Automatic Derivation of Semantic Representations for Thai Serial Verb Constructions: A Grammar-Based Approach

Vipasha Bansal
University of Washington
vipashab@uw.edu

Features of SVCs

- A clause containing two or more verbs with no overt marker of coordination, subordination, or other type of syntactic dependency[4,5]
- Monoclausal[4,6]
- Verbal Independence[4,6]
- Verbs have the same tense, aspect, and polarity value[4,6]
- Encode a single event, subevents of a larger event, or two closely related events[4,6]
- Compositional[6,7]

1. สุรี หา ของขวัญ พบ
Suri hǎː khɔ̌ːŋ-khwăn phóp
Suri seek present find
‘Suri sought then found the present’[8]

4. Aikhenvald, 2006 5. Inman, 2019 6. Haspelmath, 2016 7. Pongsutthi et al. 2013 8. Muansuwan, 2002

Distinguishing SVCs from other constructions…

2.

สุรี	กำลัง	กิน	ข้าว
Suri	kamlaŋ	kin	khâ:w
Suri	PROG	eat	rice

'Suri is eating rice/Suri is in the process of eating rice.'

Muansuwan (2002)

← Violates verbal independence

3.

สุรี	เหลว	ไหล
Suri	leĕw	lăy
Suri	be.liquefied	flow

'Suri was being silly.' Thepkanjana (1986)

← Violates compositionality

4.

ผม	อยาก	กลับ	บ้าน
phŏm	yaàk	klap	baân
I	want	return	home

'I want to return home.' Jenks (2006)

← Not monoclausal

SVCs can also be distinguished from asyndetic coordination based on this definition and a series of additional tests.

Data and Categorization

- **5 Semantic types:** Purpose, Resultative, Directional, Simultaneous, Sequential
- **3 Argument-Sharing Configurations:**
  - Shared Subject
  - Switch Function (the object of the first verb is the subject of the second verb)
  - Shared Subject and Object
- **3 Syntactic Configurations:**
  - VP → V (N) VP
  - VP → V N V
  - VP → VP VP
- **Features of component verbs:** Motion, Direction, Deictic, Posture, Intention, Stative
- Based on these distinctions, I divide SVCs into **8 categories**

(5) Sequential

สุรี	หา	ของขวัญ	พบ
Suri	hǎ:	khɔ̌:ŋ-khwăn	phóp
Suri	seek	present	find

'Suri sought then found the present'[8]

(6) Open-Purpose

สุรี	ผัด	ข้าว	กิน
Suri	phàt	khâ:w	kin
Suri	fry	rice	eat

a) 'Suri fried rice to eat (rice)'

b) 'Suri fried rice then ate (rice)'[9]

(7) Shared-Subject Resultative

สุรี	กิน	ข้าว	อิ่ม
Suri	kin	khâ:w	ʔìm
Suri	eat	rice	be.full

'Suri ate rice therefore she was full'[8]

(8) Switch-Function Resultative

สุรี	กิน	ข้าว	หมด
Suri	kin	khâ:w	mòd
Suri	eat	rice	be.gone

'Suri ate rice therefore the rice was gone'[8]

(9) Deictic Purpose

สุรี	ไป	ซื้อ	หนังสือ
Suri	paj	sɯ́ɯ	năŋsɯ́ɯ
Suri	go	buy	book

'Suri went to buy a book'

(10) Simultaneous

รี	ยืน	เคาะ	ประตู
Suri	yɨ:n	khɔ́:ʔ	pràtu:
Suri	stand	knock	door

'Suri knocked on the door while standing'[9]

(11) Direction-Deictic

สุรี	เดิน	ไป
Suri	dən	paj
Suri	walk	go

'Suri walked away from the speaker'[9]

(12) Long Directional

สุรี	เดิน	ข้าม	สะพาน	กลับ	ไป
Suri	dən	khâam	saphaan	klàb	paj
Suri	walk	cross	bridge	return	go

'Suri walked, crossing the bridge, returning, away from the speaker'[8]

8. Muansuwan, 2002 9. Sudmuk, 2005 10. Thepkanjana, 1986

SVC Interactions

SVCs can combine with additional verbs or SVCs to build longer, more complex structures, with more than one type of semantic relationship between verbs

13.

สุรี	ขี่	ม้า	ข้าม	สะพาน	เหนื่อย
Suri	khìː	mâː	khâam	saphaan	nɨ́ːy
Suri	ride	horse	cross	bridge	be.tired

‘Suri rode a horse across the bridge therefore Suri was tired.’

* ‘Suri rode a horse across the bridge therefore the horse was tired.’

14.

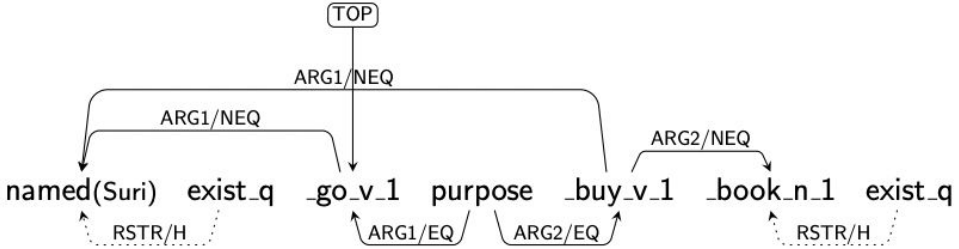
สุรี	เดิน	ข้าม	สะพาน	ไป	ซื้อ	หนังสือ	อ่าน
Suri	dəːn	khâam	saphaan	paj	sɨ́ː	nǎŋsɨ̌ː	ʔàːn
Suri	walk	cross	bridge	go	buy	book	read

‘Suri walked across the bridge to buy a book to read.’

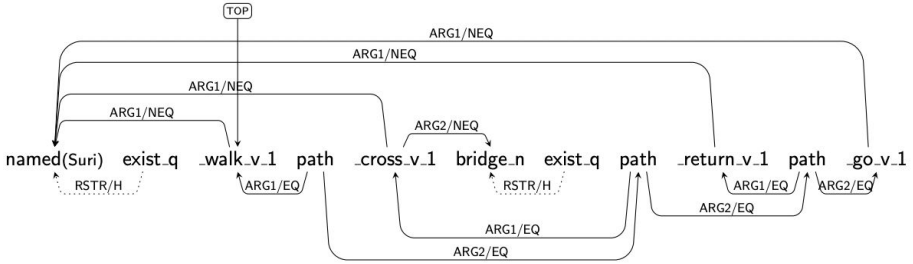
‘While walking across the bridge Suri bought a book to read.’

Target Semantic Representations

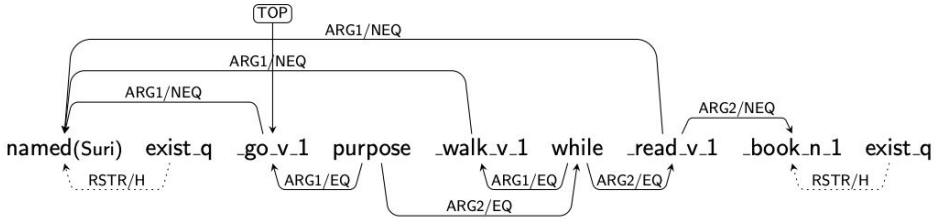
9. สุรี ไป ซื้อ หนังสือ
   Suri paj sɯ́ɯ nǎŋsɯ́ɯ
   Suri go buy book
   ‘Suri went to buy a book’



12. สุรี เดิน ข้าม สะพาน กลับ ไป
    Suri dən khâam saphaan klàb paj
    Suri walk cross bridge return go
    ‘Suri walked, crossing the bridge, returning, away from the speaker’



15. สุรี ไป เดิน อ่าน หนังสือ
    Suri paj dən ʔàːn nǎŋsɯ́ɯ
    Suri go walk read book
    ‘Suri went away from the speaker to read while walking’



Implementation Strategy

Features and Types: Added to the HEAD values of lexical entries for verbs

- **Binary Features:**
  - [SVC]
  - [AUX]
  - [NEG]
  - [INTENTION]
  - [STATIVE]
- **Additional Features:**
  - [MDDP] (motion, direction, deictic, or posture)
  - [SVTYPE] (resultative, deictic-purpose etc.)

Head Feature Principle: a mother node will have the same HEAD value as its head daughter.

Implementation Strategy

Valence-changing lexical and phrase-structure rules:

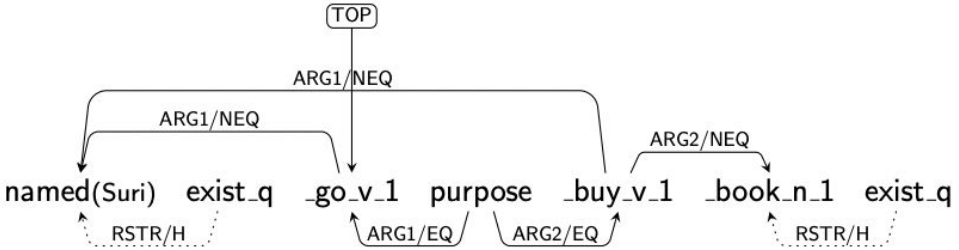
- Cross-inherit from type-hierarchies based on syntactic and semantic properties.
- Each rule subtype adds increasingly more specific constraints on component verbs

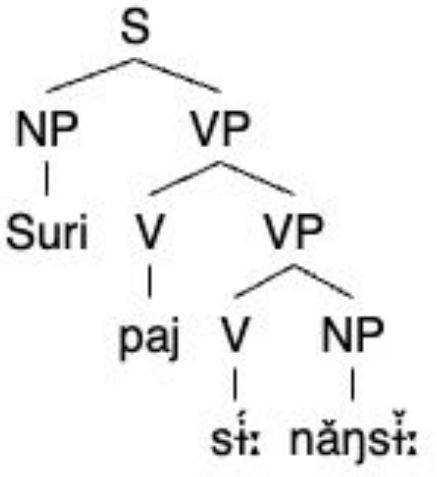
Lexical rules add a verbal complement to the initial verb's COMPS list, allowing them to combine using the existing Head Complement Rule

Phrase Structure Rules are non-headed rules and allow two distinct VPs to combine to form an overall VP

Example Analysis: Deriving the MRS for Deictic-Purpose SVC

สุรี ไป ซื้อ หนังสือ
Suri paj sɯ́ɯ nǎŋsɯ́ɯ
Suri go buy book
‘Suri went to buy a book’

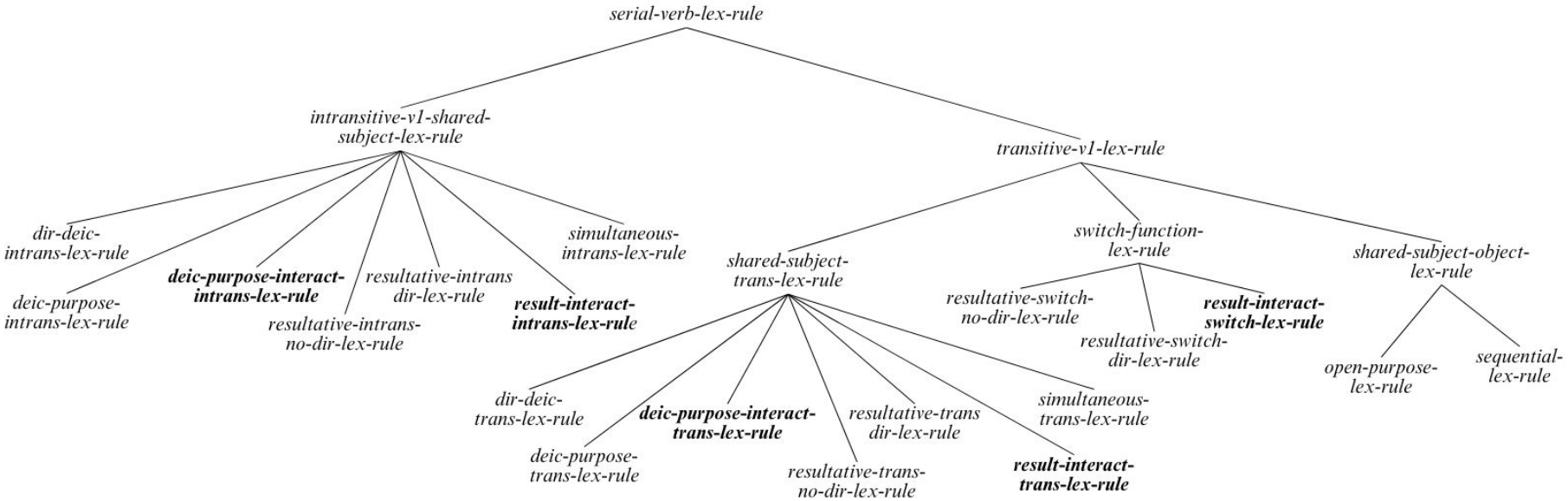




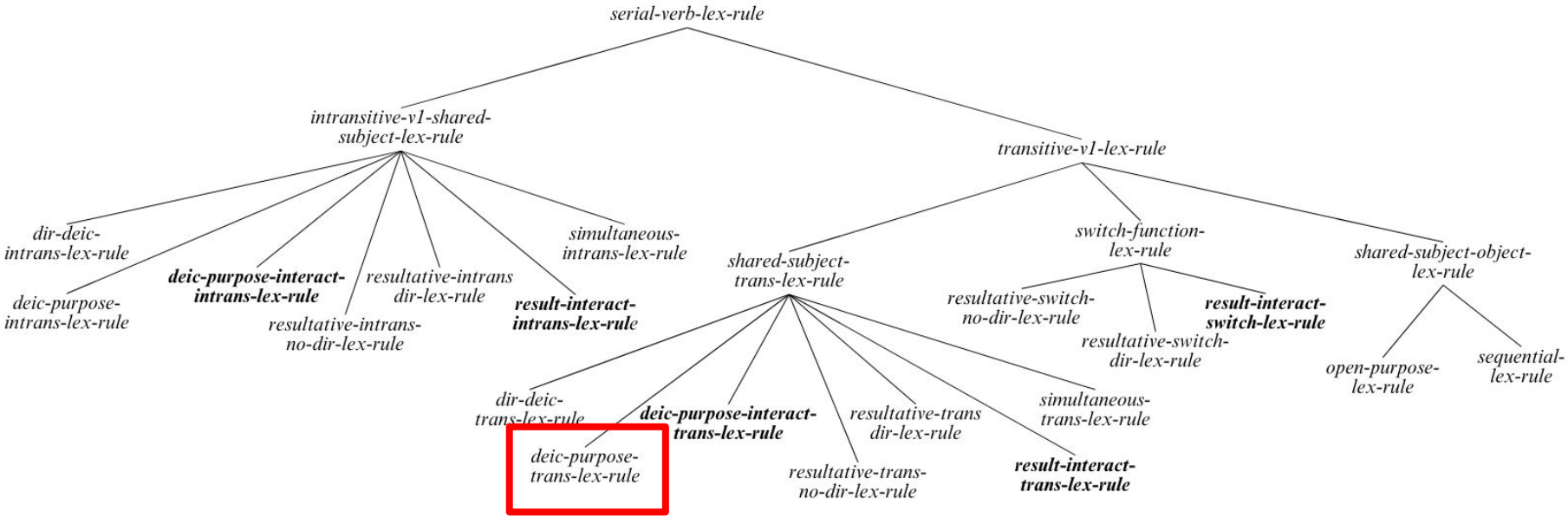
For this SVC, I use ***deictic-purpose-transitive-lex-rule***

This rule cross-inherits from
shared-subject-transitive-lex-rule and ***purpose-sem-lex-rule***

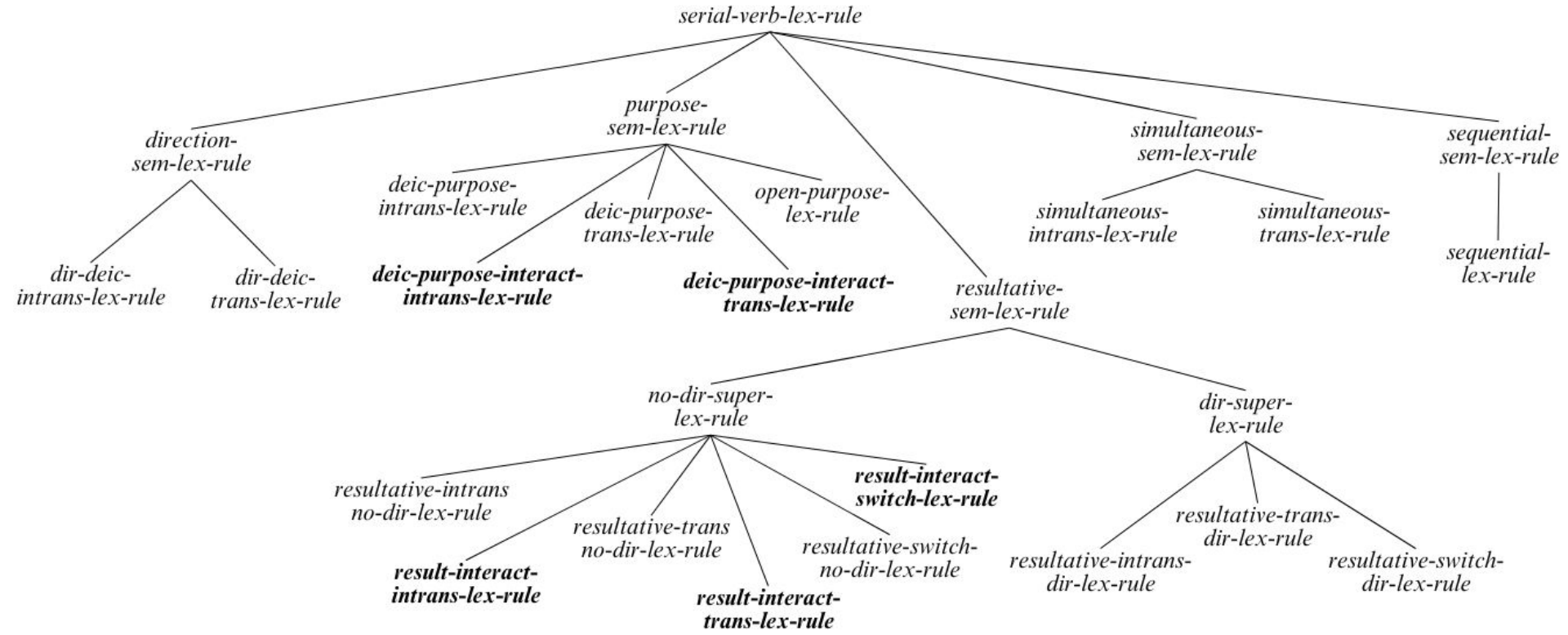
Lexical Rules: Argument-Sharing Type Hierarchy



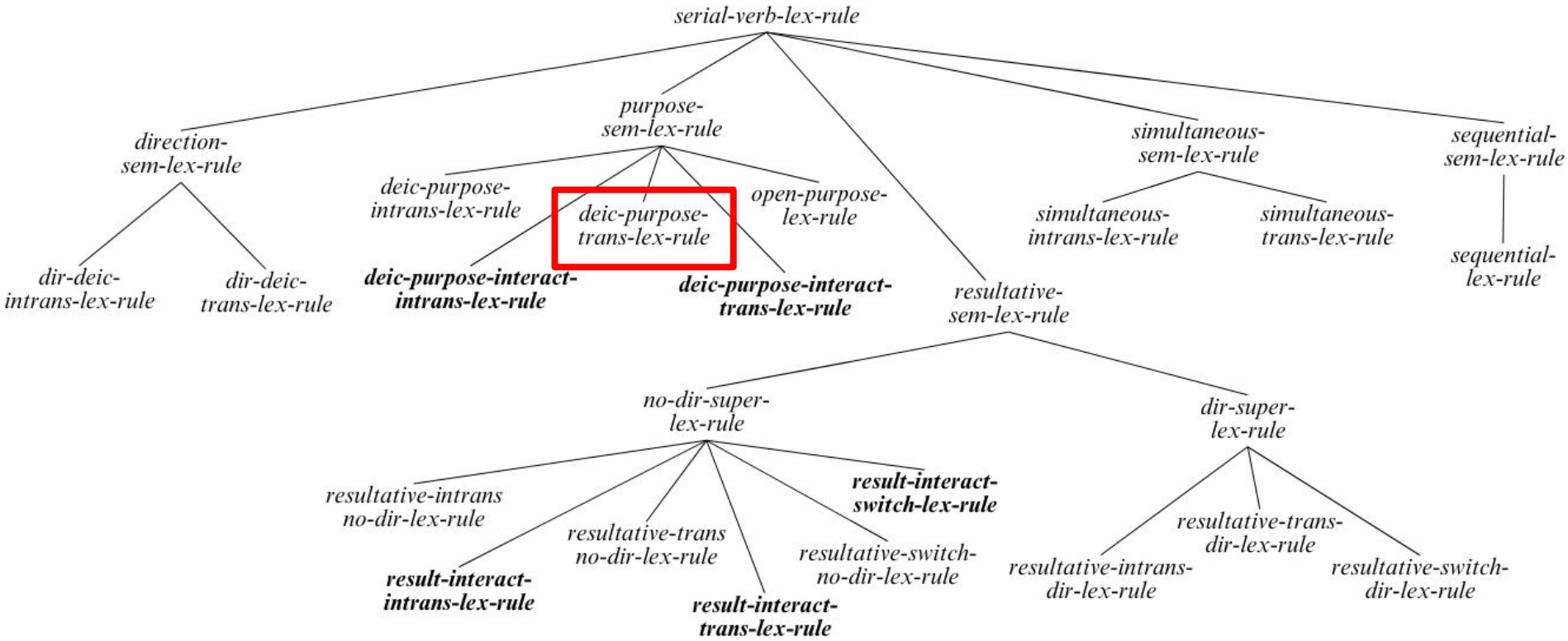
Lexical Rules: Argument-Sharing Type Hierarchy



Lexical Rules: Semantic Type Hierarchy



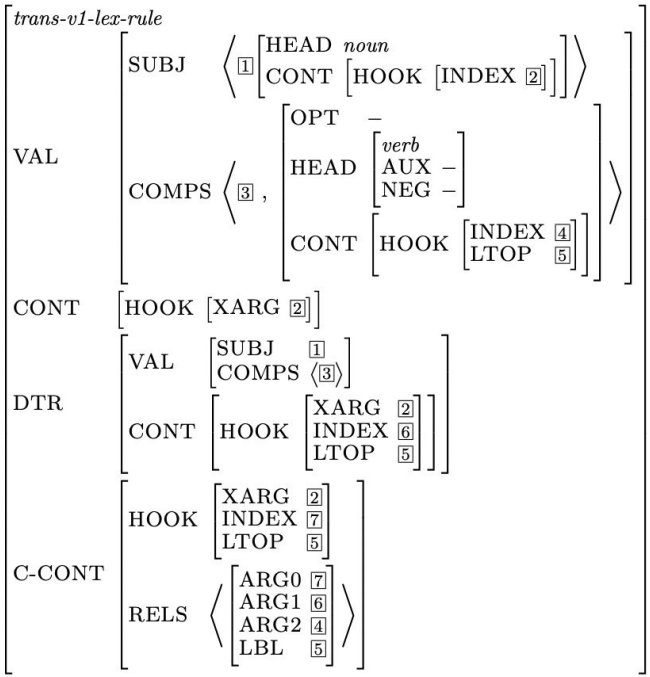
Lexical Rules: Semantic Type Hierarchy

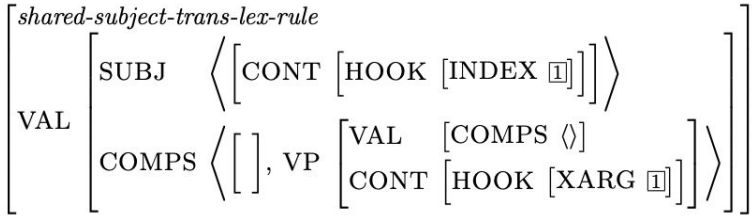


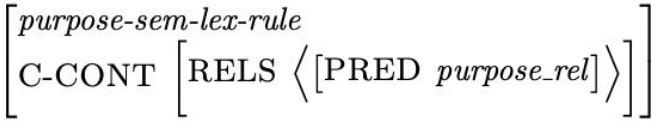
Lexical Rules: *Serial-Verb-Lex-Rule*

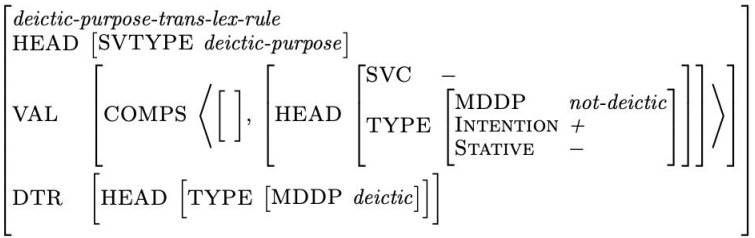
$$\begin{bmatrix} \textit{serial-verb-lex-rule} & & \\ \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{SVC} & + \\ \text{AUX} & - \\ \text{NEG} & - \end{bmatrix} & \\ \text{DTR} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{SVC} & - \\ \text{AUX} & - \\ \text{NEG} & - \end{bmatrix} \end{bmatrix} & \end{bmatrix}$$

Feature Structures to Analyze Deictic Purpose SVC

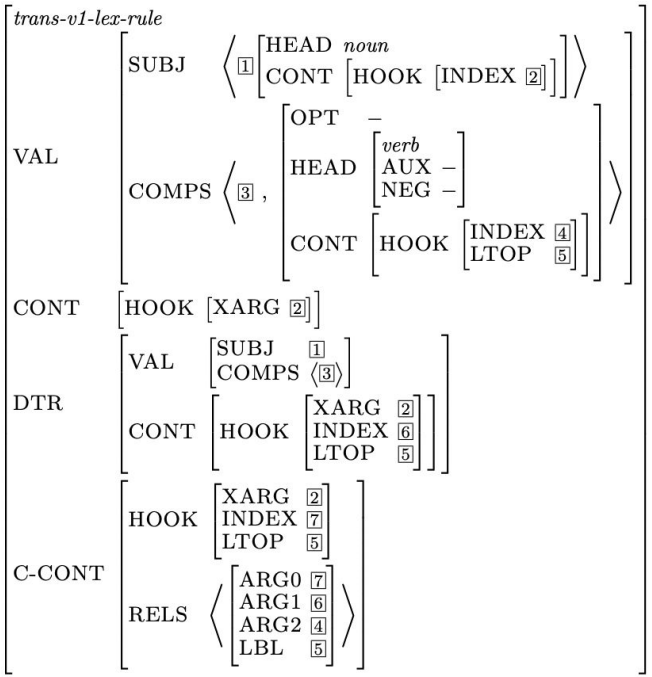




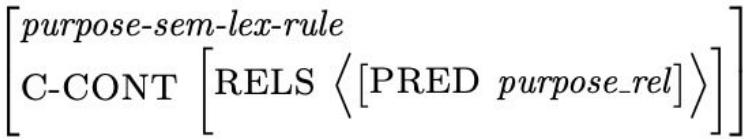




Feature Structures to Analyze Deictic Purpose SVC

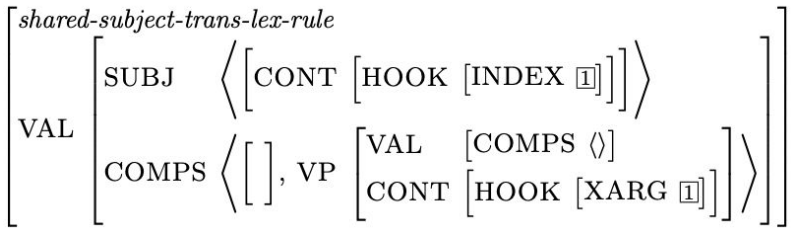


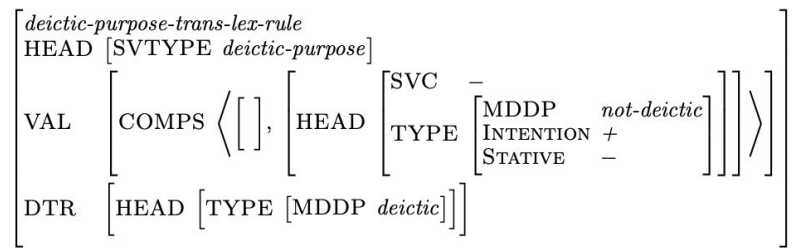
- ***trans-v1-lex-rule*** adds a verbal complement at the end of a transitive verb's COMPS list
- This added complement is [OPT –] (requiring it to be overt) and has not been negated or modified by an auxiliary verb.



- The semantic relationship between the verbs is constrained by the item on the RELS list of the C-CONT.
- This item takes each component verb as an argument ***(trans-v1-lex-rule)***, and has PRED value *purpose_rel* ***(purpose-sem-lex-rule)***.

Feature Structures to Analyze Deictic Purpose SVC

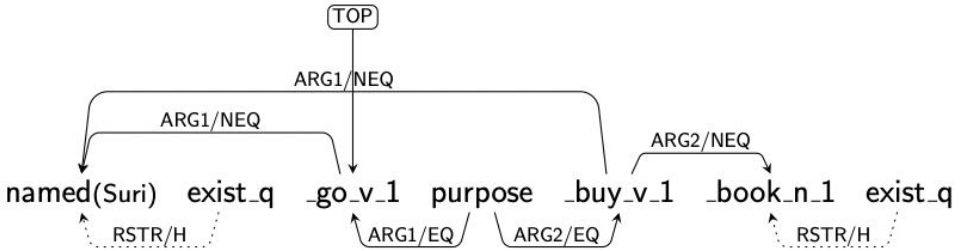




- ***shared-subject-transitive-lex-rule*** ensures that both verbs share a subject:
  - Identifies XARG of the VP complement with INDEX of the subject NP, which was already identified with XARG of the input verb in ***trans-v1-lex-rule.***
- ***deictic-purpose-transitive-lex-rule*** then constrains the individual properties of the component verbs:
  - To form a Deictic-Purpose SVC, the input verb (V1) must be deictic, while the head of the complement VP (V2) cannot be deictic, stative, or without intent.
  - The final construction is [SVTYPE *deictic-purpose*].

Example Analysis: Deriving the MRS for Deictic Purpose SVC

สุรี ไป ซื้อ หนังสือ
Suri paj sɯ́ɯ nǎŋsɯ́ɯ
Suri go buy book
‘Suri went to buy a book’



$$\left[\begin{array}{l} \textit{paj suu nangsuu} \\ \quad \left[\begin{array}{lll} \text{SVC} & + & \\ \text{AUX} & - & \\ \text{NEG} & - & \\ \text{HEAD} & \text{SVTYPE} & \textit{deictic-purpose} \\ & \text{TYPE} & \left[\begin{array}{ll} \text{MDDP} & \textit{deictic} \\ \text{INTENTION} & + \\ \text{STATIVE} & - \end{array}\right] \end{array}\right] \end{array}\right]$$

Future Dissertation Chapters

Chapter 2: How can we extend this to other languages? Both typologically related and different from Thai

- Some differences I expect:
  - Argument sharing: Subj V1 = Obj V2, No shared arguments, cumulative arguments
  - Verbal features and other factors influencing SVC semantics: in Chinese, type and position of aspect markers matters (Müller and Lipenkova 2009)
- Some typologically related languages already have SVCs implemented within the Grammar Matrix e.g. Korean, Indonesian, perhaps Vietnamese?
- Typologically different languages: many options, ideally searching for one with both an available grammar and existing SVC literature

Chapter 3: How can we use the generated MRSes to support other applications?

ขอบคุณค่ะ

Thank you!

References

Alexandra Y. Aikhenvald. 2006. Serial verb constructions in typological perspective. In Alexandra Y. Aikhenvald and R.M.W. Dixon, editors, *Serial Verb Constructions: A Crosslinguistic Typology*. Oxford University Press.

Emily Bender, Scott Drellishak, Antske Fokkens, Laurie Poulson, and Safiyyah Saleem. 2010. Grammar customization. *Research on Language and Computation*, 8:23–72.

Emily M. Bender. 2008. Evaluating a crosslinguistic grammar resource: A case study of wambaya. In *Proceedings of ACL-08: HLT*, pages 977–985, Colum- bus, Ohio. Association for Computational Linguistics.

Emily M. Bender, Dan Flickinger, and Stephan Oepen. 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*.

Emily M. Bender, Dan Flickinger, Stephan Oepen, Woodley Packard, and Ann Copestake. 2015. Layers of interpretation: On grammar and compositionality. In *Proceedings of the 11th International Conference on Computational Semantics*, pages 239–249, Lon- don, UK. Association for Computational Linguistics.

Ann Copestake. 2002. *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford, CA.

Ann Copestake. 2009. Invited Talk: slacker semantics: Why superficiality, dependency and avoidance of commitment can be the right way to go. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, pages 1–9, Athens, Greece. Association for Computational Linguistics.

Ann Copestake, Dan Flickinger, Ivan A. Sag, and Carl Pollard. 2005. Minimal recursion semantics: An introduction. *Research on Language and Computation*, 3.

A.V.N. Diller. 2006. Thai serial verbs: Cohesion and culture. In Alexandra Y. Aikhenvald and R.M.W. Dixon, editors, *Serial Verb Constructions: A Crosslinguistic Typology*. Oxford University Press.

Kira Droganova and Daniel Zeman. 2019. Towards deep Universal Dependencies. In *Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, SyntaxFest 2019)*, pages 144–152, Paris, France. Association for Computational Linguistics.

Dan Flickinger. 2000. On building a more efficient grammar by exploiting types. *Natural Language Engineering*, 6 (1):15 – 28.

References

Dan Flickinger. 2011. Accuracy vs. robustness in grammar engineering. *In E. M. Bender & J. E. Arnold (Eds.), Language from a cognitive perspective: Grammar, usage, and processing* (pp. 3150). Stanford: CSLI Publications.

Martin Haspelmath. 2016. The serial verb construction: Comparative concept and cross-linguistic generalizations. *Language and Linguistics*, 17:291–319.

David. Inman. 2019. *Multi-predicate Constructions in Nuuchahnulth*. Ph.D. thesis, The University of Washington.

Nuttanart Muansuwan. 2002. *Verb Complexes in Thai*. Ph.D. thesis, University at Buffalo, The State University of New York.

Stefan Müller, Anne Abeille, Robert Borsley, and Jean-Pierre Koenig. 2021. *Head-Driven Phrase Structure Grammar: The handbook*. Language Science Press.

Stefan Müller and Janna Lipenkova. 2009. Serial verb constructions in Chinese: An HPSG account. In *Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, University of Göttingen, Germany*, pages 234–254, Stanford. CSLI Publications

Stephan Oepen and Dan Flickinger. 1998. Towards systematic grammar profiling test suite technology 10 years after. *Journal of Computer Speech and Language*, 12:411–435.

Peter Jenks. 2006. Control in Thai. *In Variation in control structures (Online Database)*. UC San Diego.

Carl Pollard and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press.

Mancharee Pongsutthi, Nongnuch Ketui, and Chutamanee Onsuwan. 2013. Thai serial verb constructions: A corpus-based study. In *10th International Symposium on Natural Language Processing (SNLP)*.

Sebastian Schuster and Christopher D. Manning. 2016. Enhanced English Universal Dependencies: An improved representation for natural language understanding tasks. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2371–2378, Portorož, Slovenia. European Language Resources Association (ELRA).

References

Glenn Slayden. 2009. Thai resource grammar.

Cholthicha. Sudmuk. 2005. *The Syntax and Semantics of Serial Verb Constructions in Thai*. Ph.D. thesis, The University of Texas at Austin.

Kiyoko Takahashi. 2009. Basic serial verb constructions in thai. *Journal of the Southeast Asian Linguistics Society*, 1:215–229.

Kingkarn Thepkanjana. 1986. *Serial Verb Constructions in Thai*. Ph.D. thesis, University of Michigan.

Kristina Toutanova, Christopher Manning, Dan Flickinger, and Stephan Oepen. 2005. Stochastic hpsg parse disambiguation using the redwoods corpus. *Research on Language and Computation*, 3:83–105.

Olga Zamaraeva, Chris Curtis, Guy Emerson, Antske Fokkens, Michael Goodman, Kristen Howell, T.J. Trimble, and Emily M. Bender. 2022. 20 years of the grammar matrix: cross-linguistic hypothesis testing of increasingly complex interactions. *Journal of Language Modelling*, 10(1):49–137.