

Light verb constructions in Computational Grammars

TAG and LFG

IIT Delhi

17/11

Computational grammars

- Computational grammars: provide a way to model complex linguistic properties
- Light verb constructions offer a particular challenge:-
 - Two predicating elements in a monoclausal structure
 - Combinatorial properties → varied meanings
- Lexicalized grammars can be particularly relevant

Light verb construction

Light verb construction

“two or more predicational elements each contribute to a joint predication” [Butt, 2010]. E.g. *give a kiss*

The verb (aka ‘light verb’ [Jespersen, 1965]) consists of predicates e.g. *give, make, take* that are found cross-linguistically (e.g. English, Persian, Korean).

A (predicating) *preverb* combines with a *verb* to give us a joint predicate in Hindi.

- This pre-verb may be an adjective (*saaf* ‘clean’), noun (*chorii* ‘theft’) or borrowed English word (*download*).

Two predicating elements

- (1) 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) '
- (2) raam=ne mohan=par shaq ki-yaa
Ram.M.Sg=Erg Mohan.M.Sg=Loc suspicion.m do-Perf.M.Sg
'Ram suspected Mohan; (Lit: Ram did suspicion on Mohan)'

Hindi light verbs

- Not all light verb constructions occur with *kar* 'do', although it is the most productive light verb in Hindi
- Others include *ho* 'be', *de* 'give', *le* 'take', *rakh* 'keep' etc.

Alternation with *kar*, *ho* ('do' , 'be')

- (3) manriji=ne mandir=kaa udghaatan
minister.M.Sg=Erg temple.M.Sg=Gen inauguration.M
ki-yaa
do-Perf.M.Sg
'The minister inaugurated the temple'
- (4) mandir=kaa udghaatan hu-aa
temple.M.Sg=Gen inauguration.M be-Perf.M.Sg
'(The) inauguration of the temple happened'

Alternation

- Predicating nouns in light verb constructions can also *alternate* with these light verbs
- This results in causative/inchoative 'pairs' e.g. *kar* 'do' and *ho* 'be', *aa* 'come' and *dilvaa* 'cause to give' (T. Ahmed)

Alternations

- Alternation with *ho* and *kar* is a characteristic of a certain group of nominals [Ahmed and Butt, 2011]
- Also highly productive, as many as 15% of all LVCs in the Hindi Treebank undergo this alternation
- Found cross-linguistically (Korean, Persian)

Unlike other alternations (active-passive), there are *two* predicating elements

LFG Analysis

- Light verb constructions aka ‘complex predicates’ in Lexical Functional Grammar
- In this analysis, preverb and light verb are co-predicators, i.e. both come equipped with their own subcategorization frames
- Argument merger operation used to combine the argument structures of noun and light verb
- In contrast, [Grimshaw and Mester, 1988] propose ‘argument transfer’ where the light verb has no arguments of its own

Example

- (5) meNdak=ne bicchu=se bahas kii
frog.M.Sg=Erg scorpion.M.Sg=Inst argument.F.Sg do.Perf.F.S

'The frog argued with the scorpion. (Lit) The frog did (an) argument with the scorpion'

(↑ PRED) = 'do <AGENT, %PRED >'

(↑ PRED) = 'argument <AGENT, THEME >'

Figure: The form of the light verb supplies another PRED

- %PRED is an augmentation in the LFG framework, needed to compose arg structures [Butt et al., 2003]
- The Restriction Operator is required to manipulate the f-structure of *bahas* 'argument'
- In effect, the light verb *do* over-writes the argument structure of the predicating noun 'bahas'

PRED 'Nadya' PERS 3 NUM sg CASE erg	$\left[\begin{array}{l} \text{PRED 'Nadya'} \\ \text{PERS 3} \\ \text{NUM sg} \end{array} \right]$
----------------------------------------------	-----------------------------------------------------------------------------------------------------

Figure: The restriction operator has the ability to restrict out information e.g. the CASE feature from the attribute-value matrix (AVM) for the lexical item 'Nadya'. (\uparrow CASE) gives us the restricted second AVM in this figure.

PRED	‘kar	⟨	[SUBJ],	[-1]⟩]
SUBJ	[CASE	erg]		
OBJ	[GEND	fem, NUM	sg]	
CHECK	[_VMORPH	[_VTYPE	infl]	
	[_RESTRICTED	- , _VFORM	perf]	
LEX-SEM	[AGENTIVE	+]	
TNS-ASP	[ASPECT	perf]		
VTYPE	[COMPLEX-PRED-FORM	kar]		
PASSIVE	-				

Figure: F-structure for the light verb *kar* ‘do’

Co-predication

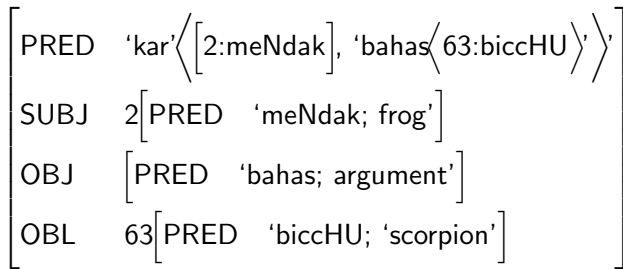


Figure: Final (abbreviated) F-structure for the NVC *bahas kar* 'debate do'. Note that this is still a monoclausal structure.

Alternations?

(6) bicchu=se bahas huii
scorpion.M.Sg=Inst argument.F.Sg be.Perf.F.S
'The debate with the scorpion happened'

- The LFG account is not explicit about light verb constructions with *ho*
- Perhaps possible to achieve argument merger with the restriction operator
- But it is not clear whether their account considers them to be LVCs

A possibility?

(↑ PRED) = 'be < %PRED >'

(↑ PRED) = 'argument <AGENT, THEME >'

Figure: This time, *be* has no arguments except %PRED ?

Restriction

- Restricting is similar to a deletion/transformation operation
- Becomes necessary due to the stipulation of co-predication in LFG
- Maintain mono-clausality and co-predication– without deletion operations?
- Recent work in LFG has attempted to do away with Restriction [Lowe, 2015]

Tree Adjoining Grammar

- Propose that noun and light verb are still co-predicators
- Although the noun will specify the arguments of the light verb construction
- The contribution of light verbs specified using features
- Framework: Feature-based tree-adjoining grammar

- The predicating noun/light verb's elementary trees comprise 'complex primitives'
- Combinatorial possibilities of argument structures constrained by features
- Two general operations of substitution and adjunction used to combine them

Tree Adjoining Grammar

- In TAG terms, the noun is the initial tree with all the arguments, the light verb is the auxiliary tree
- The initial tree remains derivationally incomplete to force adjunction of the light verb
- The light verb carries with its semantic feature [\pm AGENTIVE] (similar to LEX-SEM)

(7) *logon-ne pustak-kii tareef kii.*
people.M-Erg book.F-Gen praise.F do.Perf.F.Pl
'People praised the book '

- The initial tree for noun *tareef* 'praise' has the arguments *logon* and *pustak*
- It can only unify with an auxiliary tree which is [AGENTIVE=+]
- It is also under-specified with respect to lexical category [Han and Rambow, 2000]

Nominal: Initial tree

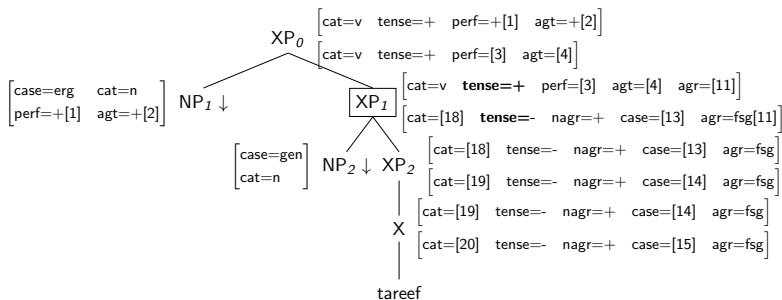


Figure: Tree for nominal *tareef* 'praise' (agentive), as seen in *logon ne pustak kii tareef kii* "People praised the book". The feature clash at XP_1 is marked with a box.

Light verb: auxiliary tree

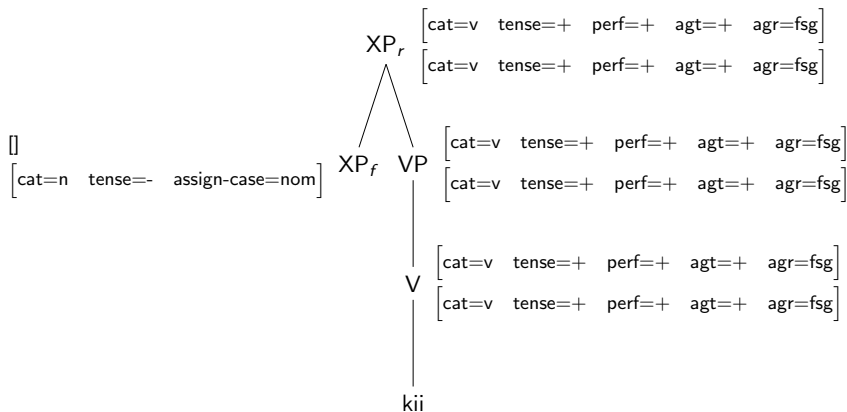


Figure: Elementary tree for light verb *kar* 'do' inflected as *kii* 'do.fem.sing.perf'

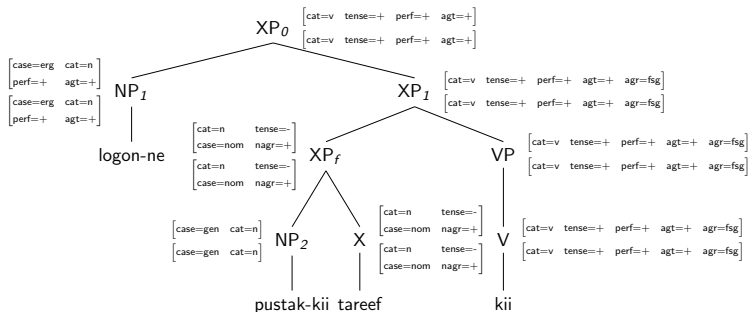


Figure: Post adjunction of the light verb's auxiliary tree into the initial tree *tareef* 'praise' at XP_2 , we get the complete argument structure. Substitution at the nodes NP_1 and NP_2 gives us *logon-ne pustak-kii tareef kii* 'People praised the book'

Alternation with *ho*

- (8) *pustak-kii tareef huii.*
book.F-Gen praise.F be.Perf.F
'The praise of the book happened'

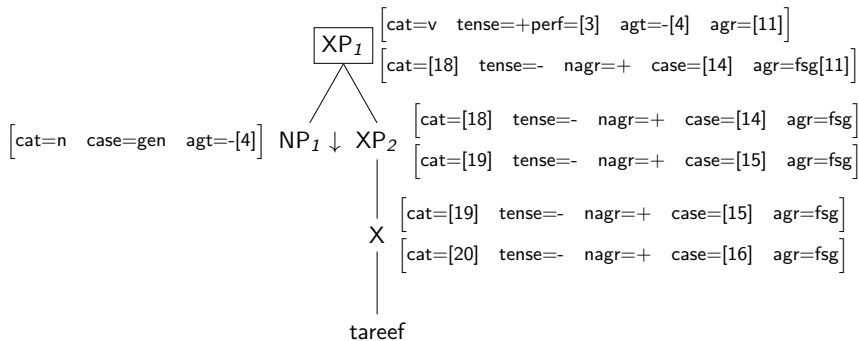


Figure: Tree for nominal *tareef* (non agentive) as seen in *pustak-kii tareef huii* '(the) book was praised'. The feature clash this time is at XP_1 and is marked with a box.

TAG




- The tree for non-agentive *tareef* will always combine with a light verb that is AGT=--,
- the auxiliary tree of the light verb *ho* 'be' will have [AGT=--].
- We can capture the light verb construction's combinatorial possibilities

TAG

- TAG carries out the “argument merger” operation without needing a deletion rule
- Two general operations of substitution and adjunction are utilized instead
- Capture other interesting properties relating to agreement facts (not covered in detail here)
- Disadvantage: many elementary trees specified for each nominal alternation

TAG

- The TAG analysis posits a separation of subcategorization and semantic information
- This captures the fact that the light verb contributes some semantics
- It also does this with greater 'burden' on the elementary trees without special operations

-  Ahmed, T. and Butt, M. (2011).
Discovering Semantic Classes for Urdu N-V Complex Predicates.
In Proceedings of the International Conference on Computational Semantics (IWCS 2011), Oxford.
-  Butt, M. (2010).
The Light Verb Jungle: Still Hacking Away.
In Amberber, M., Harvey, M., and Baker, B., editors, *Complex Predicates in Cross-Linguistic Perspective*, pages 48–78. Cambridge University Press.
-  Butt, M., King, T. H., and Maxwell III, J. T. (2003).
Productive Encoding of Urdu Complex Predicates in the ParGram Project.
In Proceedings of the EACL03 Workshop on Computational Linguistics for South Asian Languages: Expanding Synergies with Europe.
-  Grimshaw, J. and Mester, A. (1988).
Light verbs and theta-marking.

Linguistic Inquiry, 9(2):205–232.



Han, C. and Rambow, O. (2000).

The Sino-Korean light verb construction and lexical argument structure.

In *Proceedings of the Fifth International Workshop on Tree-Adjoining Grammars and Related Formalisms, TAG+5*.



Jespersen, O. (1965).

A Modern English Grammar on Historical Principles, Part VI, Morphology.

George Allen and Unwin Ltd.



Lowe, J. J. (2015).

Complex predicates: an lfg+glue analysis.

Journal of Language Modelling, 3(2):413–462.