.

This brief discussion note deals with the analysis of complex predicates consisting of a preverb and a light verb. Preverbs often have an argument structure of their own. They describe an event and the light verb can be used to realize either the full number of arguments or a reduced set of arguments. (1) provides the example from Hindi discussed by Ashwini et al. (2019).

(1) a. logon=ne pustak=kii tareef k-ii
people.M=ERG book.F=GEN praise.F do-PERF.F
'People praised the book.' Lit: 'People did praise of the book.'

b. pustak=kii tareef hu-ii
book.F.SG=GEN praise.F be.PART-PERF.F.SG/be.PRES
'The book got praised.' Lit: 'The praise of the book happened.'

Similar examples can of course be found in other languages making heavy use of complex predicates (Müller, 2010b).

Ashwini et al. (2019) assume that the structures for the examples in (1) are composed of elementary trees for *tareef* 'praise' and the respective light verbs. This is shown in Figure 1 and Figure 2 respectively. The TAG analysis is only sketched here. The authors use feature-based TAG, which makes it possible to enforce obligatory adjunction: the elementary tree for *tareef* is specified in a way that makes it necessary to take the tree apart and insert nodes of another tree. This way it can be ensured that the preverb has to be augmented by a light verb. This results in XP_f being inserted at XP_1 in Figure 1 and Figure 2.

What the analysis clearly shows is that TAG assumes two lexical items for the preverb: one with two arguments for the active case and one with just one argument for the passive. In general one would say that *tareef* is a noun that describes the praising event, that is, one person praises another one. Now this noun can be combined with a light verb and depending on which light verb is used we get an active

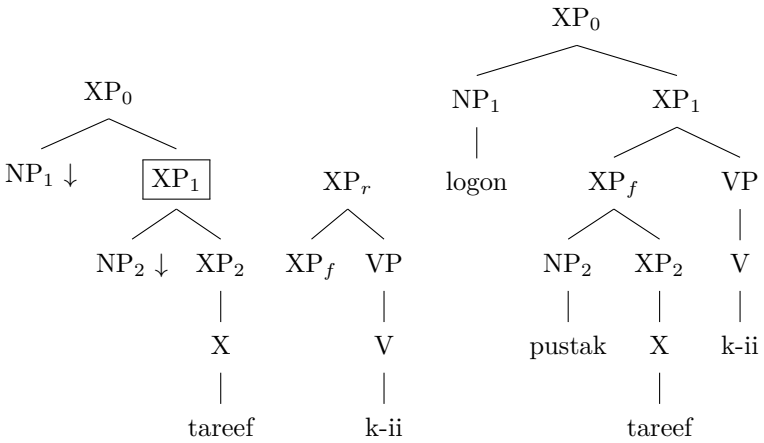

FIGURE 1 Analysis of *logon=ne pustak=kii tareef k-ii* 'People praised the book.' The tree of the light verb is adjoined into the tree of the preverb, into the XP_1 position

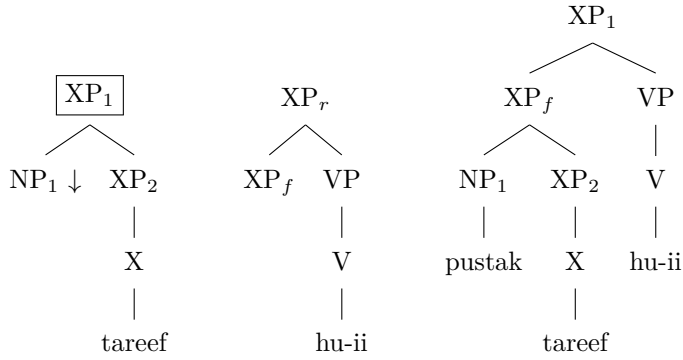

FIGURE 2 Analysis of *pustak=kii tareef hu-ii* 'The book got praised.'

sentence with both arguments realized or a passive sentence with the agent of the eventive noun suppressed. There is no morphological reflex of this active/passive alternation at the noun. It is just the same noun *tareef*: in an active sentence in (1a) and in a passive one in (1b).

And here we see a real difference between the frameworks: TAG (Joshi et al., 1975) is a framework in which structure is assembled: the basic operations are substitution and adjunction. The lexicon consists of ready-made building blocks that are combined to yield the trees we want to have in the end. This differs from Categorial Grammar (Ajdukiewicz, 1935) and HPSG (Pollard and Sag, 1994, Sag, 1997) where lexical items do not encode real structure to be used in an analysis, but potential structure: lexical items come with a list of their arguments, that is, items that are required for the lexical element under consideration to project to a full phrase. However, lexical heads may enter relations with their valents and form NPs, APs, VPs, PPs or other phrases, but they do not have to. Geach (1970) developed a technique that is called functional composition or argument composition within the framework of Categorial Grammar and this was transferred to HPSG by Hinrichs and Nakazawa (1989, 1994). Since the 90ies this technique is used for the analysis of complex predicates in HPSG for German (Hinrichs and Nakazawa, 1994, Kiss, 1995, Meurers, 2000, Müller, 1999, 2002), Romance (Miller and Sag, 1997, p. 600; Monachesi, 1998; Abeillé and Godard, 2002), Korean (Chung, 1998), and Persian (Müller, 2010b). See Godard and Samvelian, 2019 for an overview. For instance Müller (2010b, p. 642) analyzes the light verbs *kardan* 'do' and *odan* 'become' this way: both raise the subject of the embedded predicate and make it their own argument but *kardan* introduces an additional argument while *odan* does not do so.

Applying the argument composition technique to our example, we get the following lexical item for *tareef*:

(2) Sketch of lexical item for *tareef* 'praise':

$$\begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{noun} \\ \textsc{subj} & \langle\ \boxed{1}\ \rangle \end{bmatrix} \\ \textsc{comps} & \langle\ \boxed{2}\ \text{NP}\ \rangle \\ \textsc{arg-st} & \langle\ \boxed{1}\ \text{NP},\ \boxed{2}\ \text{NP}\ \rangle \end{bmatrix}$$

The ARG-ST list contains all arguments of a head. The arguments are liked to the semantic representation and are mapped to valence features like SPECIFIER and COMPLEMENTS. Depending on the langauge and the realizationability of subjects within projections, that subject may be mapped to a separate feature, which is a HEAD feature (Kiss, 1995,

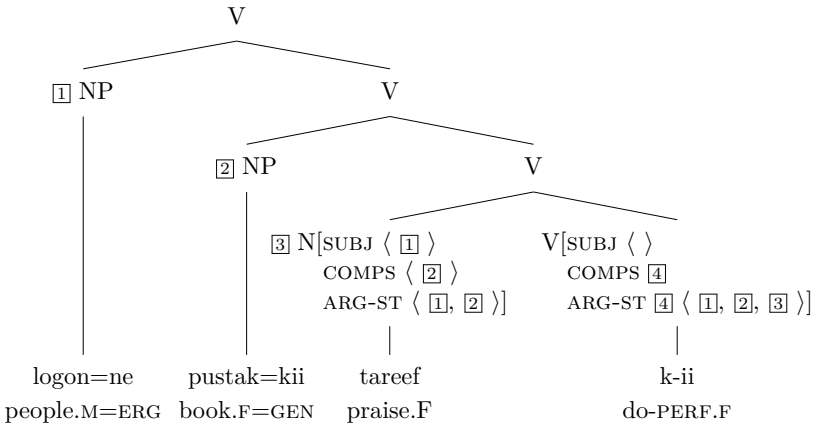

FIGURE 3 Analysis of *logon=ne pustak=kii tareef k-ii* 'People praised the book.' The arguments of the preverb are taken over by the light verb

Section 3.1.1). HEAD features are projected along the head path but the features contained under HEAD do not license combinations with the head.

The lexical items for *kar* 'do' and *ho* 'be' are:

(3) a. Sketch of lexical item for *kar* 'do':

$$\begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \boxed{1} \oplus \boxed{2} \oplus \left\langle\ \text{N}[\text{SUBJ } \boxed{1},\ \text{COMPS } \boxed{2}\ ]\ \right\rangle \end{bmatrix}$$

b. Sketch of lexical item for *ho* 'be':

$$\begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{ARG-ST} & \boxed{1} \oplus \left\langle\ \text{N}[\text{COMPS } \boxed{1}\ ]\ \right\rangle \end{bmatrix}$$

The verb *kar* 'do' selects for a noun and takes whatever the SUBJ value of this noun is ($\boxed{1}$) and concatenates the list of complements the noun takes ($\boxed{2}$) with the value of SUBJ. The result is $\boxed{1} \oplus \boxed{2}$ and it is a prefix of the ARG-ST list of the light verb. The lexical item for *ho* 'be' is similar, the difference being that the subject of the embedded verb is not attracted to the higher ARG-ST list, only the complements ($\boxed{1}$) are.

For finite verbs it is assumed that all arguments are mapped to the COMPS list of the verb, so the COMPS list is identical to the ARG-ST list.

The analysis of our example sentences is shown in Figures 3 and 4.

The conclusion is that HPSG has a representation of potential structure. When light verbs are present, they can take over valents and "ex-

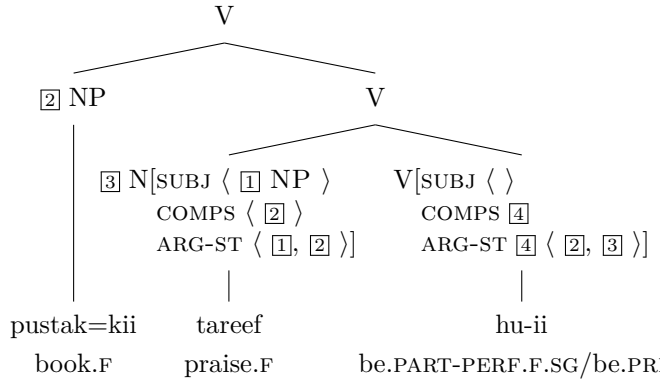

FIGURE 4 Analysis of *pustak=kii tareef hu-ii* 'The book got praised.'

ecute" them according to their own preferences. This is not possible in TAG since once structure is assembled it cannot be changed. We may insert items into the middle of an already assembled structure but we cannot take out arguments or reorder them. This is possible in Categorial Grammar and in HPSG: the governing head may choose which arguments to take over and in which order they should be represented in the valence repsresentations of the governing head.

LFG is somewhere in the middle between TAG and HPSG: the phrase structural configurations are not fully determined as in TAG since LFG does not store and manipulate phrase markers. But lexical items are associated with f-structures and these f-structures are responsible for which elements are realized in syntax. As complex predicates are assumed to be monoclausal it is not sufficient to embed the f-structure of the preverb within the f-structure of the light verb (Butt et al., 2003). Since the grammatical functions that are ultimately realized in the clause do not depend on the preverb alone the light verb may have to determine grammatical functions contributed by the preverb. In order to be able to do this Butt et al. (2003) use the restriction operator (Kaplan and Wedekind, 1993), which restricts out certain features or path equations provided by the preverb's and the light verb's f-structures. This is another instance of too strict specifications: once specified, it is difficult to get rid of it and special means like partial copying via restriction are needed. The alternative not relying on restriction was suggested by Butt (1997): embedding relations can be specified on the a-structure representation and then a mapping is defined that maps the complex a-structure to the desired f-structure. Mapping between several levels of representation is a general tool that

is also used in HPSG: for instance, Bouma, Malouf, and Sag (2001) used ARG-ST, DEPS, and COMPS in the treatment of nonlocal dependencies. See also Koenig (1999) on the introduction of arguments via additional auxiliary features. As I showed in Müller (2007, Section 7.5.2.2), one would need an extra feature for every kind of argument alternation that is to be modeled this way.

Summing up, I showed that there are indeed differences between the frameworks that are due to the basic representational formalisms they assume. While TAG assumes that the lexicon contains trees with a certain structure, HPSG assumes that lexical items come with valence specifications, that is, they have descriptions that together with dominance schemata (grammar rules) that are separate from the lexical items determine how possible structures look like. Since valence representations can be composed by superordinate predicates there is enough flexibility to deal with various light verb phenomena. LFG is a bit more constrained due to the use of f-structures, but using a restriction operator unwanted information about grammatical functions can be keept out of f-structures of matrix predicates.

Acknowledgments

I thank Miriam Butt, Timm Lichte, Rainer Osswald and Ashwini Vaidya for discussion and Sebastian Nordhoff for comments on this paper and for pointing out typos.

References

Abeillé, Anne and Danièle Godard. 2002. The syntactic structure of French auxiliaries. *Language* 78(3):404–452.

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1:1–27.

Ashwini, Vaidya, Owen Rambow, and Martha Palmer. 2019. Syntactic composition and selectional preferences in Hindi light verb constructions. *Linguistic Issues in Language Technology* 16(1).

Bouma, Gosse, Robert Malouf, and Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* 19(1):1–65.

Butt, Miriam. 1997. Complex predicates in Urdu. In A. Alsina, J. Bresnan, and P. Sells, eds., *Complex Predicates*, no. 64 in CSLI Lecture Notes, pages 107–149. Stanford, CA: CSLI Publications.

Butt, Miriam, Tracy Holloway King, and John T. Maxwell III. 2003. Complex predication via restriction. In M. Butt and T. H. King, eds., *Proceedings of the LFG 2003 Conference*, pages 92–104. Stanford, CA: CSLI Publications.

Chung, Chan. 1998. Argument composition and long-distance scrambling in Korean. In Hinrichs et al. (1998), pages 159–220.

Geach, Peter Thomas. 1970. A program for syntax. *Synthese* 22:3–17.

Godard, Danièle and Pollet Samvelian. 2019. Complex predicates. In S. Müller, A. Abeillé, R. D. Borsley, and J.-P. Koenig, eds., *Head-Driven Phrase Structure Grammar: The Handbook*, Empirically Oriented Theoretical Morphology and Syntax, pages 233–268. Berlin: Language Science Press. To appear.

Hagemann, Jörg and Sven Staffeldt, eds. 2014. *Syntaxtheorien: Analysen im Vergleich*. No. 28 in Stauffenburg Einführungen. Tübingen: Stauffenburg Verlag.

Hinrichs, Erhard W., Andreas Kathol, and Tsuneko Nakazawa, eds. 1998. *Complex Predicates in Nonderivational Syntax*. No. 30 in Syntax and Semantics. San Diego: Academic Press.

Hinrichs, Erhard W. and Tsuneko Nakazawa. 1989. Subcategorization and VP structure in German. In *Aspects of German VP Structure*, SfS-Report-01-93. Tübingen: Eberhard-Karls-Universität Tübingen.

Hinrichs, Erhard W. and Tsuneko Nakazawa. 1994. Linearizing AUXs in German verbal complexes. In J. Nerbonne, K. Netter, and C. J. Pollard, eds., *German in Head-Driven Phrase Structure Grammar*, no. 46 in CSLI Lecture Notes, pages 11–38. Stanford, CA: CSLI Publications.

Joshi, Aravind K., Leon S. Levy, and Masako Takahashi. 1975. Tree Adjunct Grammar. *Journal of Computer and System Science* 10(2):136–163.

Kaplan, Ronald M. and Jürgen Wedekind. 1993. Restriction and correspondence-based translation. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics*, pages 193–202. Columbus, Ohio, USA: Association for Computational Linguistics.

Kertész, András, Edith Moravcsik, and Csilla Rákosi, eds. 2019. *Current Approaches to Syntax: A Comparative Handbook*. No. 3 in Comparative Handbooks of Linguistics. Berlin: Mouton de Gruyter. To appear.

Kiss, Tibor. 1995. *Infinite Komplementation: Neue Studien zum deutschen Verbum infinitum*. No. 333 in Linguistische Arbeiten. Tübingen: Max Niemeyer Verlag.

Koenig, Jean-Pierre. 1999. *Lexical Relations*. Stanford Monographs in Linguistics. Stanford, CA: CSLI Publications.

Meurers, Walt Detmar. 2000. Lexical generalizations in the syntax of German non-finite constructions. Arbeitspapiere des SFB 340 No. 145, Eberhard-Karls-Universität, Tübingen.

Miller, Philip H. and Ivan A. Sag. 1997. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory* 15(3):573–639.

Monachesi, Paola. 1998. Italian restructuring verbs: A lexical analysis. In Hinrichs et al. (1998), pages 313–368.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ: Head-Driven Phrase Structure Grammar für das Deutsche*. No. 394 in Linguistische Arbeiten. Tübingen: Max Niemeyer Verlag.

Müller, Stefan. 2002. *Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German*. No. 13 in Studies in Constraint-Based Lexicalism. Stanford, CA: CSLI Publications.

Müller, Stefan. 2007. *Head-Driven Phrase Structure Grammar: Eine Einführung*. No. 17 in Stauffenburg Einführungen. Tübingen: Stauffenburg Verlag, 1st edn.

Müller, Stefan. 2010a. *Grammatiktheorie*. No. 20 in Stauffenburg Einführungen. Tübingen: Stauffenburg Verlag.

Müller, Stefan. 2010b. Persian complex predicates and the limits of inheritance-based analyses. *Journal of Linguistics* 46(3):601–655.

Müller, Stefan. 2018. *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches*. No. 1 in Textbooks in Language Sciences. Berlin: Language Science Press, 2nd edn.

Pollard, Carl J. and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Chicago: The University of Chicago Press.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33(2):431–484.