HPSG Analysis and Computational Implementation of Indonesian Passives

David **Moeljadi** and Francis **Bond**

Division of Linguistics and Multilingual Studies,
Nanyang Technological University, Singapore

The 25th International Conference on Head-Driven Phrase Structure Grammar
(HPSG 2018),
The University of Tokyo, Komaba Campus

1 July 2018

NANYANG TECHNOLOGICAL UNIVERSITY

Indonesian passives: Outline

1. Introduction
2. The data
3. Bound forms of pronouns
4. Analysis
5. Conclusion

NANYANG TECHNOLOGICAL UNIVERSITY

Aim of this study and its limitations

- Standard, formal Indonesian
- Passive type 1 (P1) and type 2 (P2) (Dardjowidjojo, 1978), (Sneddon et al., 2010)
- Monotransitive only (ditransitive will not be discussed)
- computationally implementable
  - Head Driven Phrase Structure Grammar (HPSG) (Pollard and Ivan A Sag, 1994; Ivan A. Sag, Wasow, and Bender, 2003)
  - Minimal Recursion Semantics (MRS) (Copestake et al., 2005)
  - Indonesian Resource Grammar (INDRA) (Moeljadi, Bond, and Song, 2015)

"To analyze and develop a detailed model (rules) of syntax and semantics of passive sentences in standard Indonesian, explicit enough to be interpreted by a computer, focusing on implementation rather than theory"

NANYANG TECHNOLOGICAL UNIVERSITY

Indonesian passives type 1 and type 2

- Two passive constructions: Dardjowidjojo, 1978 Sneddon et al., 2010 Chung, 1976 Cole, Hermon, and Tjung, 2006
  1. Passive Type 1 (P1) / Canonical passive
  2. Passive Type 2 (P2) / Object preposing passive
- P1 and P2 are in (near) complementary distribution
  - P1: 3 agents (both pronouns and nouns)
  - P2: 1 and 2 agents *or* pronouns (including 3 agent, also pronoun substitutes, including proper nouns)
  - P1 and P2 overlap with respect to 3 pronoun agent
- Syntax
  - Word order: patient/undergoer in active → passive subject
- Semantics
  - Semantically transitive clause
  - The semantic roles of the arguments in active and its passive counterpart are the same
  - The promoted arguments are *focus-or-topic*, the demoted arguments are *non-topic* (Song, 2017)

NANYANG TECHNOLOGICAL UNIVERSITY

Passive Type 1

$$\text{NP}_{\text{Theme/Patient}} + \text{Neg/Aux} + \textit{di}\text{-V} + \textit{oleh} + \text{NP}_{\text{Agent}}$$

- *di-* is prefixed to the verb, phonologically and orthographically part of the verb, always immediately to its left
- NP agent
  - third person pronoun, common noun, proper name
  - follows the verb (post-verbal)
  - demoted to adjunct, optional
  - can be preceded by a semantically empty preposition *oleh* "by"
  - with 3SG agent, the bound form *-nya* may occur
- auxiliaries and negation precede the passive *di-* verb

Passive type 1

(1) a. *Budi sudah dijemput.*
Budi PRF PASS-pick.up
"Budi has been picked up."

b. *Budi sudah dijemput (oleh) dia/mereka/istrinya/Adi.*
Budi PRF PASS-pick.up by 3SG/3PL/wife=3SG/Adi
"Budi has been picked up by him/them/his wife/Adi."

c. *Budi sudah dijemputnya.*
Budi PRF PASS-pick.up=3SG

d. *Budi sudah dijemput olehnya.*
Budi PRF PASS-pick.up by=3SG

e. * *Budi sudah dijemput aku/kamu/kami/kalian.*
Budi PRF PASS-pick.up 1SG/2SG/1PL.EXCL/2PL

Passive type 1 with coordination

(2) a. *Budi sudah dijemput (oleh) Adi dan Ani.*
Budi PRF PASS-pick.up by Adi and Ani
"Budi has been picked up by Adi and Ani."

b. * *Budi sudah dijemputnya dan Ani.*
Budi PRF PASS-pick.up=3SG and Ani

c. * *Budi sudah dijemput olehnya dan Ani.*

d. *Budi sudah dijemput dan disuntik (oleh) Adi.*
Budi PRF PASS-pick.up and PASS-inject by Adi
"Budi has been picked up and injected by Adi."

e. *Budi sudah dijemput dan disuntiknya.*

f. *Budi sudah dijemput oleh Adi dan disuntik Ani.*

Passive Type 2

$$\text{NP}_{\text{Theme/Patient}} + \text{Neg/Aux} + \text{NP}_{\text{Agent}} + \text{bare V}$$

- the verb appears in bare stem form
- NP agent
  - all pronouns, 'pronoun substitute' (kinship terms such as *bapak* "father", *ibu* "mother"), personal names (Budi, David, ...) meaning "you" or "I"
  - precedes the verb (pre-verbal)
  - not demoted to adjunct, obligatory, an argument (Arka, Manning, et al., 1998)
  - cannot be preceded by any prepositions such as *oleh* "by", nothing can intervene the NP agent and the bare verb
  - with 1SG and 2SG agents, the bound form *ku-* and *kau-* may occur
- auxiliaries and negation precede the agent (cannot come between the agent and the bare verb)

NANYANG TECHNOLOGICAL UNIVERSITY

Passive type 2

(3) a. *Budi sudah dia/mereka/aku/kamu/kami/kalian jemput.*
Budi PRF 3SG/3PL/1SG/2SG/1PL.EXCL/2PL pick.up
"Budi has been picked up by him/them/me/you/us."

b. *Budi sudah bapak jemput.*
Budi PRF father pick.up
"Budi has been picked up by me/you (father/old man)."

c. *Budi sudah Adi jemput.*
Budi PRF Adi pick.up
"Budi has been picked up by me/you (Adi)."

d. *Budi sudah kujemput.*
Budi PRF 1SG-pick.up
"Budi has been picked up by me."

e. *Budi sudah kaujemput.*
Budi PRF 2SG-pick.up

NANYANG TECHNOLOGICAL UNIVERSITY

Passive type 2 with coordination

It is impossible to coordinate two pronouns as actors or two verbs with a single actor (Musgrave 2001)

(4) a. ?*Budi sudah kami dan mereka jemput.*
Budi PRF 1PL.EXCL and 3PL pick.up
"Budi has been picked up by us and them."

b. ?*Budi sudah dia tunggu dan jemput.*
"Budi has been waited and picked up by him."

c.??*Budi sudah kutunggu dan jemput.*

d. *Budi sudah kutunggu dan kujemput.*

e. *Budi sudah kutunggu dan kaujemput.*
"Budi has been waited by me and picked up by you."

Bound forms *ku-*, *kau-*, *-nya*

Clitics vs inflectional affixes (Zwicky & Pullum 1983) (Zwicky and Pullum, 1983)

- clitics can attach to words of any category (or part-of-speech), while the inflectional affixes are quite specific in their selections of stems
- no syntactic operations exist which treat a word combined with clitics as a unit, while words with inflectional affixes are treated as units

ku- "1SG" and *kau-* "2SG" are inflectional affixes, *-nya* "3SG" is a clitic

- *-nya*
  - treated as a word (tokenized)
- *ku-* and *kau-*
  - treated as prefixes (not tokenized)
  - the relationship of the agent and the verb is morphological rather than syntactic

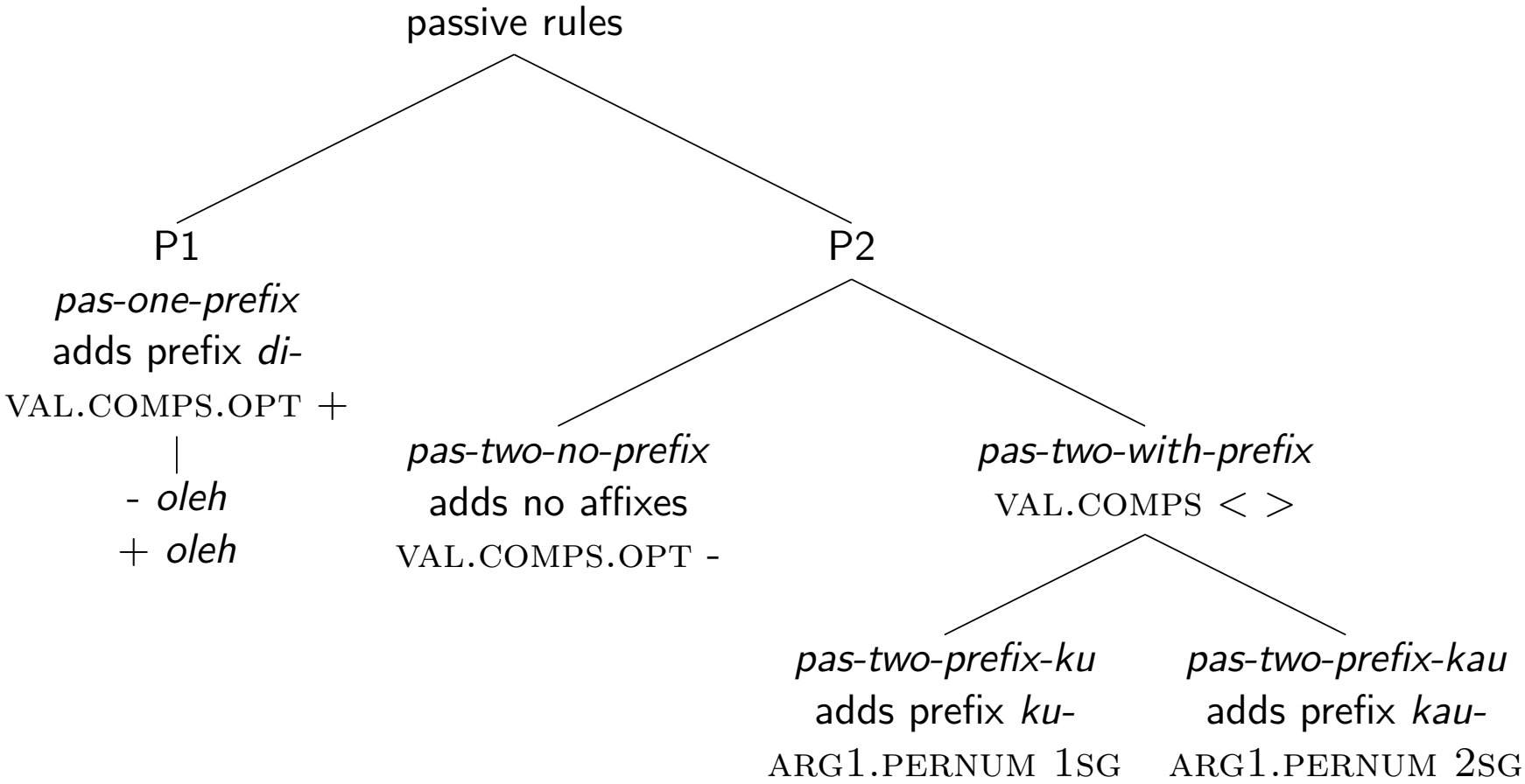
NP coordination rule: cannot take bound forms *-nya*, *ku-*, *kau-*

Passive as an inflectional rule

(5)

$$\left[\begin{array}{ll} \textsc{input} & \left\langle \text{X} \;,\; \left[\begin{array}{l} \text{lexeme } (\textit{tr-verb-lex}) \\ \textsc{arg-st} \left\langle [\ldots], [\ldots] \right\rangle \end{array}\right] \right\rangle \\ \textsc{output} & \left\langle \text{(di-/ku-/kau-)X} \;,\; \left[\begin{array}{l} \text{word } (\textit{passive-transitive-lex-item}) \\ \textsc{arg-st} \left\langle \left[\begin{array}{ll} \textsc{index} & \boxed{1} \\ \textsc{icons-key} & \boxed{3} \end{array}\right], \left[\begin{array}{ll} \textsc{index} & \boxed{2} \\ \textsc{icons-key} & \boxed{4} \end{array}\right] \right\rangle \\ \textsc{lkeys.keyrel} \left[\begin{array}{ll} \textsc{arg}1 & \boxed{2} \\ \textsc{arg}2 & \boxed{1} \end{array}\right] \\ \textsc{icons} \left\langle ! \; \boxed{3} \left[\begin{array}{ll} \textit{focus-or-topic} \\ \textsc{iarg}2 & \boxed{1} \end{array}\right], \boxed{4} \left[\begin{array}{ll} \textit{non-topic} \\ \textsc{iarg}2 & \boxed{2} \end{array}\right] ! \right\rangle \end{array}\right] \right\rangle \end{array}\right]$$

Indonesian passive lexical rules

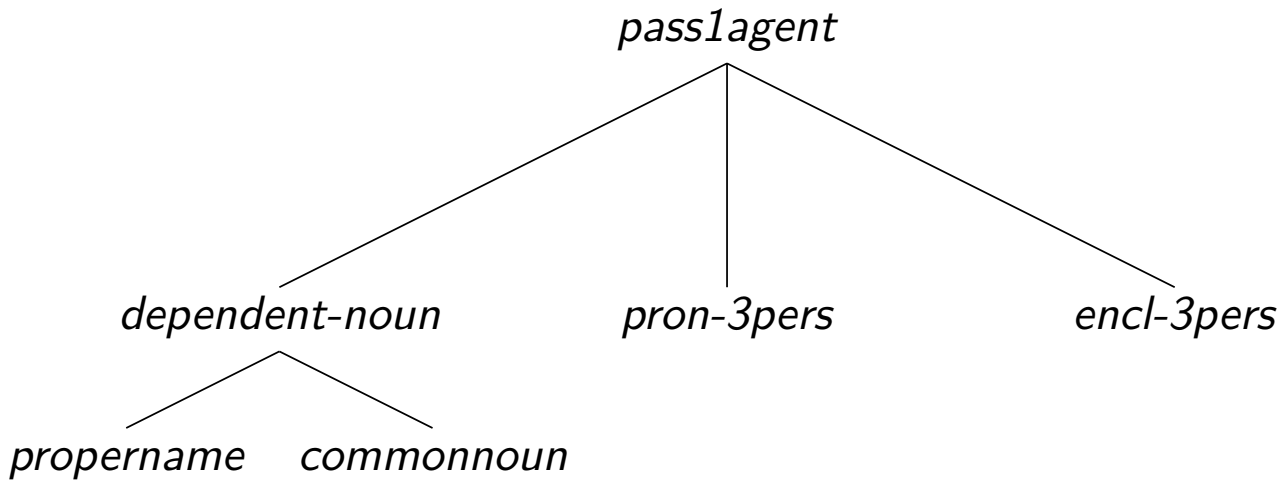


(1) P1 without *oleh*

(6)

$$\begin{bmatrix} \text{HEAD } \textit{pass1} \\ \text{VAL.COMPS } \left\langle \boxed{1} \right\rangle \\ \text{ARG-ST } \left\langle \begin{bmatrix} \text{HEAD } \textit{subj-noun} \\ \text{VAL } \begin{bmatrix} \text{SPR } \langle\,\rangle \\ \text{COMPS } \langle\,\rangle \end{bmatrix} \end{bmatrix}, \boxed{1} \begin{bmatrix} \text{HEAD } \textit{pass1agent} \\ \text{POSTHEAD } + \\ \text{VAL } \begin{bmatrix} \text{SPR } \langle\,\rangle \\ \text{COMPS } \langle\,\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

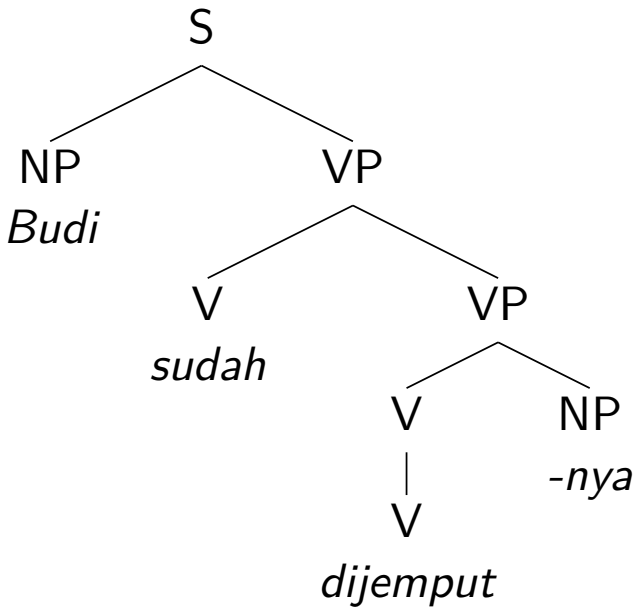
(7)

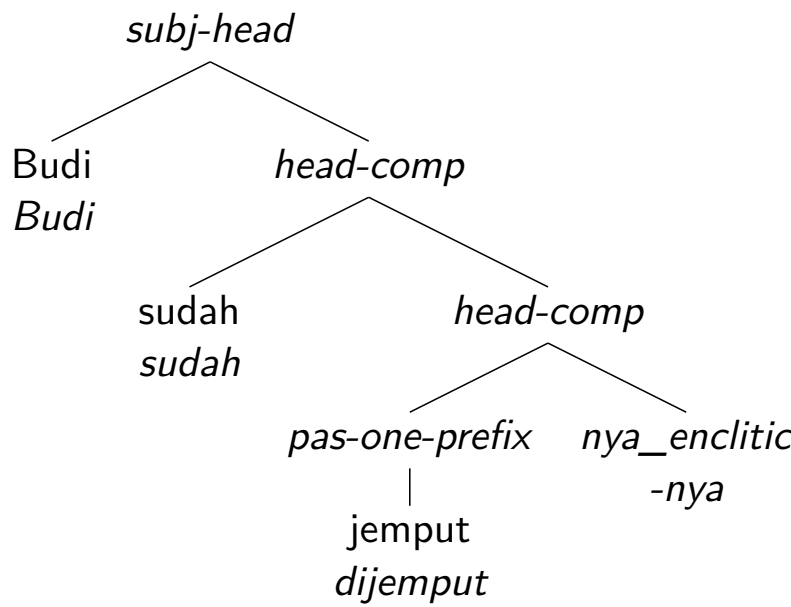


NANYANG TECHNOLOGICAL UNIVERSITY

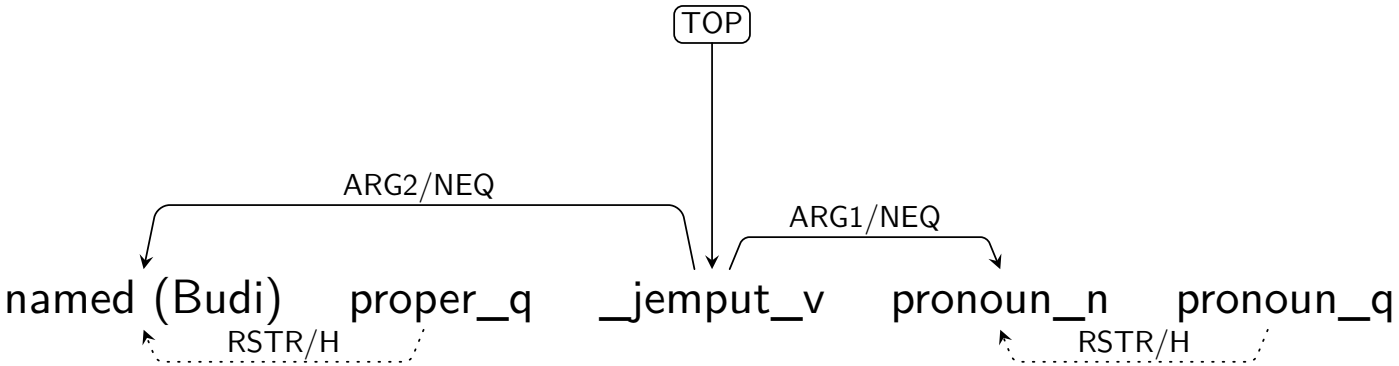
(1) P1 without *oleh* (tree and DMRS)

(8)





(9)



(2) P1 with *oleh* "by"

(10)
$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{oleh-adp} & \\ \text{MOD.LOC.CAT} & \begin{bmatrix} \text{HEAD} & \textit{pass1} \\ \text{VAL.COMPS} & \langle \boxed{1} \rangle \end{bmatrix} \end{bmatrix} \\ \text{VAL.COMPS} & \langle \boxed{1} \rangle \end{bmatrix}$$

- *oleh* adds nothing to the meaning except the information that the COMPS of the passive verb is coindexed with the COMPS of *oleh*
- The semantics of the PP headed by *oleh* is identical to the semantics of *oleh*'s NP complement

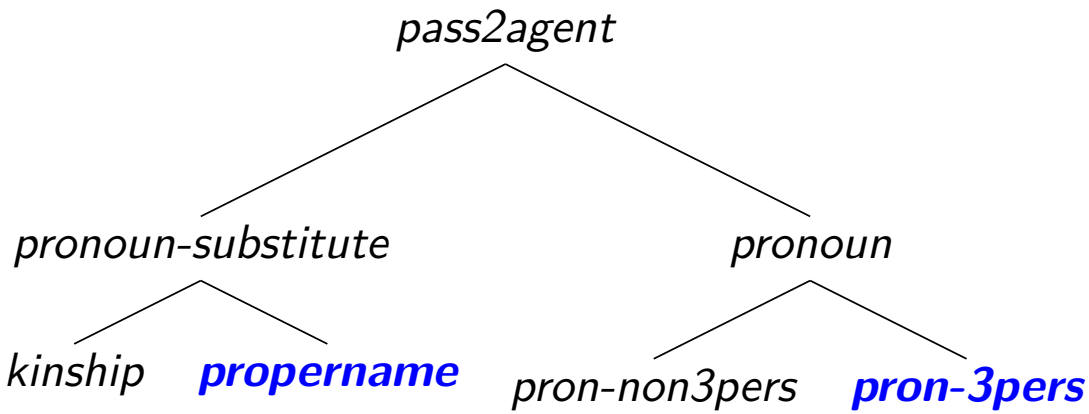
NANYANG TECHNOLOGICAL UNIVERSITY

(3) P2 without bound forms

(11)

$$\begin{bmatrix} \text{HEAD } \textit{passive-two} \\ \text{VAL.COMPS } \left\langle \boxed{1} \right\rangle \\ \text{ARG-ST } \left\langle \begin{bmatrix} \text{HEAD } \textit{subj-noun} \\ \text{VAL } \begin{bmatrix} \text{SPR } \langle\,\rangle \\ \text{COMPS } \langle\,\rangle \end{bmatrix} \end{bmatrix}, \boxed{1} \begin{bmatrix} \text{HEAD } \textit{pass2agent} \\ \text{POSTHEAD } - \\ \text{VAL } \begin{bmatrix} \text{SPR } \langle\,\rangle \\ \text{COMPS } \langle\,\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

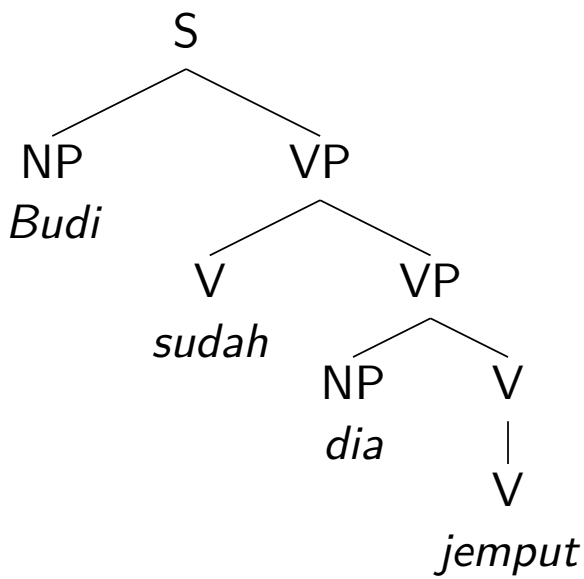
(12)

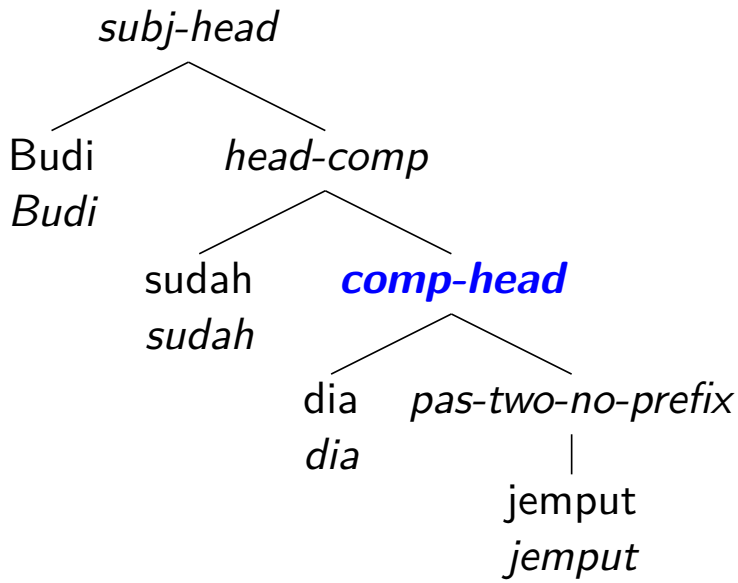


NANYANG TECHNOLOGICAL UNIVERSITY

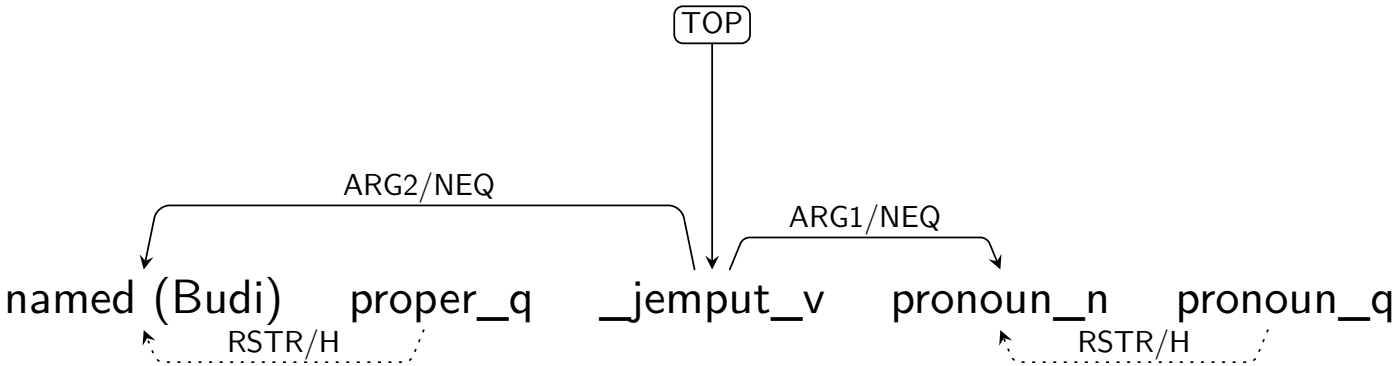
(3) P2 without bound forms (tree and DMRS)

(13)



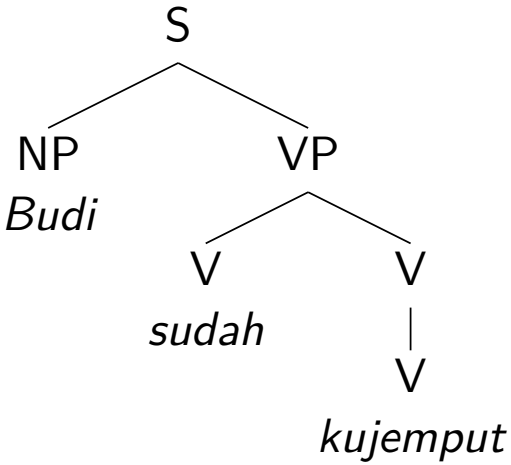


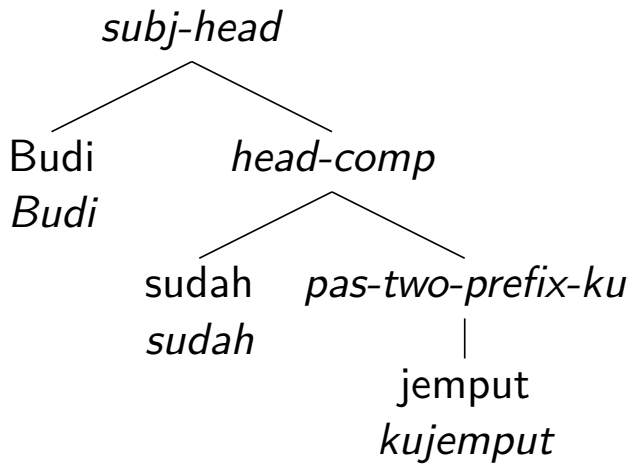
(14)



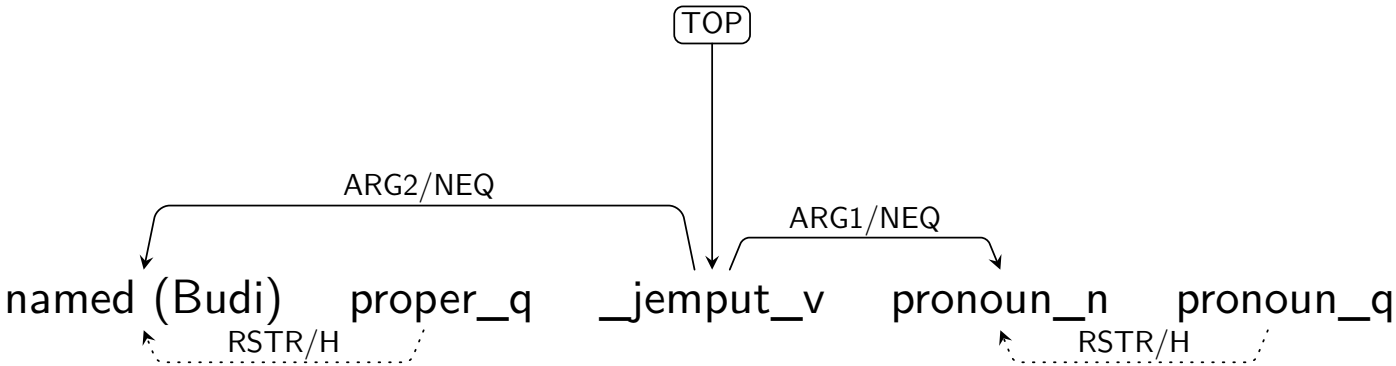
(4) P2 with bound forms (tree and DMRS)

(15)





(16)



NANYANG TECHNOLOGICAL UNIVERSITY

Summary and conclusions

- Bound forms *-nya*, *ku-*, *kau-*: tokenization
- constraint in NP coordination rule
- type hierarchies for nouns
- inflectional rules for P1 and P2
- *oleh* "by" as a semantically empty preposition
- *comp-head* rule, HEAD-DTR is of type *passive-two*
- Future work: evaluation

NANYANG TECHNOLOGICAL UNIVERSITY

Thank you

Terima kasih

NANYANG TECHNOLOGICAL UNIVERSITY

References I

I Wayan Arka, Christoffer Manning, et al. *Voice and grammatical relations in Indonesian: A new perspective*. Citeseer, 1998.

Sandra Chung. "On the subject of two passives in Indonesian". In: *Subject and topic* 57 (1976), p. 99.

Peter Cole, Gabriella Hermon, and Yassir Tjung. "Is there pasif semu in Jakarta Indonesian?" In: *Oceanic Linguistics* 45.1 (2006), pp. 64–90.

Ann Copestake et al. "Minimal Recursion Semantics: An Introduction". In: *Research on Language and Computation* 3.4 (2005), pp. 281–332.

Soenjono Dardjowidjojo. *Sentence patterns of Indonesian*. University of Hawaii Press, 1978.

David Moeljadi, Francis Bond, and Sanghoun Song. "Building an HPSG-based Indonesian Resource Grammar (INDRA)". In: *Proceedings of the GEAF Workshop, ACL 2015*. 2015, pp. 9–16. URL: http://aclweb.org/anthology/W/W15/W15-3302.pdf.

NANYANG TECHNOLOGICAL UNIVERSITY

References II

Carl Pollard and Ivan A Sag. *Head-driven phrase structure grammar*. University of Chicago Press, 1994.

Ivan A. Sag, Thomas Wasow, and Emily M. Bender. *Syntactic Theory: A Formal Introduction*. 2nd ed. Stanford: CSLI Publications, 2003.

James Neil Sneddon et al. *Indonesian Reference Grammar*. 2nd ed. New South Wales: Allen & Unwin, 2010.

Sanghoun Song. *Modeling information structure in a cross-linguistic perspective*. 2017.

Arnold M. Zwicky and Geoffrey K. Pullum. "Cliticization vs. Inflection: English N'T". In: *Language* 59.3 (Sept. 1983), pp. 502–513.

NANYANG TECHNOLOGICAL UNIVERSITY