Modeling Light Verb Constructions in the LinGO Grammar Matrix

Tara Wueger
University of Washington - Seattle

July 7, 2025

What is a Light Verb Construction?

- light verb construction (LVC): a type of complex predicate comprised of a light verb and a coverb, where most (but not necessarily all) of the lexical meaning of the combined complex predicate comes from the coverb

  - light verb: a verb that is semantically 'light' (Butt, 2010, p. 48) to some degree, ranging from contributing no lexical meaning (i.e. semantically bleached) to some but never all of the lexical meaning in an LVC
  - coverb: noun, verb, or adjective co-occurring with the light verb that could be used as its intended part of speech in the language and/or could be used to provide some or all of the lexical meaning in an LVC

- e.g. English *take a shower*

Overview of Analysis

- light verb is the head daughter
- coverb (or the constituent it heads) is non-head daughter

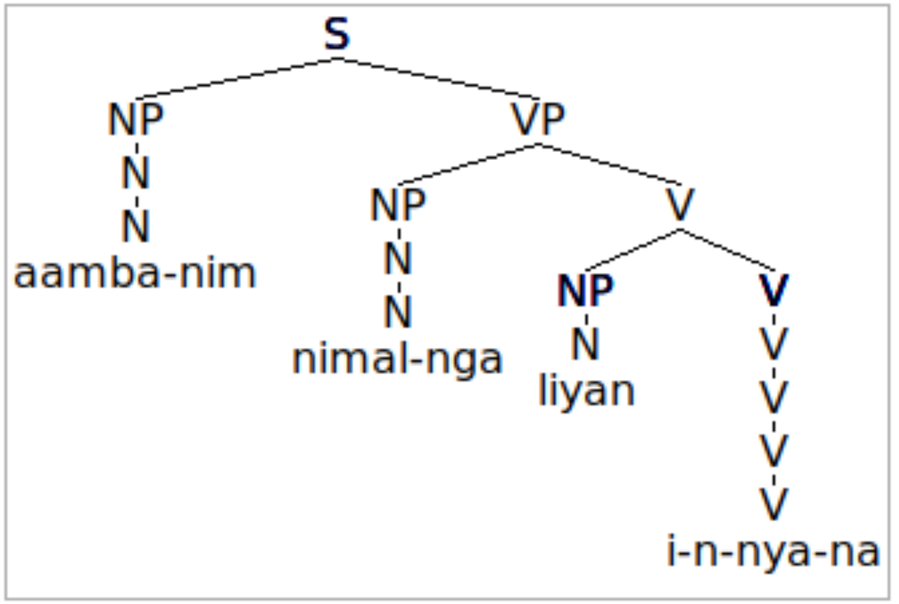
- 2 new phrase structure rules (specialized variants of head-complement rules) are used to combine them
  - *head-comp-lvc*
  - *comp-head-lvc*

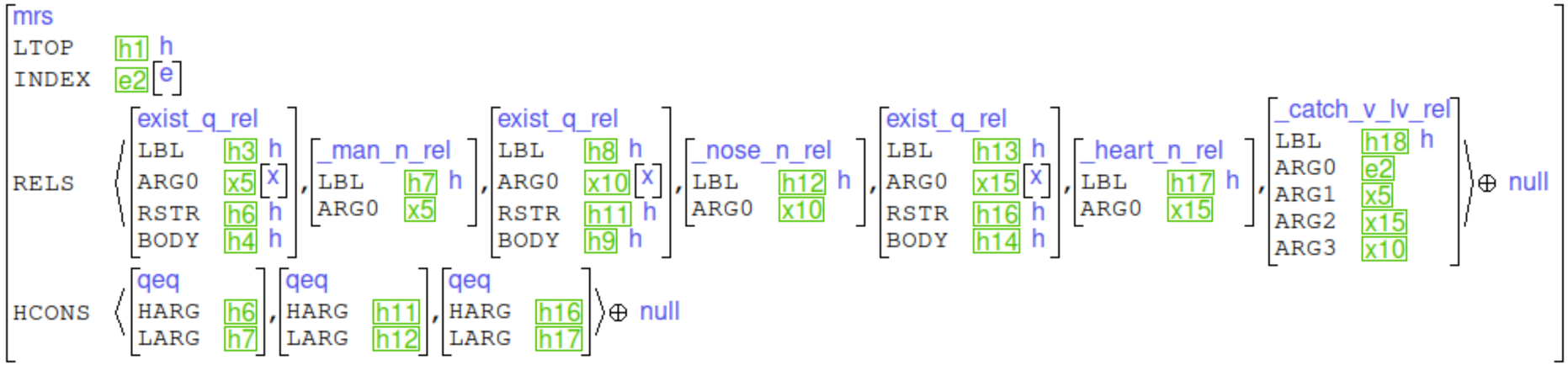
- generally, the light verb combines with the coverb (or constituent it heads) first, before it combines with any other constituents
  - exception for languages that allow elements to intervene between the light verb and coverb (like Persian), which will be discussed later

Example Sentences

Bardi [bcj] - Example with Desired Tree + MRS

Aamba nimalnga liyan innyana.
aamba-nim nimal-nga liyan i-n-nya-na
man-ERG nose-INS heart 3-TR-catch-PST
“The man breathed through his nose.” [bcj]
(Bowern, 2012, p. 440)





Bardi [bcj] - Example with Desired Tree + MRS

- transitive light verb *-nya-* “to catch”
- noun coverb *liyan* “heart”

- *comp-head-lvc* rule is used to form the LVC *liyan innyana* “breathed”

- in the MRS
  - the relation for the light verb (*_catch_v_lv_rel*) introduces the event *e2*
  - the coverb (*_heart_n_rel*) is identified with the ARG2 of the light verb (indicated by the *x15* instance)
  - the object (*nimalnga* “through the nose”) has a the *_nose_n_rel* relation and is identified with the ARG3 of the light verb (indicated by the *x10* instance)
  - the subject (*aambanim* “man”) is identified with the ARG1 of the light verb (this is the same as a heavy verb)

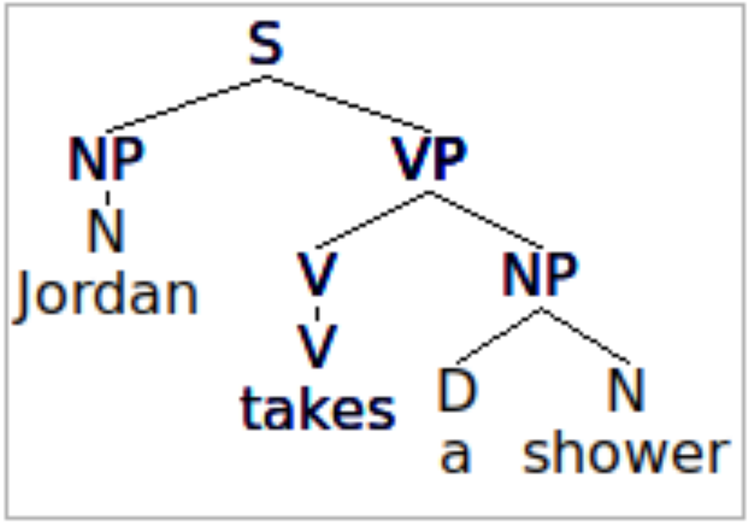
- unbleached light verbs (like *-nya-*) introduce an event and their ARG2s are always identified with the coverb
- in constructions with heavy verbs, the object would be identified with the ARG2 of the verb
  - since the coverb is identified with the ARG2, the object is instead identified with the ARG3

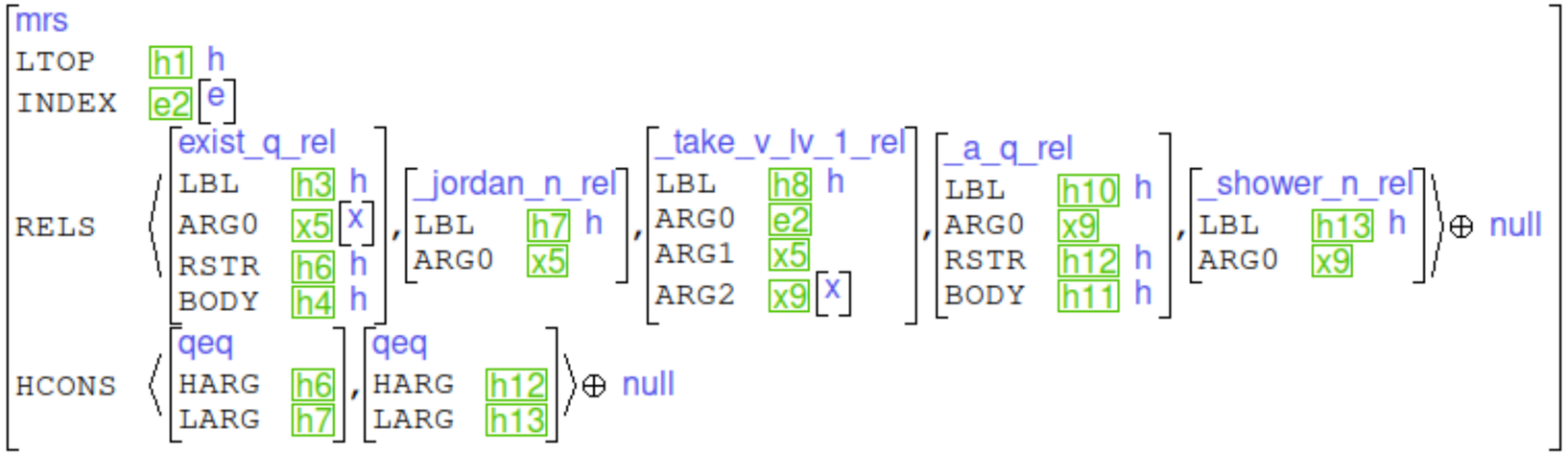
English [eng] - Example with Desired Tree + MRS

Jordan takes a shower.

Jordan	takes	a	shower
Jordan	take.PRS-3SG	DET.INDF	shower

“Jordan takes a shower.” [eng]





English [eng] - Example with Desired Tree + MRS

- intransitive light verb *take*
- noun coverb *shower*

- *head-comp-lvc* rule is used to form the LVC *take a shower*

- in the MRS
  - the coverb (*_shower_n_rel*) and its determiner (*_a_q_rel*) is identified with the ARG2 of the light verb (indicated by the *x9* instance)

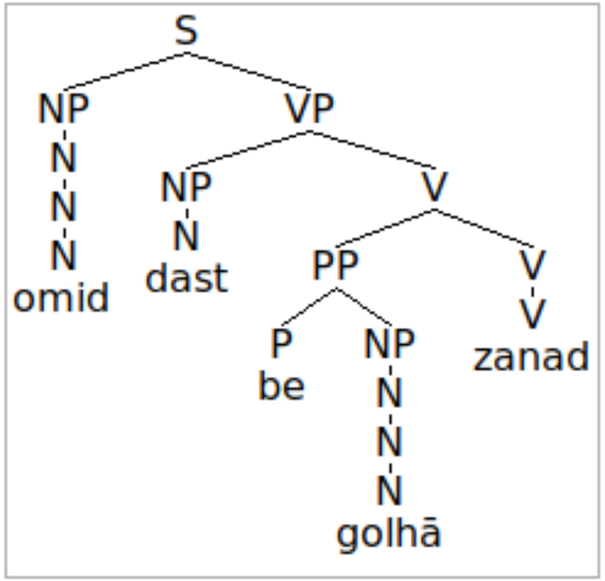
- elements are allowed to intervene between the light verb and coverb (in this case, a determiner) if they are constituents of the coverb
  - if this behavior is undesired, it can be blocked, which will be discussed later
  - cases where the intervening element is not a constituent of the coverb will also be discussed later

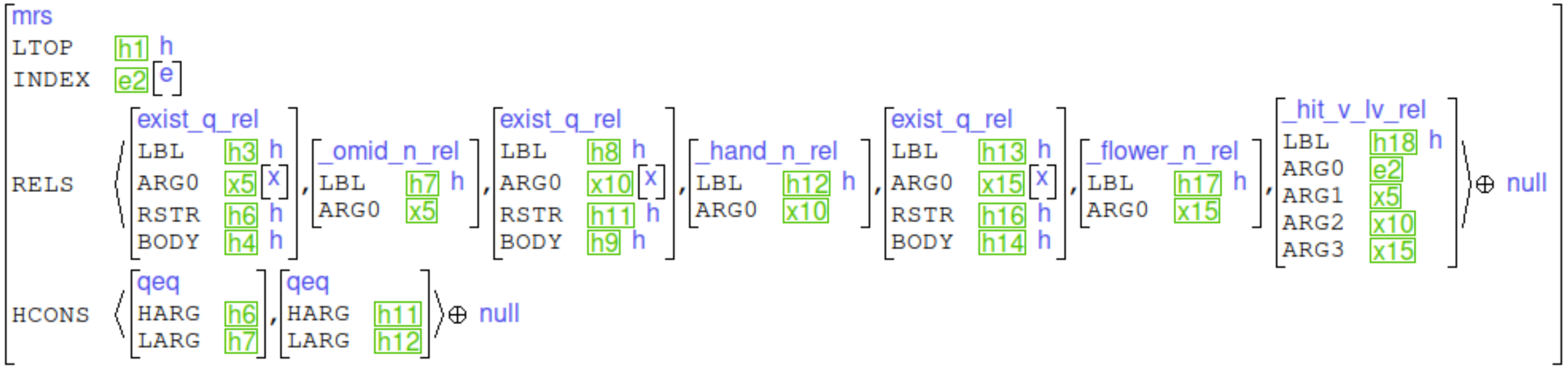
Persian [per] - Example with Desired Tree + MRS

Omid dast be golhā zanad.

Omid	dast	be	gol-hā	zan-ad
Omid	hand	to	flower-PL	hit-3SG

“Omid touches the flowers.” [per] (Godard & Samvelian, 2021, p. 470)





Persian [per] - Example with Desired Tree + MRS

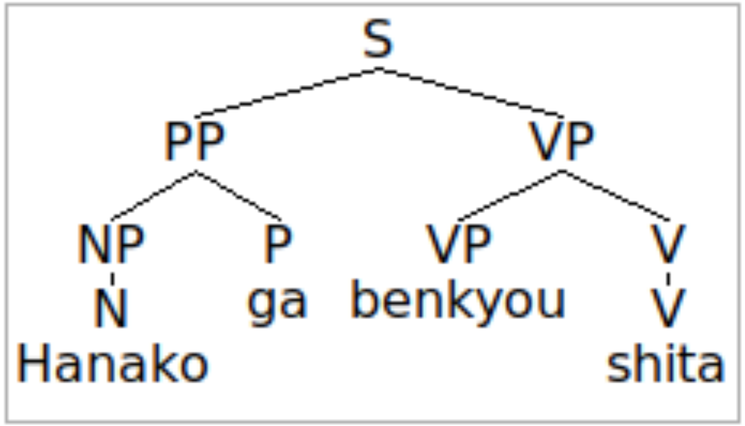
- transitive light verb *zad* “to hit”
- noun coverb *last* “hand”

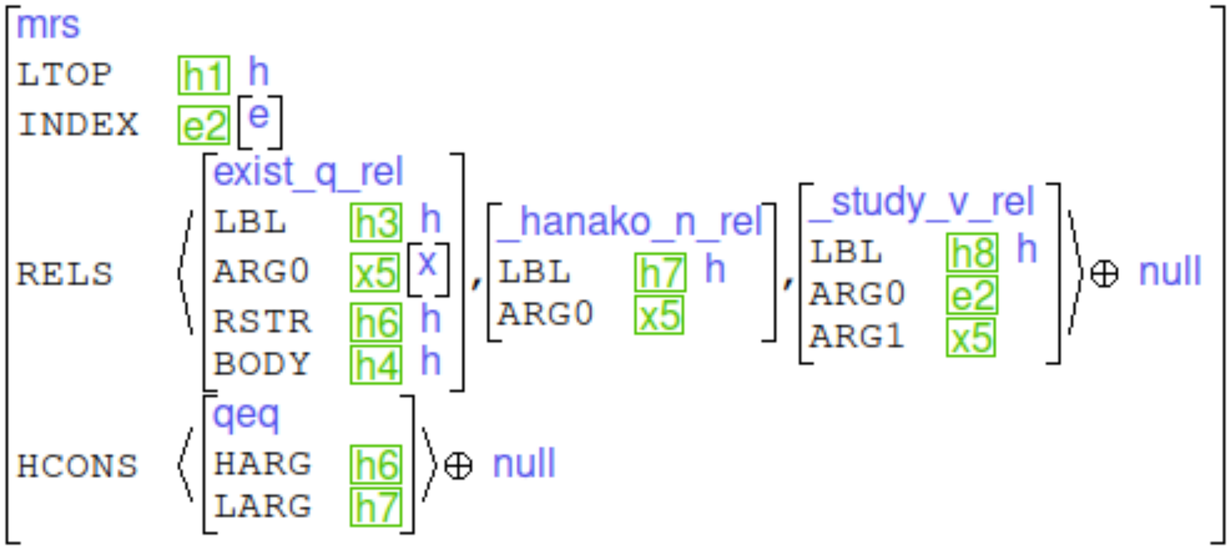
- an element (that is not a dependent of the coverb) is allowed to intervene between the coverb and light verb
  - this means that the light verb and coverb cannot combine first
  - the light verb and the object (*be golhā* “the flowers”) combine first (using the *comp-head-2* rule)
  - then it combines with the coverb using *comp-head-lvc*

- in the MRS
  - there is no difference with a sentence where the object is not intervening

Japanese [jpn] - Example with Desired Tree + MRS

Hanako ga benkyou shita.
Hanako ga benkyou shi-ta
Hanako NOM study do-PST
“Hanako studied.” [jpn]





*I analyze *benkyou* “study” as an intransitive coverb (taking only an ARG1 semantically), which might not be accurate for Japanese.

Japanese [jpn] - Example with Desired Tree + MRS

- bleached intransitive light verb *shi* “to do”
- verb coverb *benkyou* “study”

- *comp-head-lvc* rule is used to form the LVC *benkyou shita* “studied”
  - syntactically, not different from the unbleached cases (is different semantically)

- in the MRS
  - no relation for bleached light verb so arguments normally identified with the light verb are instead identified with the coverb
  - the coverb (*_study_v_rel*) introduces the event and the subject is identified with the ARG1 of the coverb (instead of the ARG1 of both the light verb and the coverb as with unbleached cases)

Analysis and Implementation

The LVC Feature

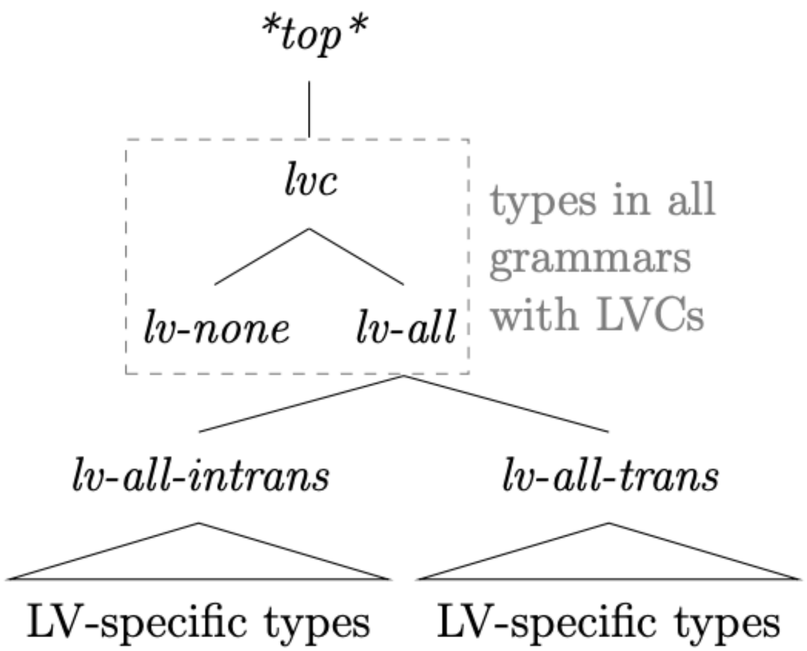
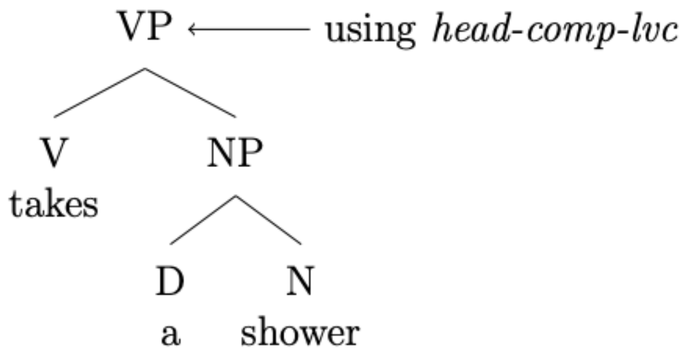


Figure 4.5: LVC Type Hierarchy

- specified on the HEAD of the coverb and allows a coverb to indicate which light verb(s) it can combine with
- e.g. in the Bardi *liyan innyana* “breathed”, the coverb *liyan* “heart” inherits from a lexical type which specifies a value of *lv-nya-tr-noun-trans* for its LVC feature
  - this value corresponds to the transitive light verb *-nya-* “to catch” that can combine with noun coverbs

New Phrasal Types

$$\left[\begin{array}{l} \textit{head-comp-phrase-lvc} \\ \text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{LOCAL} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS} \quad \left\langle \left[\text{LOCAL} \mid \text{CAT} \mid \text{HEAD} \mid \text{LVC} \quad \textit{lv-all}\right] \right\rangle \end{array}\right]$$



- *head-comp-phrase-lvc*
  - when the coverb appears after the light verb
  - makes first element of COMPS of head daughter be [LVC *lv-all*]
- some languages allow an element (that is not a dependent of the coverb) to intervene between the light verb and the coverb
  - to block this from happening in languages where it is not allowed, LIGHT + is added to the head daughter
  - (this applies to *head-comp-phrase-lvc* and/or *comp-head-phrase-lvc* and only works for languages with free or v2 word order
- also added [LVC *lv-none*] to non-head daughter of other head-complement rules to prevent being used by LVCs

New Phrasal Types

$$\left[\begin{array}{l} \textit{comp-head-phrase-lvc} \\ \text{HEAD-DTR} \mid \text{SYNSEM} \mid \text{LOCAL} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS} \quad \left\langle \text{LOCAL} \mid \text{CAT} \mid \text{HEAD} \mid \text{LVC} \quad \textit{lv-all} \right\rangle \end{array}\right]$$

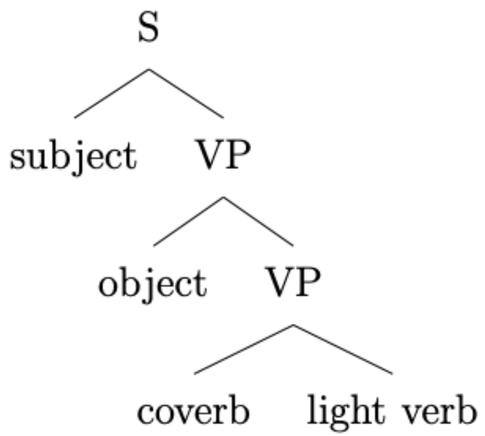
$$\left[\begin{array}{l} \textit{head-final-lvc} \\ \text{SYNSEM} \mid \text{ATTACH} \quad \textit{lmod} \end{array}\right]$$

- *comp-head-phrase-lvc*
  - when the coverb appears before the light verb
  - inherits from *head-final* or *head-final-lvc*
- *head-final-lvc*
  - for languages with free word order
  - allows the coverb to combine with the light verb first (i.e. before combining with other elements)

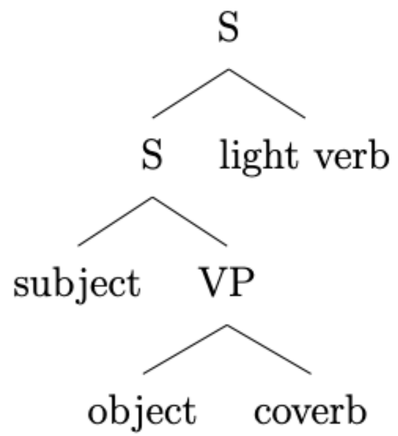
New Phrasal Types

- change to *decl-head-subj-phrase*
  - to prevent objects from combining with the coverb before the coverb has combined with the light verb (specifically in cases where the word order is subject, object, coverb, and then light verb)
  - requires *lv-none* for the LVC feature on the head daughter

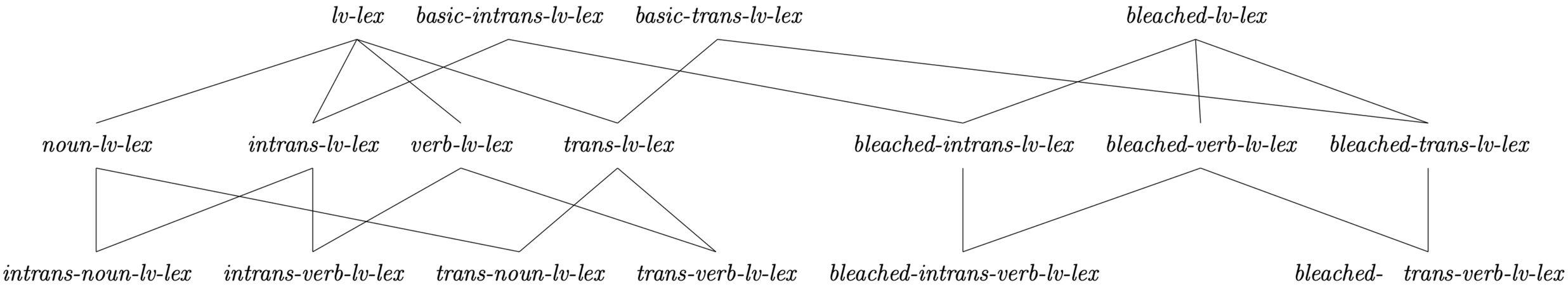
(a) Correct Combination Order



(b) Incorrect Combination Order



Hierarchy for LVC-related Types



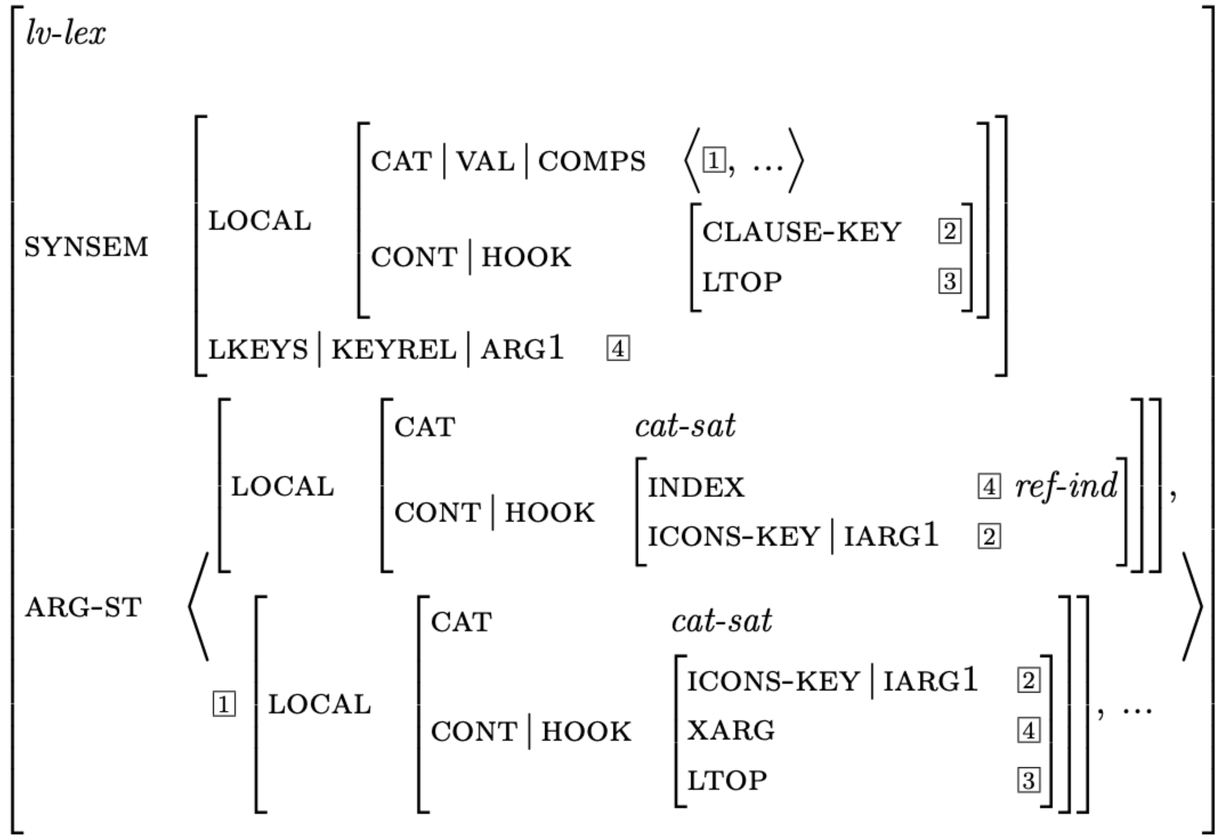
Persian [per] - Walkthrough of Lexical Types

- light verb types:
  - transitive light verb *zad* “to hit”
    - *zadan-past-tr-trans-noun-lv-lex*
      includes word-specific features (e.g. case), LVC value of acceptable coverb(s), inflection flags
      - *trans-noun-lv-lex*
        - *noun-lv-lex*
          - *lv-lex*
        - *trans-lv-lex*
          - *basic-trans-lv-lex*
          - *lv-lex*

Omid dast be golhā zanad.				
Omid	dast	be	gol-hā	zan-ad
Omid	hand	to	flower-PL	hit-3SG
“Omid touches the flowers.” [per] (Godard & Samvelian, 2021, p. 470)				

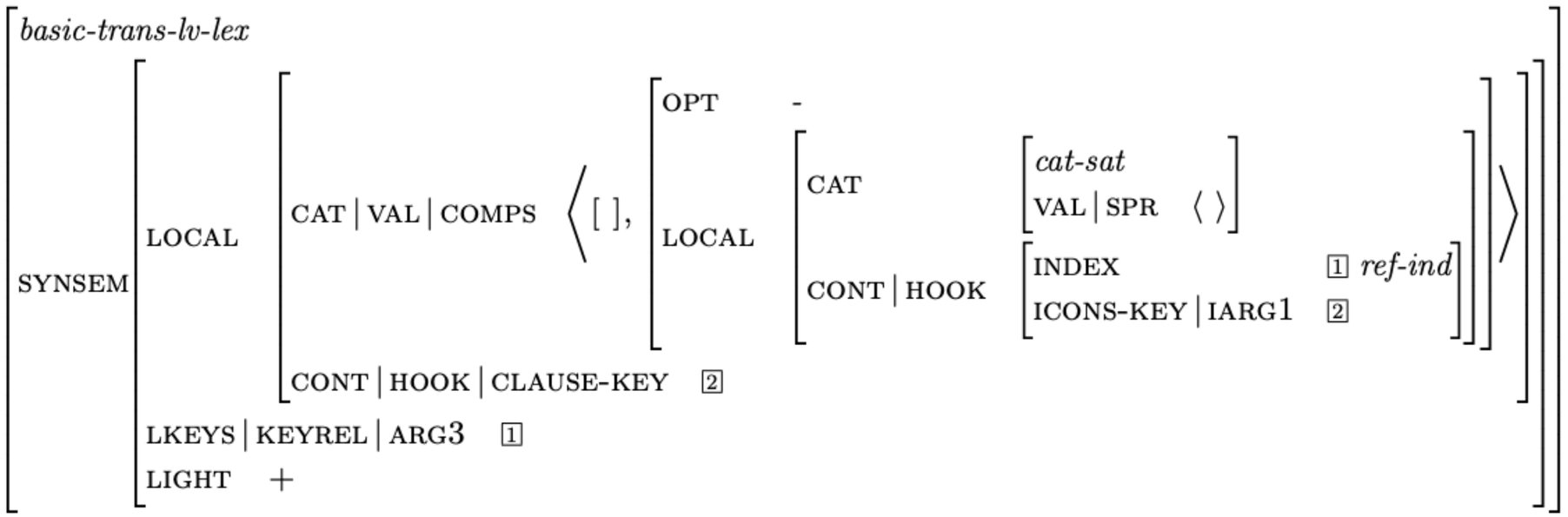
- coverb types:
  - noun coverb *dast* “hand”
    - *lv-zadan-past-tr-trans-coverb-noun-lex*
      includes LVC value for this coverb
      - *noun2-coverb-noun-lex*
        - *coverb-noun-lex* (inherits from *basic-noun-lex*)

Persian [per] - Walkthrough of Lexical Types



- also specifies the head type for the subject:
  - `[ SYNSEM.LOCAL.CAT.VAL.SUBJ.FIRST.LOCAL.CAT.HEAD +np ]`

Persian [per] - Walkthrough of Lexical Types



*identifying the index of the object with the ARG3 of the light verb needs to be moved to *trans-lv-lex* as it will not work properly for bleached light verbs (where it should be ARG2)

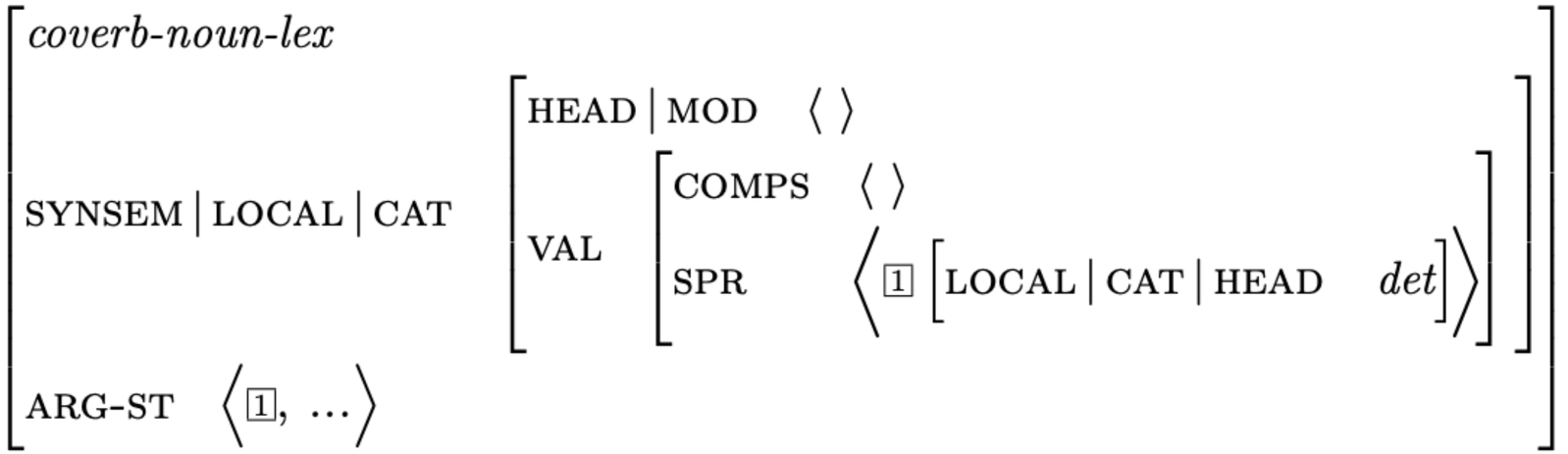
Persian [per] - Walkthrough of Lexical Types

$$\begin{bmatrix} \textit{noun-lv-lex} \\ \text{SYNSEM} \quad \begin{bmatrix} \text{LOCAL} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS} & \left\langle \boxed{1}, \ \ldots \right\rangle \\ \text{LKEYS} \mid \text{KEYREL} \mid \text{ARG2} & \boxed{2} \end{bmatrix} \\ \text{ARG-ST} \quad \left\langle [\ ], \boxed{1} \begin{bmatrix} \text{LOCAL} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{cat-sat} \\ \text{VAL} \mid \text{SPR} \quad \langle\ \rangle \end{bmatrix} \\ \text{CONT} \mid \text{HOOK} \mid \text{INDEX} & \boxed{2}\ \textit{ref-ind} \end{bmatrix} \end{bmatrix}, \ \ldots \right\rangle \end{bmatrix}$$

- Persian doesn't allow coverbs to take on dependents (similar to their non-coverb counterparts) so this must also be added:

$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOCAL} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS} & \left\langle \boxed{1}, \ \ldots \right\rangle \\ \text{ARG-ST} & \left\langle [\ ], \boxed{1} \begin{bmatrix} \text{LIGHT} & + \\ \text{MODIFIED} & \textit{notmod} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Persian [per] - Walkthrough of Lexical Types



- noun coverbs will get the same value as regular nouns for OPT on the first element in their SPR list (either + or −) (this is to account for optional, required, and impossible determiners)
- in Persian, this is OPT +

* current analysis does not account for complement-taking nouns (which is why COMPS is empty)
* noun coverbs should also be SPEC empty

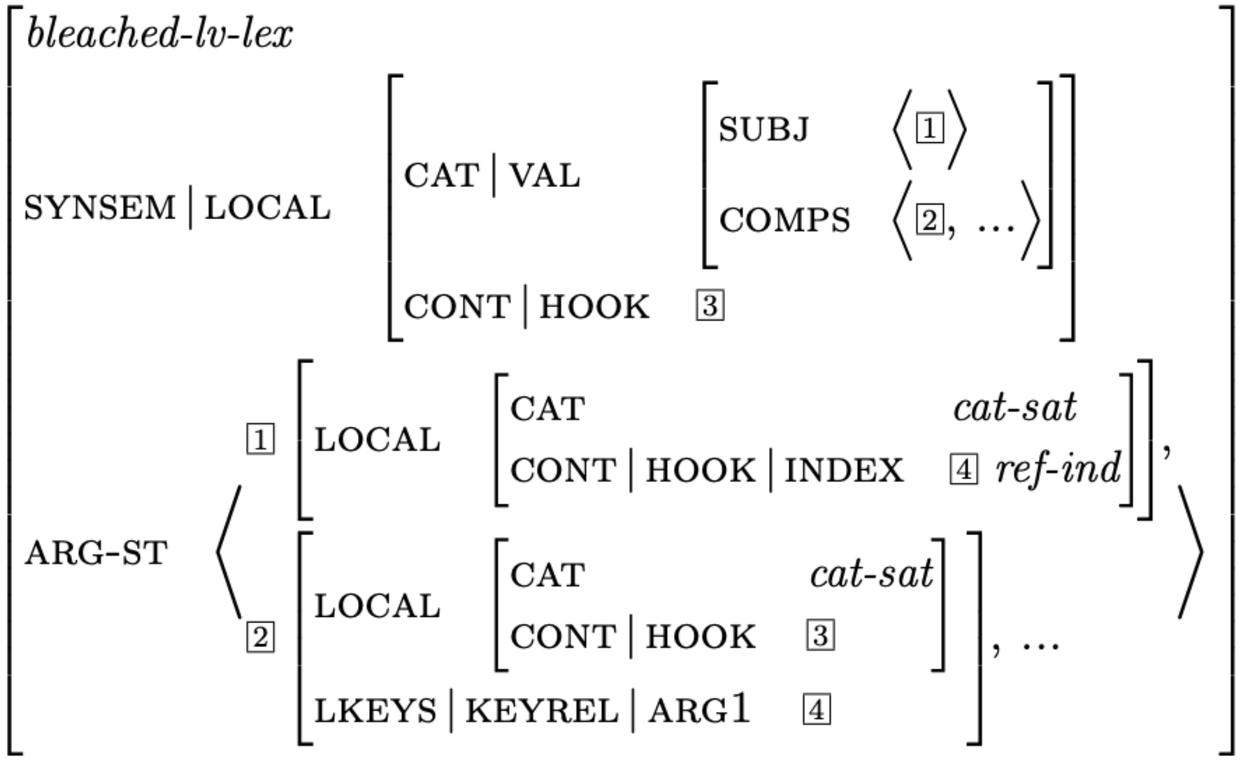
Japanese [jap] - Walkthrough of Lexical Types

- light verb types:
  - bleached intransitive light verb *shi* “to do”
    - *shi-it-bleached-intrans-verb-lv-lex*
      includes word-specific features (e.g. case), LVC value of acceptable coverb(s), inflection flags
      - *bleached-intrans-verb-lv-lex*
        - *bleached-verb-lv-lex*
          - *bleached-lv-lex*
        - *bleached-intrans-lv-lex*
          - *basic-intrans-lv-lex*
          - *bleached-lv-lex*

Hanako ga benkyou shita.			
Hanako	ga	benkyou	shi-ta
Hanako	NOM	study	do-PST
“Hanako studied.” [jpn]			

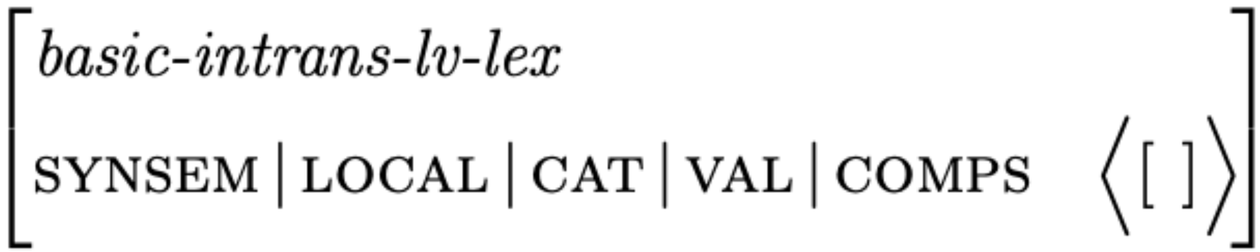
- coverb types:
  - verb coverb *benkyou* “study”
    - *verb3-lv-shi-it-intrans-coverb-verb-lex*
      includes LVC value for this coverb
      - *verb3-coverb-verb-lex*
        - *coverb-intrans-verb-lex* (inherits from *intransitive-verb-lex*)

Japanese [jap] - Walkthrough of Lexical Types



- also specifies the head type for the subject:
  - `[ SYNSEM.LOCAL.CAT.VAL.SUBJ.FIRST.LOCAL.CAT.HEAD +np ]`

Japanese [jap] - Walkthrough of Lexical Types



Japanese [jap] - Walkthrough of Lexical Types

$$\begin{bmatrix} \textit{bleached-verb-lv-lex} \\ \text{SYNSEM} \mid \text{LOCAL} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS} \quad \left\langle \boxed{1}, \ \ldots \right\rangle \\ \text{ARG-ST} \quad \left\langle [\ ], \ \boxed{1} \begin{bmatrix} \text{LOCAL} \mid \text{CONT} \mid \text{HOOK} \mid \text{INDEX} \quad \textit{event} \end{bmatrix}, \ \ldots \right\rangle \end{bmatrix}$$

- Japanese doesn't allow coverbs to take on dependents (similar to their non-coverb counterparts) so this must also be added:

$$\begin{bmatrix} \text{SYNSEM} \quad \begin{bmatrix} \text{LOCAL} \mid \text{CAT} \mid \text{VAL} \mid \text{COMPS} \quad \left\langle \boxed{1}, \ \ldots \right\rangle \\ \text{ARG-ST} \quad \left\langle [\ ], \ \boxed{1} \begin{bmatrix} \text{LIGHT} & + \\ \text{MODIFIED} & \textit{notmod} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Questionnaire

LVC Subpage

Light Verb Constructions [documentation]

If your language uses light verb constructions (LVCs), mark the appropriate options below.

What possible coverbs are allowed in your language?
(selecting one or more of these options will allow you to add coverbs of that type to the lexicon)

- [x] Noun Coverbs
- [x] Verb Coverbs

Can noun coverbs pick up dependents on their own (e.g. determiner, modifier, complement)?

- (•) yes
- ( ) no

Can verb coverbs pick up dependents on their own (e.g. modifier, complement)?

- ( ) yes
- (•) no

The word order within an LVC is:

- (•) The coverb is before the light verb
- ( ) The coverb is after the light verb
- ( ) The coverb can be before or after the light verb

Does a coverb (or the constituent it heads) have to be immediately adjacent to a light verb?

- ( ) yes
- (•) no

What possible valence options are allowed for light verbs in your language?

- [x] A light verb can take just a coverb as an argument (i.e. intransitive)
- [x] A light verb can take a coverb plus an additional complement as arguments (i.e. transitive)

Are bleached (semantically empty) verbs possible in your language?

- (•) yes
- ( ) no

Are all light verbs bleached in your language?

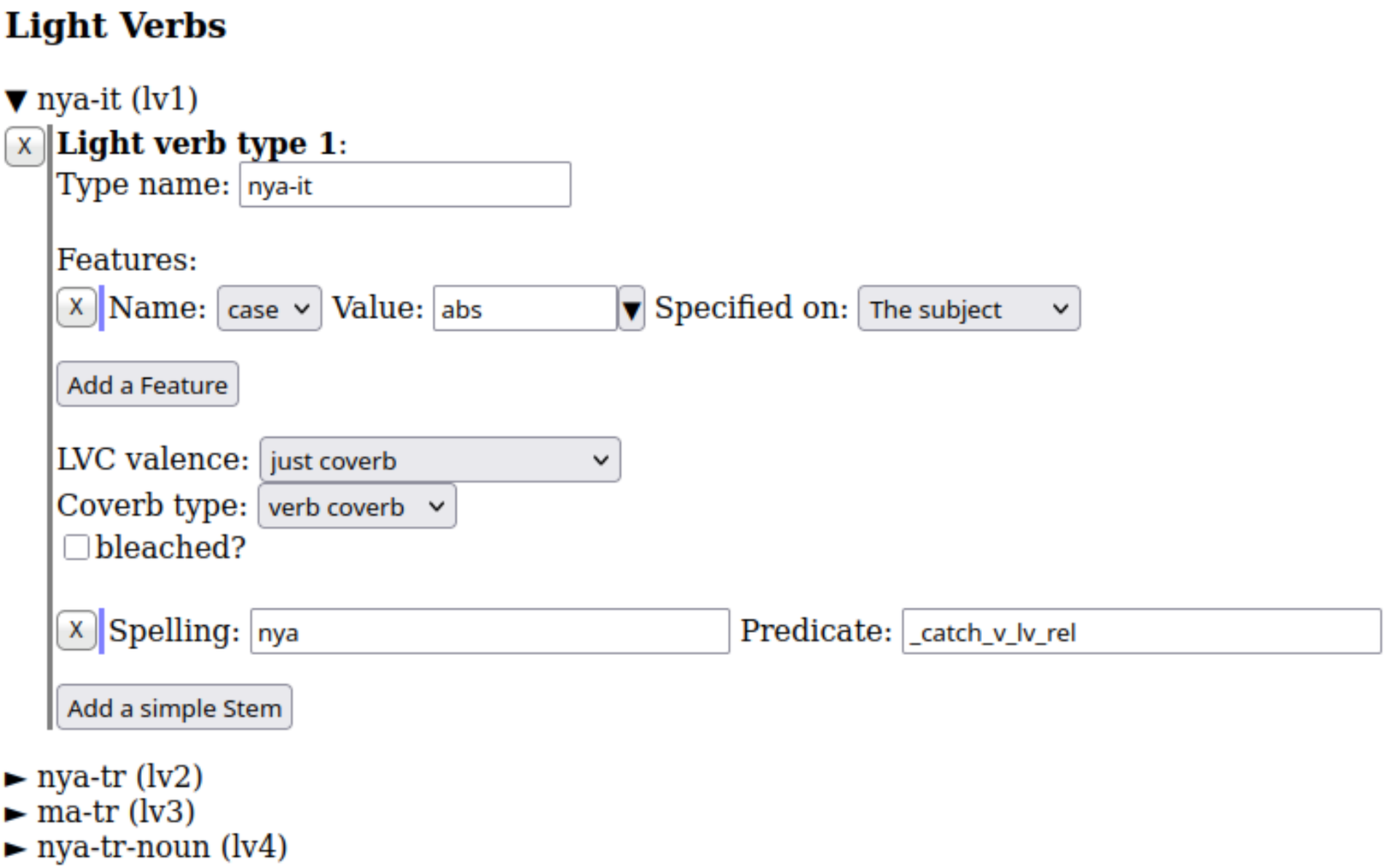
- ( ) yes
- (•) no

Main page

*Gen Info
*Word Order
Number
*Person
Gender
*Case
Poss
Dir-inv
TAM
Evidentials
Features
Neg
Coord
Y/N Qs
Wh-Qs
Info Str
Arg Opt
Nmz
Light Verb Constructions
Embed Claus
Clausal Mod
?Lexicon
Morph
Toolbox Import
Test S
TbG Options

Choices file (right-click to download)
Save & stay
Clear current subpage
Create grammar: # tgz, zip

Lexicon Subpage



Lexicon Subpage

▼ noun10

X **Noun type 10**:

Type name:

Supertypes: ▼

☐ This is a personal pronoun type

☐ This is a question pronoun (like *who/what*)

Features:

Add a Feature

For nouns of this type, a determiner is: ○ obligatory ○ optional ◉ impossible

Stems:

X Spelling: liyan Predicate: _heart_n_rel

Add a Stem

Morphotactic Constraints:

Add a Require constraint

Add a Forbid constraint

Light Verb Constructions:

○ This can be a coverb
◉ This is only a coverb

Light verbs that take this coverb class as an argument:

Light verb(s): nya-tr-noun (lv4) ▼

▼ verb10

X **Verb type 10**:

Type name:

Supertypes: ▼

Features:

Add a Feature

Argument structure: Intransitive

If this verb class includes bipartite stems, select the position class for the affix portion of the stems:

Stems:

X Spelling: daab Predicate: _go.up.to_v_rel

X Spelling: joornk Predicate: _run_v_rel

Add a simple Stem

Add a bipartite Stem

Morphotactic Constraints:

Add a Require constraint

Add a Forbid constraint

Light Verb Constructions:

○ This can be a coverb
◉ This is only a coverb

Light verbs that take this coverb class as an argument:

Light verb(s): lv1, lv2 ▼

Evaluation

Languages

- Illustrative Languages
  - Bardi (bcj, Nyulnuylan)
  - English (eng, Indo-European)
  - Japanese (jpn, Japonic)
  - Persian (per, Indo-European)
- Pseudo Languages
  - Pseudo 1
  - Pseudo 2
  - Pseudo 3
- Held-out Languages
  - Ch’ol (ctu, Mayan)
  - Daasanach (dsh, Afro-Asiatic)
  - Korafe-Yegha (kpr, Nuclear Trans New Guinean)

Results

Language	Positive	Coverage	Negative	Overgeneration	Ambiguity
Bardi [bcj]	7	100%	3	0%	1.86
English [eng]	7	100%	2	0%	2.86
Japanese [jpn]	6	100%	3	0%	1.0
Persian [per]	6	100%	2	0%	1.0
Pseudo 1 [pso1]	9	100%	5	0%	1.44
Pseudo 2 [pso2]	11	100%	2	0%	3.91
Pseudo 3 [pso3]	10	100%	2	0%	1.80
Ch'ol [ctu]	10	90%	5	0%	3.11
Daasanach [dsh]	9	77.8%	4	0%	1.0
Korafe-Yegha [kpr]	8	100%	4	75%	2.13

References

Bender, E. M., Flickinger, D., & Oepen, S. (2002). The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In COLING-02: Grammar Engineering and Evaluation.

Bender, E. M., Drellishak, S., Fokkens, A., Poulson, L., & Saleem, S. (2010). Grammar customization. Research on Language and Computation, 8, 23-72.
Bowern, C. (2012). A grammar of Bardi. de Gruyter.

Butt, M. (2010). The light verb jungle: Still hacking away. Complex predicates in cross-linguistic perspective, 48-78.

Godard, D., & Samvelian, P. (2021). Complex predicates. HeadDriven Phrase Structure Grammar, 419.

Wueger, Tara. (2024). Modeling Light Verb Constructions in the LinGO Grammar Matrix. University of Washington MS thesis.

Questions?