

HPSG analysis and computational implementation of Indonesian passives

David Moeljadi 

Nanyang Technological University

Francis Bond 

Nanyang Technological University

Proceedings of the 25th International Conference on
Head-Driven Phrase Structure Grammar

University of Tokyo


Stefan Müller, Frank Richter (Editors)

2018

Stanford, CA: CSLI Publications

pages 127–137

Keywords: HPSG, Indonesian, passives, INDRA

Moeljadi, David & Francis Bond. 2018. HPSG analysis and computational implementation of Indonesian passives. In Stefan Müller & Frank Richter (eds.), *Proceedings of the 25th International Conference on Head-Driven Phrase Structure Grammar, University of Tokyo*, 127–137. Stanford, CA: CSLI Publications. DOI: 10.21248/hpsg.2018.8. 

Abstract

This study aims to analyze and develop a detailed model of syntax and semantics of passive sentences in standard Indonesian in the framework of Head-Driven Phrase Structure Grammar (HPSG) (Pollard & Sag, 1994; Sag et al., 2003) and Minimal Recursion Semantics (MRS) (Copestake et al., 2005), explicit enough to be interpreted by a computer, focusing on implementation rather than theory. There are two main types of passive in Indonesian, following Sneddon et al. (2010, pp. 256-260) and Alwi et al. (2014, pp. 352-356), called ‘passive type 1’ (P1) and ‘passive type 2’ (P2). Both types were analyzed and implemented in the Indonesian Resource Grammar (INDRA), a computational grammar for Indonesian (Moeljadi et al., 2015).

1 Introduction

A passive is a semantically transitive (two-participant) clause. Typically, the agent is either omitted or demoted to an oblique role, the other core participant possesses all properties of subjects, and the verb possesses formal properties of intransitive verbs (Payne, 2008, p. 204). Passive constructions are far more frequent in Indonesian than in English; an Indonesian passive is often naturally translated into English by an active construction (Sneddon et al., 2010, pp. 256, 263-264). Passive constructions in Indonesian are used in imperatives and for politeness, as well as in relative clauses which can only relativize subjects on defining relative clauses.

Research on Indonesian passives has been done by many linguists, such as McCune (1979), Voskuil (2000), Arka & Manning (2008), Cole et al. (2008), and Nomoto (2013). There has been a lot of linguistic work on Indonesian voice, in particular the status of passive-like structures in Indonesian and Austronesian languages (Musgrave, 2001; Riesberg, 2014). However, to the best of our knowledge, no research on Indonesian passives has been done in the HPSG framework. Our analysis is implemented in the Indonesian Resource Grammar (INDRA), a computational grammar for Indonesian (Moeljadi et al., 2015), which can parse and generate sentences.

There are two main types of passive in Standard Indonesian,¹ following Sneddon et al. (2010, pp. 256-260) and Alwi et al. (2014, pp. 352-356). They are called ‘passive type 1’ (P1) and ‘passive type 2’ (P2).² Both types are available for monotransitive and ditransitive verbs. They promote an object to subject. If there are two objects in an active ditransitive clause, only the one immediately following the verb (which has semantic role as patient or recipient) can be promoted to subject of the passive (Sneddon et al., 2010, p. 260).³ P1 and P2 are in (near) complementary

¹Indonesian is a diglossic language. This paper only deals with the ‘High’ variety of Indonesian, also known as the standard or formal Indonesian.

²Other types such as passives with prefix *ter-* and circumfix *ke-...-an* have not been analyzed and implemented in INDRA. They are for future work.

³This study only describes passives for monotransitive verbs. However, the analysis proposed here can be applied to ditransitive verbs as well.

distribution. P1 takes only a third person agent, while P2 may take first, second, and third person agent. P1 and P2 overlap with respect to the third person agent (Sneddon et al., 2010, p. 256).

2 Basic Data

2.1 Passive type 1

The verb in P1 is morphologically built by attaching a prefix *di-* to a transitive verbal stem (lexeme) in the lexicon. The subject (which usually has semantic role as agent) in the active sentence becomes an optional complement, immediately follows the passive verb (post-verbal), and it is optionally marked by a semantically empty preposition *oleh* ‘by’. Its PERNUM is third person, i.e. pronoun *dia* ‘3SG’, *mereka* ‘3PL’, enclitic *=nya* ‘3SG’, (common) noun, or proper name (Sneddon et al., 2010, p. 256-257). The position of the components of the predicate, such as auxiliaries and temporal markers, as well as the negative word *tidak* ‘NEG’ remain unchanged, i.e. they immediately precede the verb predicate both in active and passive voice.

Example (1a) shows a transitive sentence in active voice.⁴ An aspect marker *sudah* ‘PRF’ immediately precedes the active voice verb *menjemput* ‘ACT-pick.up’. Its corresponding P1 constructions are shown in Example (1b) to (1e). The position of the aspect marker is the same in all example sentences in (1). Example (1b), (1c), and (1d) show the optional preposition *oleh* ‘by’. Example (1c) and (1d) show that the enclitic *=nya* ‘3SG’ can attach directly to the passive verb or to the preposition *oleh* ‘by’. Example (1e) shows that a P1 construction may occur without a complement.

- (1) a. *Dia sudah menjemput Budi.*
 3SG PRF ACT-pick.up Budi
 ‘He has met Budi.’ (lit. ‘He has picked Budi up.’) (based on Sneddon et al., 2010, p. 256)
- b. *Budi sudah dijemput (oleh) dia.*
 Budi PRF PASS-pick.up by 3SG
 ‘Budi has been picked up by him.’ (based on Sneddon et al., 2010, p. 257)
- c. *Budi sudah dijemputnya.*
 Budi PRF PASS-pick.up=3SG
 ‘Budi has been picked up by him.’ (based on Sneddon et al., 2010, p. 257)

⁴A number of nasalization (sound changes) or morphology process occurs when *meN-* ‘ACT’ combines with stems, listed up in Moeljadi et al. (2015).

- d. Budi sudah dijemput olehnya.
 Budi PRF PASS-pick.up by=3SG
 ‘Budi has been picked up by him.’ (based on Sneddon et al., 2010, p. 257)
- e. Budi sudah dijemput.
 Budi PRF PASS-pick.up
 ‘Budi has been picked up.’ (based on Sneddon et al., 2010, p. 257)

In a coordinative construction with two or more passive verbs, the agent (both full forms and the bound form or enclitic =*nya*) can appear only once, following the last passive verb, as shown in (2).

- (2) Budi sudah ditunggu dan dijemputnya.
 Budi PRF PASS-wait and PASS-pick.up=3SG
 ‘Budi has been waited and picked up by him.’

2.2 Passive type 2

The verb in P2 is morphologically built by not attaching any affixes to a transitive verb lexeme in the lexicon. The verbs appear in bare stem form. Different from P1, the subject (agent) in the active sentence becomes an obligatory complement (argument), immediately preceding the verb (pre-verbal), without any prepositions such as *oleh* ‘by’. The agent is a pronoun such as *aku* ‘1SG’, *engkau* ‘2SG’, *dia* ‘3SG’ etc. or ‘pronoun substitute’, i.e. kinship terms such as *bapak* ‘father’, *ibu* ‘mother’, and personal names which can refer to the addressee, meaning ‘you’, or to the speaker, meaning ‘I’ (Sneddon et al., 2010, pp. 257, 259). No other component of the clause, such as negative and temporal marker, can come between the NP agent and the P2 verb (Sneddon et al., 2010, p. 258). They must occur before the agent.

Example (3) shows the corresponding P2 construction of Example (1a). The aspect marker *sudah* ‘PRF’ precedes the agent *dia* ‘3SG’.

- (3) Budi sudah dia jemput.
 Budi PRF 3SG pick.up
 ‘Budi has been picked up by him.’ (based on Sneddon et al., 2010, p. 257)

If the agent is *aku* ‘1SG’ or *engkau* ‘2SG’, the bound forms (also called as ‘proclitics’ by some grammarians) *ku-* ‘1SG’ and *kau-* ‘2SG’ usually occur (Sneddon et al., 2010, p. 258), as shown in (4).

- (4) Budi sudah kujemput.
 Budi PRF 1SG-pick.up
 ‘I have met Budi.’ (lit. ‘Budi has been picked up by me.’) (based on Sneddon et al., 2010, p. 257)

$$(6) \quad \left[\begin{array}{l} \text{INPUT} \left\langle X, \left[\begin{array}{l} \text{lexeme (tr-verb-lex)} \\ \text{ARG-ST} \left\langle [\dots], [\dots] \right\rangle \end{array} \right] \right\rangle \\ \\ \text{OUTPUT} \left\langle (\text{di-/ku-/kau-})X, \left[\begin{array}{l} \text{word (passive-transitive-lex-item)} \\ \text{ARG-ST} \left\langle \left[\begin{array}{l} \text{INDEX} \quad \boxed{1} \\ \text{ICONS-KEY} \quad \boxed{3} \end{array} \right], \left[\begin{array}{l} \text{INDEX} \quad \boxed{2} \\ \text{ICONS-KEY} \quad \boxed{4} \end{array} \right] \right\rangle \\ \text{LKEYS.KEYREL} \left[\begin{array}{l} \text{ARG1} \quad \boxed{2} \\ \text{ARG2} \quad \boxed{1} \end{array} \right] \\ \text{ICONS} \left\langle ! \boxed{3} \left[\begin{array}{l} \text{focus-or-topic} \\ \text{IARG2} \quad \boxed{1} \end{array} \right], \boxed{4} \left[\begin{array}{l} \text{non-topic} \\ \text{IARG2} \quad \boxed{2} \end{array} \right] ! \right\rangle \end{array} \right] \right\rangle \end{array} \right]$$

In a coordinative construction with two or more passive verbs, the bound forms usually occur before each passive verb, as shown in (5a).

- (5) a. Budi sudah kutunggu dan kujemput.
 Budi PRF 1SG-wait and 1SG-pick.up
 ‘Budi has been waited and picked up by me.’
- b. ??Budi sudah kutunggu dan jemput.
 Budi PRF 1SG-wait and pick.up

3 Analysis

We treat passive as an inflectional rule, as shown in (6). The input is a lexeme, of type *tr-verb-lex*, which has two arguments. The output is a word, of type *passive-transitive-lex-item* which adds the semantic information for passives, i.e. its ARG1 is coindexed with the ARG0 of the complement (agent) and its ARG2 with the subject. The prefix *di-*, *ku-*, or *kau-* may be attached. Following Song (2017, pp. 211-214), we added information in the ICONS. The promoted argument or the subject is marked as *focus-or-topic*, while the demoted argument is marked as *non-topic*.

We treat *ku-* ‘1SG’, *kau-* ‘2SG’, and *=nya* ‘3SG’ differently because of the difference in their occurrence in coordinative constructions and their optionality. Following Zwicky & Pullum (1983) who distinguish clitics from inflectional affixes, we tokenize *=nya*, treating it as a word which belongs to a type *encl-3pers*. One of the reasons is because *=nya* can attach both to the verb or to a preposition. On the other hand, we do not tokenize *ku-* and *kau-* and treat them as inflectional affixes.

We made four lexical rules for P1 and P2, as shown in Figure 1. The first rule is for P1 (having an optional complement) without *oleh* ‘by’ and the second one

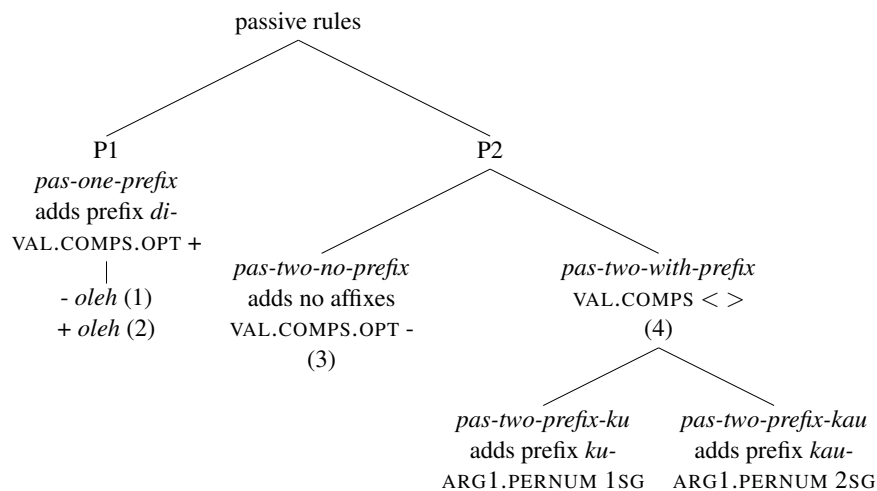
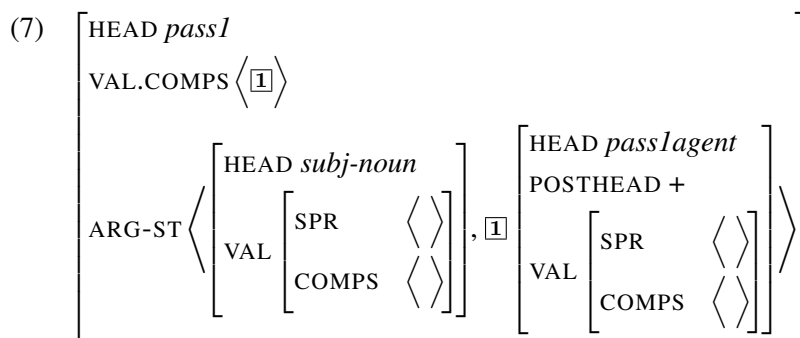


Figure 1: Type hierarchy for passive lexical rules

is with *oleh*. The third rule is for P2 (having an obligatory complement) without affixes and the fourth one is for P2 with a saturated complement and a prefix *ku-* or *kau-*. The details of each rule will be discussed in the next section.

3.1 Passive type one

We define a rule for P1, called *pas-one-prefix*. It is a rule which adds a prefix *di-* ‘PASS’. It inherits from the inflectional rule in (6). The output is a word, of type *passive-one-verb-lex*, with an optional complement, which inherits from *passive-transitive-lex-item*. It contributes the HEAD value, which is of type *pass1*. The COMPS has one item as its value. It has a feature POSTHEAD whose value is plus, as shown in (7). The COMPS’s HEAD is of type *pass1agent*. Its type hierarchy is shown in Figure 2.



The parse tree of (1c) is shown in Figure 3. It shows the *pas-one-prefix* rule changes the lexeme *jemput* ‘pick.up’ to an inflected passive word *dijemput* ‘PASS-pick.up’. The inflected passive word is combined with its optional complement

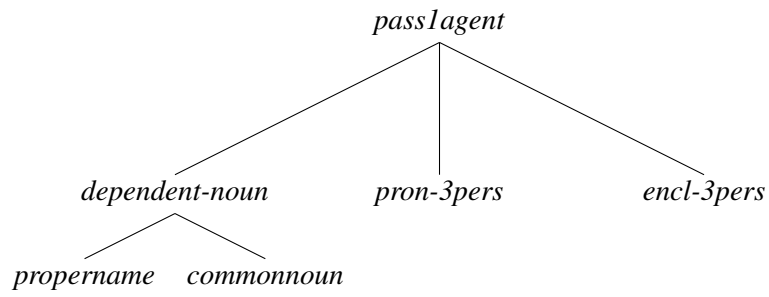


Figure 2: Type hierarchy for P1 agent

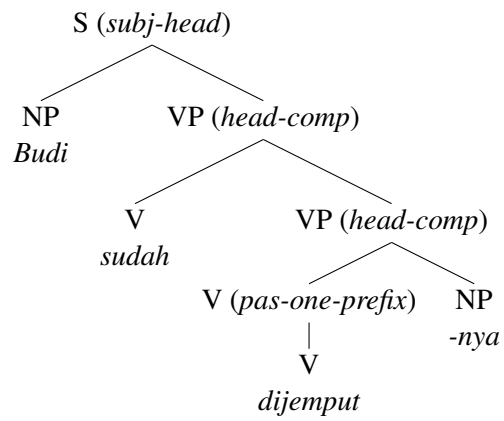


Figure 3: Parse tree of *Budi sudah dijemputnya*

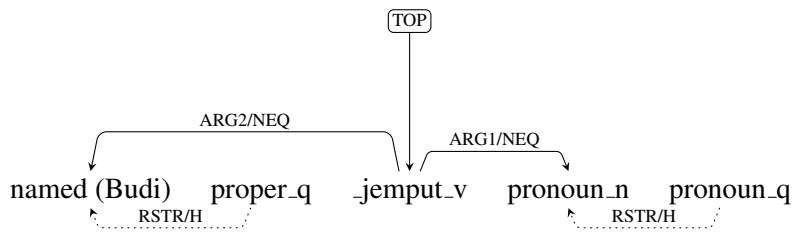


Figure 4: DMRS of *Budi sudah dijemputnya*

via *head-comp* rule. The semantics of the passive sentences in examples (1b) to (1d) look very much like the semantics of their active sentence counterpart in (1a), as shown in Figure 4, with additional information on the information structure. The ARG1 is linked to the optional agent complement and the ARG2 linked to the subject.

We treat *oleh* ‘by’ as a semantically empty preposition. It adds nothing to the meaning except the information that the COMPS of the passive verb is coindexed with the one of *oleh*, as shown in (8). The semantics of the PP headed by *oleh* is identical to that of *oleh*’s NP complement.

$$(8) \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{oleh-adp} \\ \text{MOD.LOC.CAT} \left[\begin{array}{l} \text{HEAD} \quad \textit{pass1} \\ \text{VAL.COMPS} \quad \langle \mathbf{1} \rangle \end{array} \right] \end{array} \right] \\ \text{VAL.COMPS} \langle \mathbf{1} \rangle \end{array} \right]$$

3.2 Passive type two

We made a rule *pas-two-no-prefix*, which adds no affixes for P2. It inherits from the same inflectional rule and the output is a word, of type *passive-two-verb-lex*, with an obligatory complement. Its AVM is shown in (9). It takes two saturated noun phrase arguments: the first argument is the subject whose HEAD’s value is of type *subj-noun* and the second argument is the sole item in the COMPS whose HEAD’s value is of type *pass2agent* and it has a feature *POSTHEAD* whose value is minus, i.e. it must occur before the head verb. The type hierarchy for *pass2agent*, which is the head type for agent in P2, is shown in Figure 5.⁵ The type *pass1agent* (see Figure 2) and *pass2agent* have *propername* and *pron-3pers* as their subtypes.

⁵Another approach is to analyze P2 agents as “lite” pronouns (Abeillé & Godard, 2001) because they must be adjacent to the P2 verbs but can be coordinated, like *aku atau dia* ‘me or him/her’ or modified, like *aku sendiri* ‘me alone’. At present, we are still analyzing the possibility of this approach.

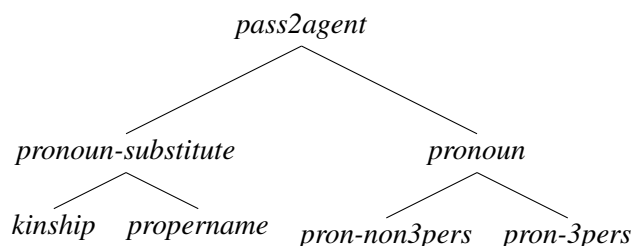


Figure 5: Type hierarchy for P2 agent

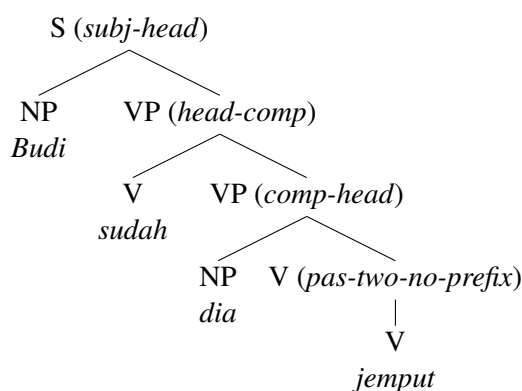
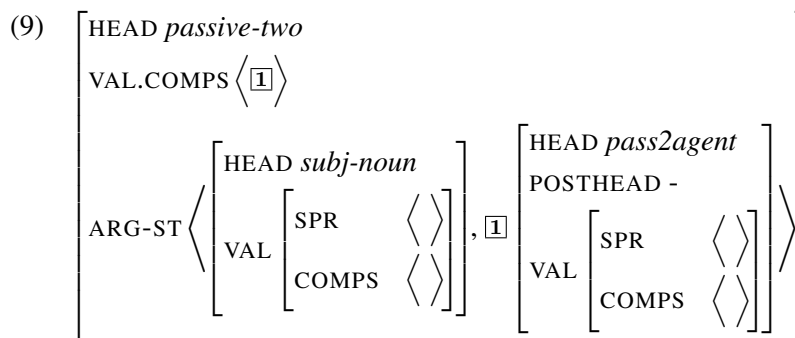


Figure 6: Parse tree of *Budi sudah dia jemput*



In addition, we made a new phrase rule called *complement-head* rule, which is constrained to lexical P2 head only. The HEAD value of its HEAD-DTR is of type *passive-two*. Parse tree of (3) is shown in Figure 6. The complement (agent) and P2 verb are combined by *complement-head* rule, the result is combined with the aspect marker by *head-complement* rule. Its semantics is similar to the one shown in Figure 4.

For P2 with *ku-* ‘1SG’, we made a rule *pas-two-prefix-ku* which adds *ku-*. It adds the semantic information that the PERNUM of the ARG1 is first person singular. The COMPS is saturated.

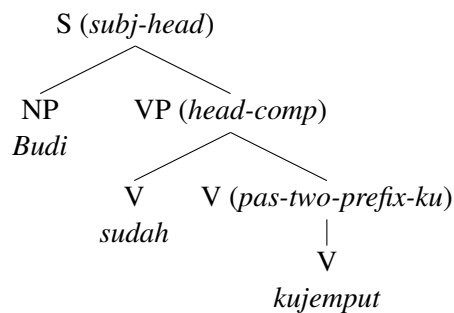


Figure 7: Parse tree of *Budi sudah kujemput*

The result is a passive verb with *ku-* whose COMP's value is empty (saturated) but still needs a subject. The verb's ARG2 is coindexed with the INDEX of the subject, whose HEAD's value is of type *subj-noun*.

Parse tree of (4) is shown in Figure 7. It shows *pas-two-prefix-ku* rule makes the lexeme *jemput* 'pick.up' become *kujemput* '1SG-pick.up'. The result is the verb *kujemput* '1SG-pick.up' which has *aku* '1SG' in the semantics, coindexed with the ARG1 of the verb. This verb is then combined with an aspect marker *sudah* 'PRF' by *head-complement* rule. Its semantics is similar to the one shown in Figure 4.

For P2 with *kau-*, we treat it similarly as for P2 with *ku-*. We made a rule *pas-two-prefix-kau* which adds *kau-* with the PERNUM of ARG1 is second person singular.

4 Conclusion

We made four rules for two types of passive (P1 and P2) and type hierarchies for the complement nouns (agent). Due to the optionality of the complements in coordinative constructions, the bound pronouns *-nya* '3SG', *ku-* '1SG', and *kau-* '2SG' are treated differently: *-nya* is treated as a word, while *ku-* and *kau-* are treated as affixes. We made a *complement-head* rule which combines a complement with a P2 verb without affixes.

References

- Abeillé, Anne & Daniele Godard. 2001. A Class of 'Lite' Adverbs in French. *Romance syntax, semantics and their L2 acquisition* 9–26.
- Alwi, Hasan, Soenjono Dardjowidjojo, Hans Lapoliwa & Anton M. Moeliono. 2014. *Tata Bahasa Baku Bahasa Indonesia*. Jakarta: Balai Pustaka 3rd edn.
- Arka, I Wayan & C. D. Manning. 2008. Voice and grammatical relations in Indone-

- sian: a new perspective. In *Voice and Grammatical Relations in Austronesian Languages*, 45–69. Stanford: CSLI Publications.
- Cole, Peter, Gabriella Hermon & Yanti. 2008. Voice in Malay/Indonesian. *Lingua* 118(10). 1500–1553.
- Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan A. Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research on Language and Computation* 3(4). 281–332.
- McCune, Keith. 1979. Passive Function and the Indonesian Passive. *Oceanic Linguistics* 18(2). 119–169.
- Moeljadi, David, Francis Bond & Sanghoun Song. 2015. Building an HPSG-based Indonesian Resource Grammar (INDRA). In *Proceedings of the GEAF Workshop, ACL 2015*, 9–16. <http://aclweb.org/anthology/W/W15/W15-3302.pdf>.
- Musgrave, Simon. 2001. *Non-subject arguments in Indonesian*. Melbourne: The University of Melbourne PhD dissertation.
- Nomoto, Hiroki. 2013. On the optionality of grammatical markers: A case study of voice marking in Malay/Indonesian. *Voice variation in Austronesian languages of Indonesia* 121–143.
- Payne, Thomas E. 2008. *Describing Morphosyntax: A Guide for Field Linguists*. New York: Cambridge University Press.
- Pollard, Carl & Ivan A. Sag. 1994. *Head Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Riesberg, Sonja. 2014. *Symmetrical Voice and Linking in Western Austronesian* Pacific Linguistics 646. Berlin: Mouton de Gruyter.
- Sag, Ivan A., Thomas Wasow & Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Stanford: CSLI Publications 2nd edn.
- Sneddon, James Neil, Alexander Adelaar, Dwi Noverini Djenar & Michael C. Ewing. 2010. *Indonesian Reference Grammar*. New South Wales: Allen & Unwin 2nd edn.
- Song, Sanghoun. 2017. *Modeling information structure in a cross-linguistic perspective*. Berlin: Language Science Press.
- Voskuil, Jan E. 2000. Indonesian voice and A-bar movement. In *Formal issues in austronesian linguistics*, 195–213. Springer.
- Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. Inflection: English N'T. *Language* 59(3). 502–513.