

THAI GRAMMAR: HPSG ALTERNATIVES

GLENN SLAYDEN

DELPH-IN SUMMIT – AUGUST 9, 2017

DISSERTATION PROSPECTUS

- DELPH-IN Joint Reference Formalism (JRF) grammar of Thai
 - Matrix-based
 - Metagrammar facilitates contrastive treatments for a given phenomena
- A principled approach to ad-hoc modularity
 - General-purpose framework for declaring, naming, manipulating and enabling/disabling fragmentary structure in TFS grammars
 - Define arbitrary new axes for expressing persistent structural enrichment within the grammar artifact

RESEARCH QUESTIONS - GRAMMAR

- *What are the syntactic and semantic implications of alternative treatments of Thai grammatical aspect?*
- *What are the syntactic entailments of modeling the predicative adjective role with a distinct lexical type, versus a pumping rule?*
- *For the Thai verbal complex, what alternative structural approaches suffice for constraining word order amongst (e.g. serialized) main verb(s), modals/auxiliaries, and variegated functional elements such aspect and negation markers?*

RESEARCH QUESTIONS - METAGRAMMAR

- *What benefits accrue from introducing a principled, uniform approach to modularity to the unification grammar regime?*
 - *How does a modular approach grammar development affect the grammar engineering task?*
 - *How does the introduction of persistent modularity transform the expressive capacity of the grammar artifact itself?*

ADJECTIVE: ATTRIBUTIVE VS. PREDICATIVE

หมา หิว เห่า
mǎ: hǐw hàw
dog hungry bark

Hungry dogs bark; A hungry dog is barking.

หมา หิว
mǎ: hǐw
dog hungry

The dog is hungry.

เขา เป็น ครู
kʰǎw pen kʰru:
he [cop] teacher

He is a teacher.

* หมา เป็น หิว
mǎ: pen hǐw
dog [cop] hungry

The dog is hungry.

นี่ คือ หมา
nî: kʰu: mǎ:
this [cop] dog

This is a dog.

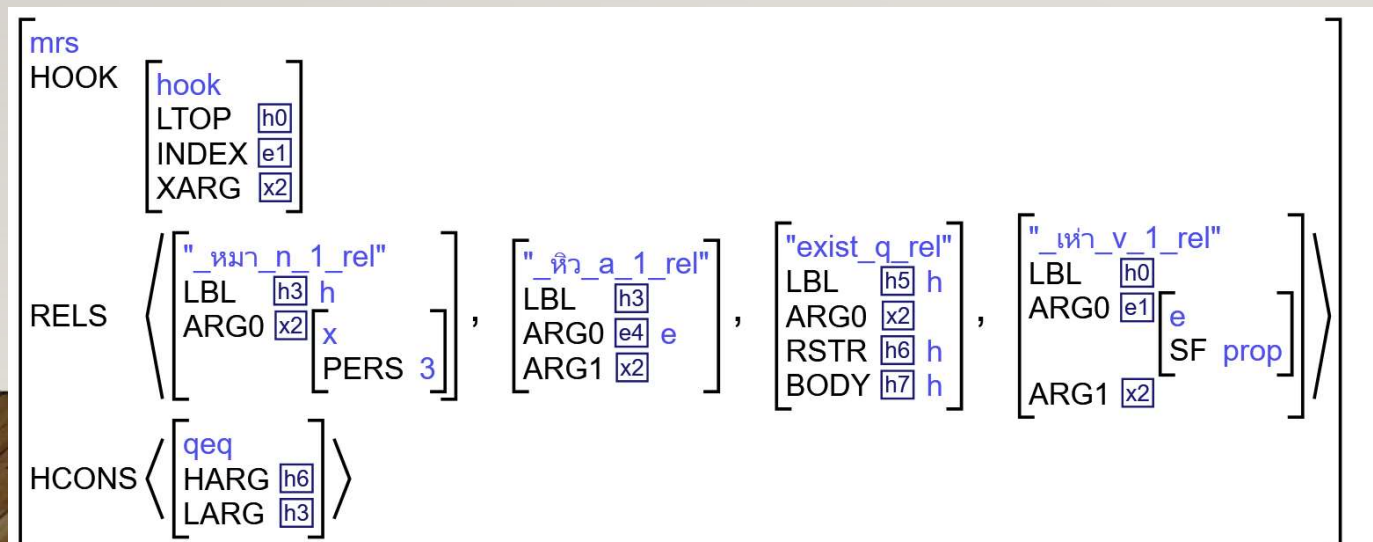
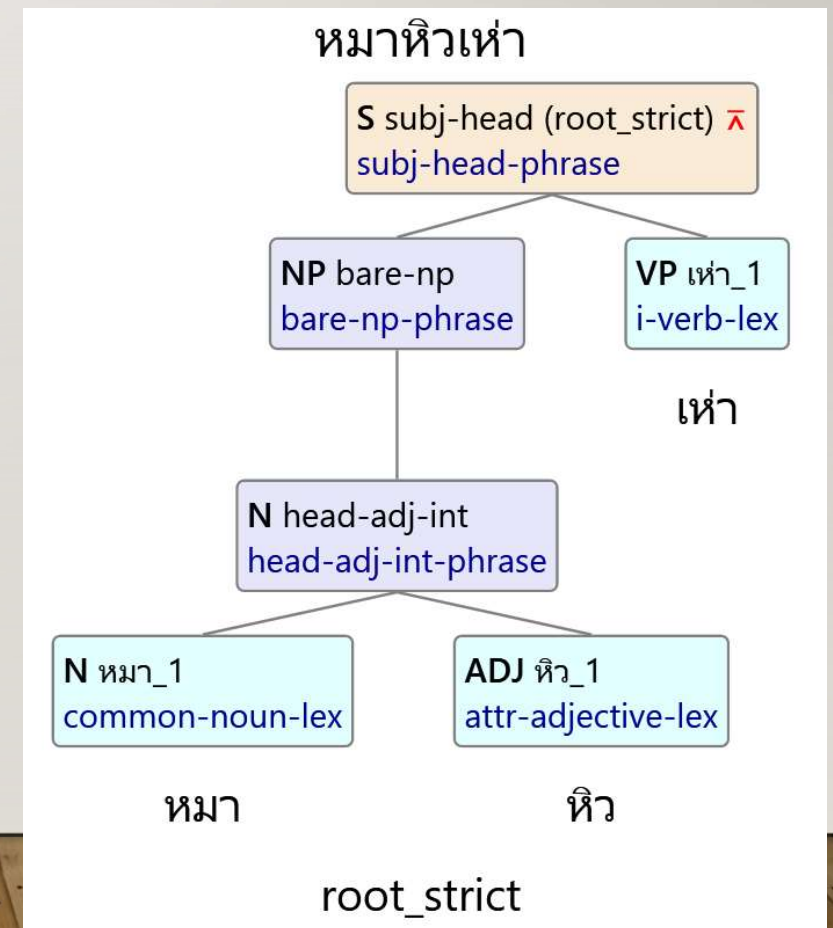
* หมา คือ หิว
mǎ: kʰu: hǐw
dog [cop] hungry

The dog is hungry.

ATTRIBUTIVE ADJECTIVE LEXICAL RULE

```

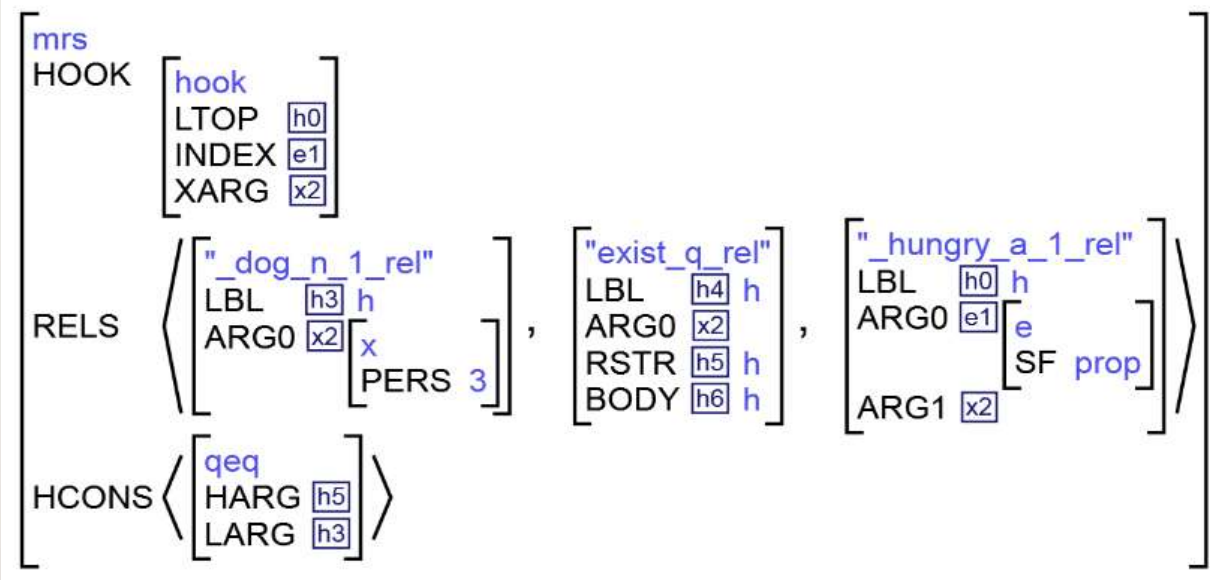
attr-adjective-lex := basic-adjective-lex & intersective-mod-lex & norm-ltop-lex-item &
  [ SYNSEM [ LOCAL [ CAT [ HEAD [ PRD -,
                                MOD < [ LOCAL.CAT [ HEAD noun,
                                                VAL.SPR cons ] ] > ],
                                VAL [ SPR < >,
                                      SUBJ < >,
                                      COMPS < >,
                                      SPEC < > ],
                                POSTHEAD + ],
    CONT.HOOK [ INDEX #ix,
                XARG #arg ] ],
  LKEYS.KEYREL [ ARG0 #ix,
                 ARG1 #arg ] ] ].
  
```



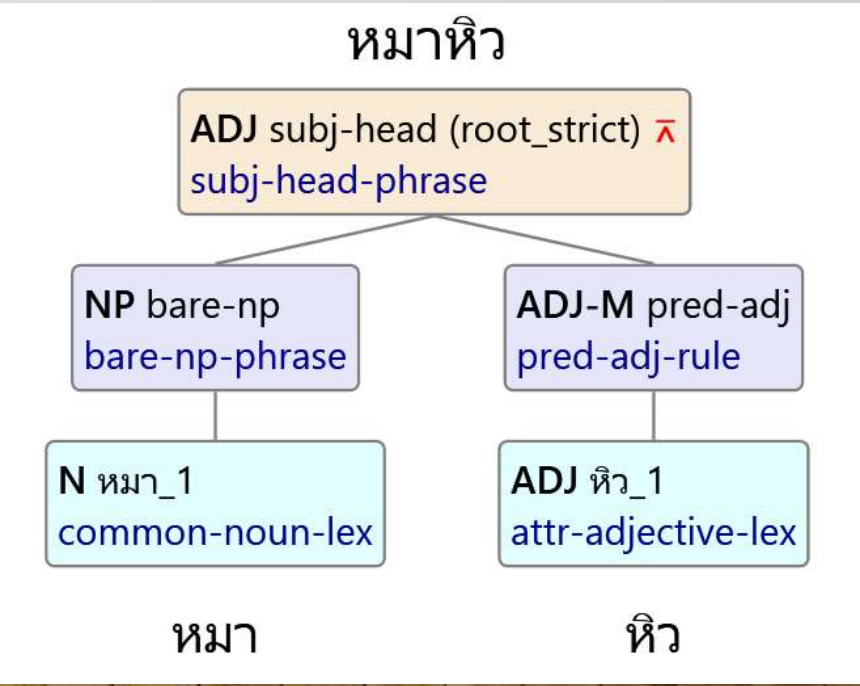
```

pred-adj-rule := unary-phrase &
  [ SYNSEM [ LOCAL [ CAT [ HEAD adj &
    [ PRD +,
      MOD < > ],
    VAL [ SPR < >,
      SUBJ < [ OPT +,
        LOCAL [ CAT [ HEAD noun,
          VAL.SPR < > ],
          CONT.HOOK.INDEX #ix ] ] >,
      COMPS < >,
      SPEC < > ],
    POSTHEAD #ph ],
    CONT.HOOK #hk ],
    NONLOC #nl,
    MODIFIED #mod,
    LIGHT #light ],
  ARGS < lex-item &
    [ SYNSEM [ LOCAL [ CAT [ HEAD adj &
      [ MOD < [ LOCAL.CONT.HOOK.INDEX #ix ] > ],
      VAL [ SPR < >,
        SUBJ < >,
        COMPS < >,
        SPEC < > ],
      POSTHEAD #ph ],
      CONT.HOOK #hk ],
      NONLOC #nl,
      MODIFIED #mod,
      LIGHT #light ] ] >,
  C-CONT [ RELS <! !>,
    HCONS <! !>,
    ICONS <! !> ] ].

```



PREDICATIVE ADJECTIVE via unary rule



THAI ASPECT

QUITE EXUBERANT

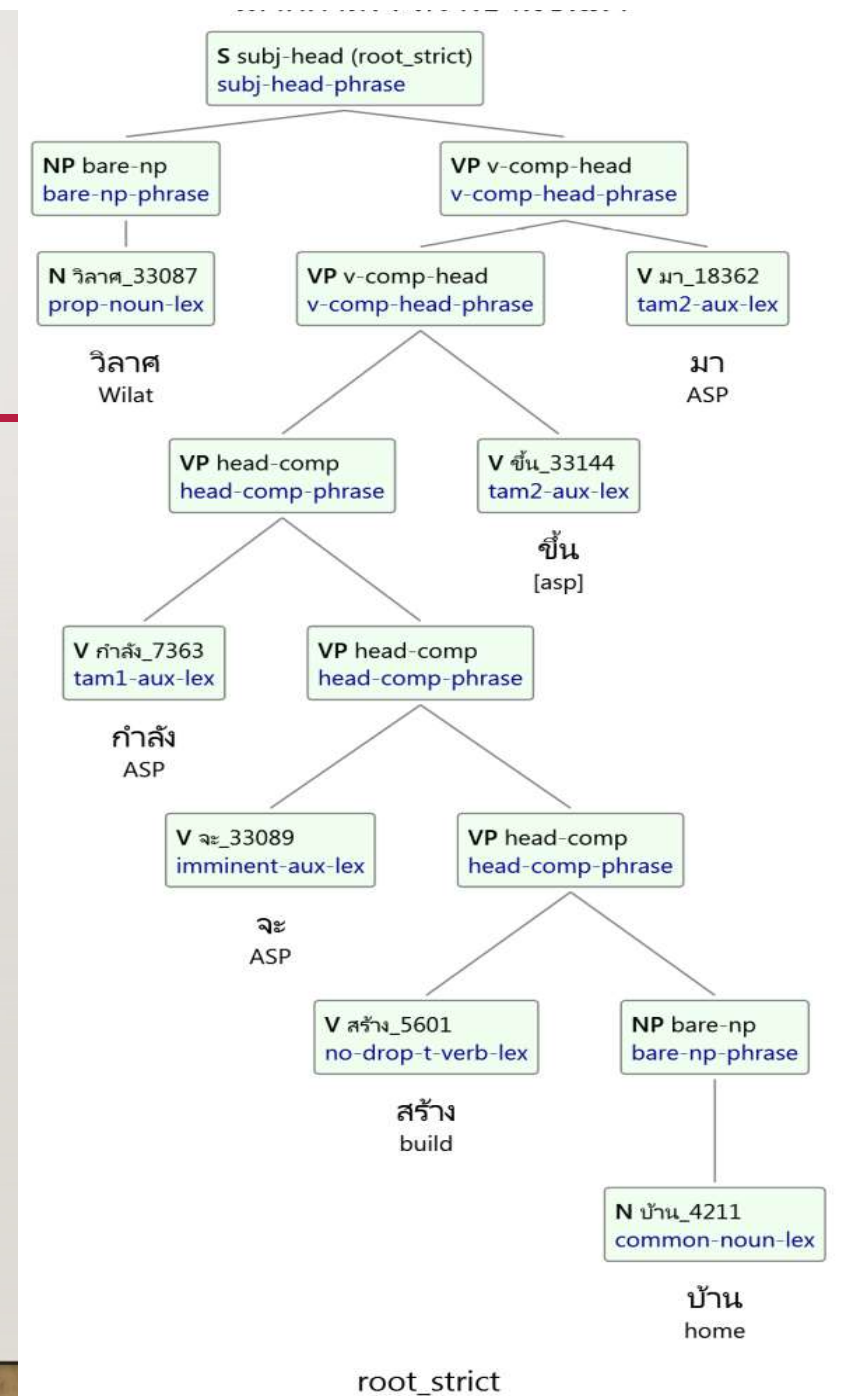
วิลาศกำลังจะเย็บ	wíʔlâ:t kamlan tɕàʔ jép	w PRG WILL sew	Wilat is going to sew.
วิลาศกำลังจะเย็บเสร็จ	wíʔlâ:t kamlan tɕàʔ jép sèt	w PRG WILL sew FIN	Wilat is about to finish sewing.
วิลาศกำลังเย็บจะเสร็จ	wíʔlâ:t kamlan jép tɕàʔ sèt	w PRG sew WILL FIN	Wilat is going to finish sewing.
วิลาศกำลังเย็บเสร็จ	wíʔlâ:t kamlan jép sèt	w PRG sew FIN	Wilat is going to finish sewing.
วิลาศจะกำลังเย็บ	wíʔlâ:t tɕàʔ kamlan jép	w WILL PRG sew	Wilat was about to be sewing.
วิลาศจะเย็บขึ้น	wíʔlâ:t tɕàʔ jép k ^h ûn	w WILL sew UP	Wilat is about to sew (to its end).
วิลาศเย็บกำลังจะเสร็จ	wíʔlâ:t jép kamlan tɕàʔ sèt	w sew PRG WILL FIN	Wilat is about to finish sewing.
วิลาศเย็บเสร็จ	wíʔlâ:t jép sèt	w sew [fin]	Wilat finished sewing.

THAI ASPECT

วิลิตกำลังจะสร้างบ้านขึ้นมา

วิลิต	กำลัง	จะ	สร้าง	บ้าน	ขึ้น	มา
wíʔlâ:t	kamlan	tɕàʔ	sâ:ŋ	bâ:n	kʰûn	ma:
Wilat	[asp]	[asp]	build	house	[asp]	[asp]

Wilat is about to build a (complete) house (from scratch).



THAI ASPECT MARKERS

Thai Aspect Markers

Group 1 TAM	Group 2 TAMs
p ^h əŋ: POST-INC	k ^h u̯m: SEMI-PERFV (lit. 'ascend')
rəəm: start, INCH	lɔŋ: SEMI-PERFV (lit. 'descend')
k ^h əj: experience	ʔòk: SEMI-PERFV (lit. 'exit')
càʔ: be about to	k ^h âw IMPFV (lit. 'enter')
kamləŋ: PROG	jùu: IMPFV (lit. 'be located')
	paj: IMPFV (lit. 'go')
	paj: PERFV (lit. 'go')
	sǎa: PERFV (lit. 'lose, waste')
	tòw: continue
	sèd: finish
	còb: end
	maa: PERF (lit. 'come')

Koenig, J.-P., & Muansuwan, N. (2005). *The Syntax of Aspect in Thai*. *Natural Language & Linguistic Theory*. 23 (2), 335-380.

ขึ้น	k ^h u̯m	ascend	[semi-perfective]
เข้า	k ^h âw	enter	[imperfective]
เคย	k ^h ɤj	ever	
จบ	tɔp	stop	
จะ	tɔʔ	"will"	
ได้	daɿ		[present perfect]
ต่อ	tò:	until	[present continuative]
ไป	paj	go	[perfective/imperfective]
มา	ma:	come	[perfect]
เริ่ม	rɤm	begin	[inceptive]
ลง	lɔŋ	descend	[semi-perfective]
ไว้	wáj	trust	
เสร็จ	sèt	finish	
เสีย	sǎa	break	[perfective]
อยู่	jù:		[progressive perfect]
ออก	ʔòk	exit	[semi-perfective]
เอา	ʔaw	take	
แล้ว	lé:w	"already"	
กำลัง	kamləŋ	"-ing"	[present progressive]
เพิ่ง	phɤj	"just"	

THREE STRATEGIES FOR THAI ASPECT

1. Binarized features: PERFV+, PROGRESS-, etc.
2. Hierarchy of aspect types
3. EP functors

THAI ASPECT – 1. BINARIZED FEATURES

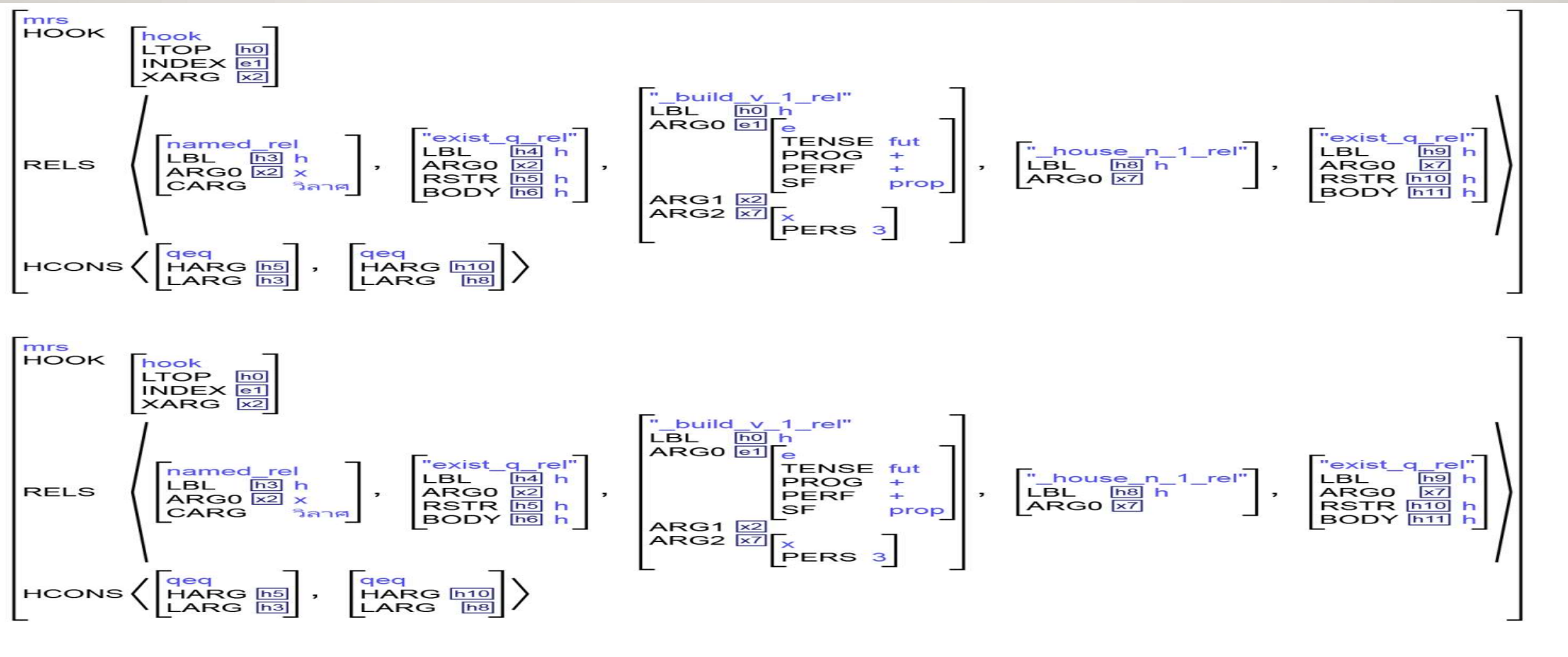
- A model of limited linguistic insight w.r.t. aspect interactions
- Currently implemented approach
- Will be extended in step to match coverage of options 2. & 3.

```
aspect := aspect-min &  
[ SOON bool,  
  PROGRESS bool,  
  GET bool,  
  EVER bool,  
  ALREADY bool,  
  CONTIN bool,  
  SEMIPERFV bool,  
  PERF bool,  
  PERFV bool ].
```


วิลาคกำลังจะสร้างบ้านขึ้นมา

วิลาศ	กำลัง	จะ	สร้าง	บ้าน	ขึ้น	มา
wí?lâ:t	kamlan	tɔ̀a?	sâ:n	bâ:n	kʰûn	ma:
Wilat	[asp]	[asp]	build	house	[asp]	[asp]

Wilat is about to build a (complete) house (from scratch).



2. HIERARCHY OF ASPECT TYPES

- Are there constraining generalizations about aspect that can be encoded?
- Which aspect markers are semantically incompatible and to what degree of strength is any corresponding mutual exclusion guaranteed?
- That is, is it the purview of the grammar to strongly assert that an event cannot be both completed [perfect] and underway [progressive]?
 - Hard to demonstrate on a single event in English: “It has stopped and is raining.”
 - But what about languages that can more directly mark contradictory properties on the same event unambiguously?

3. ASPECT FUNCTORS (EPs)

- Aspect markers contribute EP to MRS
- Allows semantic composition in the usual way
- Query/extract a Corpus of Thai Aspect Usage
- Or perhaps more generally, a Corpus of Thai VP Variation
- Currently searching corpora for minimal pairs of aspect interacting with negation, modals, and verb serialization.

ALSO TREATED

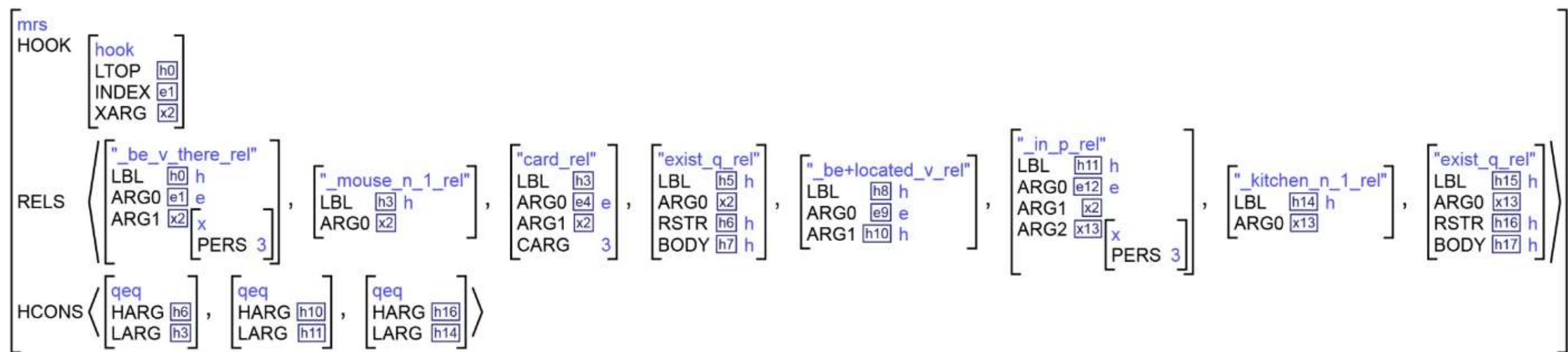
- Negation, Copula, Negative copula
- Modal auxiliaries
- Explicit coordination (N, NP, VP, S, Adj, Adv)
- Implicit VP coordination (main verb serialization)
- Case-marking adposition
- Demonstrative/numerical classifier interaction (previous DELPH-IN talk)

NEW ADDITION : PRESENTATIONAL

มีหนูสามตัวอยู่ในครัว

มี	หนู	สาม	ตัว	อยู่	ใน	ครัว
mi:	nǎ:	sǎ:m	tu:a	jù:	naj	k ^h ru:a
have	mouse	three	CLS	LOC	in	kitchen

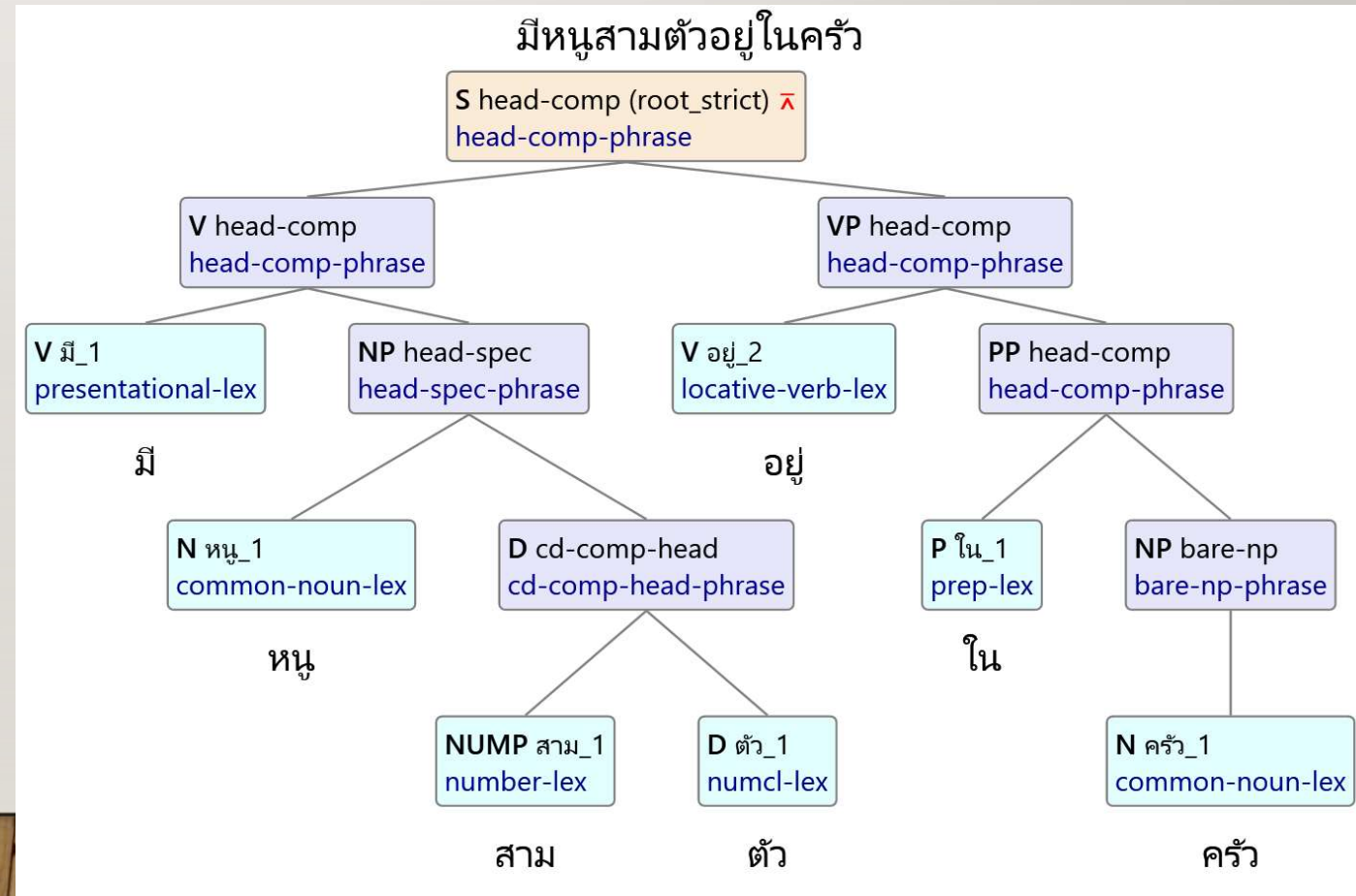
There are three mice in the kitchen.



```

presentational-lex := main-verb-base &
  [ SYNSEM.LOCAL [ CAT [ VAL [ SUBJ < >,
                           COMPS < #c1 &
                               [ OPT - ],
                               #c2 > ] ],
    CONT [ HOOK [ LTOP #h0,
                  INDEX #x2,
                  XARG #x1 ],
          RELS <! [ PRED "_be_v_there_rel",
                    LBL #h0,
                    ARG0 #x2,
                    ARG1 #x1 ] !>,
          HCONS <! !> ] ],
  ARG-ST < #c1 &
    [ LOCAL [ CAT.HEAD noun,
              CONT.HOOK.INDEX #x1 ] ],
    #c2 &
    [ LOCAL [ CAT [ HEAD verb,
                    VAL [ SPR < >,
                          COMPS < >,
                          SUBJ < [ ] > ] ],
              CONT.HOOK.XARG #x1 ] ] > ].

```



CONTINUING WORK

- Query 12K sentence corpus for VP complexes, categorizing cases of aspect, auxiliary, negation, and main verb positional constraints.
- Develop treebanking workflow – because “if you’re not treebanking, you’re not doing grammar engineering...”

SELECTED REFERENCES (1/3)

- Bender, E. M. (2008). *Grammar Engineering for Linguistic Hypothesis Testing*. (N. Gaylord, A. Palmer, & E. Ponvert, Eds.) CSLI Publications ONLINE.
- Bender, E. M., & Flickinger, D. (2005). Rapid Prototyping of Scalable Grammars: Towards Modularity in Extensions to a Language-Independent Core. *Proceedings of IJCNLP-05 (Posters/Demos)*. Jeju Island, Korea.
- Bender, E. M., Drellishak, S., Fokkens, A. S., Poulson, L., & Saleem, S. (2010). Grammar Customization. *Research on Language and Computation* 8(1), 23–72.
- Bender, E. M., Flickinger, D., & Oepen, S. (2011). Grammar Engineering and Linguistic Hypothesis Testing: Computational Support for Complexity in Syntactic Analysis. In E. Bender, & A. J.E. (Eds.), *Language from a Cognitive Perspective: Grammar, Usage and Processing* (pp. 5-29). Stanford: CSLI Publications.
- Bender, E. M., Flickinger, D., Oepen, S., Packard, W., & Copestake, A. (2015). Layers of Interpretation: On Grammar and Compositionality. *Proceedings of the 11th International Conference on Computational Semantics, London UK, April 15-17 2015* (pp. 239-249). Association for Computational Linguistics.

SELECTED REFERENCES (2/3)

- Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems. Cambridge Textbooks in Linguistics, Volume 2.*
- Costa, F. N. (2012). *Processing Temporal Information in Unstructured Documents. Ph.D. Dissertation. Universidade De Lisboa.*
- Fokkens, A. S. (2014). *Enhancing Empirical Research for Linguistically Motivated Precision Grammars. Ph.D. Thesis, Universitat des Saarlandes.*
- Koenig, J.-P., & Muansuwan, N. (2000). How to end without ever finishing: Thai semi-perfectivity. *Journal of Semantics*, 147-184.
- Koenig, J.-P., & Muansuwan, N. (2005). The Syntax of Aspect in Thai. *Natural Language & Linguistic Theory*. 23 (2), 335-380.
- Muansuwan, N. (2002). *Verb complexes in Thai.* State University of New York at Buffalo.

SELECTED REFERENCES (3/3)

- Poulson, L. (2009). Meta-modeling of Tense and Aspect in a Cross-linguistic Grammar Engineering Platform. *UWWPL, University of Washington Working Papers in Linguistics*.
- Prasithrathsint, A. (2000). Adjectives as Verbs in Thai. *Linguistic Typology* 4, 251-271.
- Sookgasem, P. (1990). *Morphology, syntax, and semantics of auxiliaries in Thai*. University of Arizona: Ph.D. Dissertation.
- Sygal, Y., & Wintner, S. (2011). Towards Modular Development of Typed Unification Grammars. *Association for Computational Linguistics*.
- Thepkanjana, K. (1986). Serial Verb Constructions in Thai. *Ph.D. Dissertation*. University of Michigan.