

インドネシア語の受身文における 動作主の表示形式について

佐近 優太

(東京外国語大学大学院博士後期課程/日本学術振興会特別研究員)

研究対象

- ・ 受身文における動作主の表示形式: oleh 「によって」とzero

本研究が取り組む問題

- ・ 接頭辞di-とter-という受身形式があり、接頭辞di-の方は研究が進んでいるが、その分析をter-に当てはめることができない

本研究の提案

- ・ 動詞の種類や動作主の特性が選択要因となる(=接頭辞di-とは異なる説明を与える)

本研究の意義

- ・ 記述の少ない接頭辞ter-についての分析
- ・ zero標示の分析

本発表の目的

研究対象

- ・ 受身文における動作主の表示形式: oleh 「によって」 と zero

日本語の受身文におけるマーカー

- (1) 彼は友達{二 / カラ}好かれている。
- (2) この論文は多くの人{二 / ニヨッテ}引用された。

インドネシア語の受身文におけるマーカー

- (3) **Buku** **itu** **di-baca** **{oleh / Ø}** **adik.**
book that PASS-read by / zero younger.brother
「その本は弟に読まれた」 (Jeoung & Biggs 2017: 85)

研究対象

- ・ 受身文における動作主の表示形式: oleh 「によって」 とzero

本研究が取り組む問題

- ・ 接頭辞di-とter-という受身形式があり、接頭辞di-の方は研究が進んでいるが、その分析をter-に当てはめることができない

本研究の提案

- ・ 動詞の種類や動作主の特性が選択要因となる(=接頭辞di-とは異なる説明を与える)

本研究の意義

- ・ 記述の少ない接頭辞ter-についての分析
- ・ zero標示の分析

研究対象

- ・ 受身文における動作主の表示形式: oleh 「によって」とzero

本研究が取り組む問題

- ・ 接頭辞di-とter-という受身形式があり、接頭辞di-の方は研究が進んでいるが、その分析をter-に当てはめることができない

本研究の提案

- ・ 動詞の種類や動作主の特性が選択要因となる(=接頭辞di-とは異なる説明を与える)

本研究の意義

- ・ 記述の少ない接頭辞ter-についての分析
- ・ zero標示の分析

研究対象

- ・ 受身文における動作主の表示形式: oleh 「によって」とzero

本研究が取り組む問題

- ・ 接頭辞di-とter-という受身形式があり、接頭辞di-の方は研究が進んでいるが、その分析をter-に当てはめることができない

本研究の提案

- ・ 動詞の種類や動作主の特性が選択要因となる(=接頭辞di-とは異なる説明を与える)

本研究の意義

- ・ 記述の少ない接頭辞ter-についての分析
- ・ zero標示の分析

1-1 インドネシア語の特徴

類型論的特徴

- ・オーストロネシア語族マレー・ポリネシア語派に属するマレー語を基盤とする
- ・語順 S V O
- ・形態的に膠着語的特徴を持つ – 接辞による派生が多い

1-1 インドネシア語の特徴

インドネシア語の態

(4) a. **Saya** **mem-bawa** **buku** **saudara**.

1SG ACT-take book brother

「私はあなたの本を持って行った」

b. **Buku** **saudara** {**di/ter**}-bawa **oleh** **saya**.

book brother {PASS/TER}-take by 1SG

「あなたの本は私が持って {行った / 行ってしまった}」 (cf. Sneddon et al. 2010: 118)

→(意味面で)ter-は有標の受動態標識(結果状態・可能・非意図)

1-2 先行研究

接頭辞di-における動作主表示

(3) **Buku itu di-baca {oleh / Ø} adik.**

book that PASS-read by / zero

younger.brother

「その本は弟に読まれた」

(Jeoung & Biggs 2017: 85)

→oleh 「によって」は

1. 接頭辞di-以外が使われているとき

2. 前置詞句の移動が起こるとき

3. 挿入句があるとき

にzeroをつかうことができない

1-2 先行研究

1. 接頭辞di-以外が使われているとき (Jeoung & Biggs 2017: 85)

(4) **Pintu itu ter-buka *(oleh) angin.**

door that ACCID-open by wind

「そのドアは風によって開いてしまった」

2. 前置詞句の移動が起こるとき (Jeoung & Biggs 2017: 86)

(5) ***(Oleh) Pak Joko buku itu di-tulis pada tahun 2001.**

by Mr Joko book that PV-write at year 2001

「ジョコさんによってその本は2001年に書かれた」

3. 挿入句があるとき (Jeoung & Biggs 2017: 86)

(6) **Semua anak di-jahit-kan baju *(oleh) Ibu Mindy.**

all child PV-sew-BEN shirt by Ms. Mindy

「すべての子供たちはミンディさんに服を縫ってもらった」

1-2 先行研究

Olehの省略に関する先行研究の分析

- ・ 特定の条件に反しない場合olehが〇形をとることができる
- ・ olehと省略(zero標示)は意味的に等価
 - (i) Competition: a form α or a form β may be used
(Kroch 1989)
 - (ii) Variable rule: α may take the form β (in some environment)**
(Labov 1969)

1-2 問題点

1. 接頭辞di-以外が使われているとき

- (4) **Pintu itu ter-buka *(oleh) angin.**
door that ACCID-open by wind
「そのドアは風によって開いてしまった」

2. 前置詞句の移動が起こるとき

- (5) ***(Oleh) Pak Joko buku itu di-tulis pada tahun 2001.**
by Mr Joko book that PV-write at year 2001
「ジョコさんによってその本は2001年に書かれた」

3. 挿入句があるとき

- (6) **Semua anak di-jahit-kan baju *(oleh) Ibu Mindy.**
all child PV-sew-BEN shirt by Ms. Mindy
「すべての子供たちはミンディさんに服を縫ってもらった」

1-2 問題点

接頭辞ter-の場合の記述不足という問題

・確かに先行研究の例は省略できないが、接頭辞ter-でもolehがない(方が自然な)場合がある

(7) Se-banyak 198 anak di Jawa Barat ter-infeksi Covid-19.

one-many 198 child in Jawa west TER-infect Covid-19
「198名の西ジャワの子供がコロナウイルスに感染した」 (Kompas.id)

<https://www.kompas.id/baca/nusantara/2020/07/07/198-anak-di-jabar-terinfeksi-covid-19-mayoritas-tertular-dari-orangtua/>

1-2 問題点

接頭辞ter-の場合の記述不足という問題

(4) **Pintu itu ter-buka *(oleh) angin.**
door that ACCID-open by wind

「そのドアは風によって開いてしまった」

(7) **Se-banyak 198 anak di Jawa Barat ter-infeksi Covid-19.**
one-many 198 child in Jawa west TER-infect Covid-19

「198名の西ジャワの子供がコロナウイルスに感染した」

→ 「このような選択の要因は何か」 を考えるべき

1-2 先行研究

本研究の提案

コーパス調査を行い、以下のことを主張する

- ・ 接頭辞 *ter-* では *oleh* / \emptyset は等価ではなく、特定の要因によって交替する別の機能を持つ標示である

(i) Competition: a form α or a form β may be used
(Kroch 1989i.a.)

(ii) Variable rule: α may take the form β (in some environment)
(Labov 1969i.a.)

2-1. データ

使用コーパス (cf. Shiohara et al. 2019)

Leipzig corpus collection (Goldhahn, Eckhart & Quasthoff 2012) 内の

Mixed-tufs4(ジャンルを問わないウェブサイト: 30万文、5,428,067語)

Web-tufs13(ブログなどのサイト: 30万文、5,540,573語)

Wikipedia-tufs14(Wikipediaの記事: 30万文、5,634,138語)

(cf: Nomoto, Akasegawa and Shiohara 2018)

2-1. データ

データ整形

- ・テキストデータから検索を行い、手作業でノイズを除去

	Verb ter-	Other ter-	Non-prefix ter-	All
Type size	2,095	154	1,045	3,294
Token size	102,642	162,282	25,752	209,676

→olehが共起する頻度の高い35個の動詞を調査対象とする

2-2. 調査対象

調査対象

Word	Oleh	Zero	Total				
terpengaruh 「影響を受ける」	108	65	326	terasa 「感じられる」	19	0	2,711
terinspirasi 「感化される」	108	17	245	terpisah 「分けられる」	19	4	1,231
terlihat 「見られる」	81	12	4,506	terdesak 「詰められる」	19	10	163
terganggu 「邪魔される」	72	18	673	terkesan 「感じられる」	18	1	601
tertutup 「閉まる」	58	119	920	terhambat 「妨げられる」	18	6	156
terbunuh 「殺される」	54	27	557	terbatas 「制限される」	15	8	2,076
tertangkap 「捕まる」	48	4	355	terbakar 「燃やされる」	15	22	522
terikat 「結ばれる」	45	67	537	tertekan 「抑圧される」	15	3	168
terdengar 「聞かれる」	40	6	1,403	terjebak 「罠にかけられる」	13	37	327
tersentuh 「触られる」	40	47	180	terhubung 「関係づけられる」	12	0	376
terbentuk 「形成される」	35	3	1,879	tergantung 「吊られる」	11	0	2,238
terancam 「脅される」	29	48	448	terluka 「傷つけられる」	11	5	723
terjangkau 「手を差し伸べられる」	28	19	216	terpilih 「選ばれる」	9	8	1,371
terdorong 「押される」	27	5	203	tersusun 「積まれる」	9	0	320
tertarik 「引かれる」	25	13	1,212	tercemar 「汚される」	9	31	161
terinfeksi 「感染する」	20	68	278	terbukti 「証明される」	6	0	2,714
terbawa 「持っていかれる」	20	85	194	terbuka 「開く」	2	0	1,251

2-3. 調査方法

分析手法

動作主標示の選択の要因

- ・ 動作主の有生性 (Vangunu [Austronesian, Solomon Island]: Zúñiga & Kittilä 2019: 94)
- ・ 動作主の長さ(語数) (Indonesian: Nomoto 2021: 67, Moeliono et al. 2017:470)
- ・ **動詞の種類** (Japanese: Oshima 2003: 258など多数)

→どれくらいoleh/Øの選択に影響を及ぼしているか

2-3. 調査方法

動詞の種類決め方

できるだけ客観的な方法で決めたい (cf. Stefanowistch & Gries 2009: 948)

→空間ベクトルを用いた階層的クラスタ分析

→共起語を基にして意味の近さを測る (Landauer and Dumais 1997)

2-3. 調査方法

動詞の種類の決め方

Python (3.8.5)

○ベクトル化 : word2vec

(Mikolov, Yih & Zweig 2013; Mikolov, Chen, Corrado, & Dean 2013; Mikolov, Sutskever, Chen, Corrado, & Dean 2013)

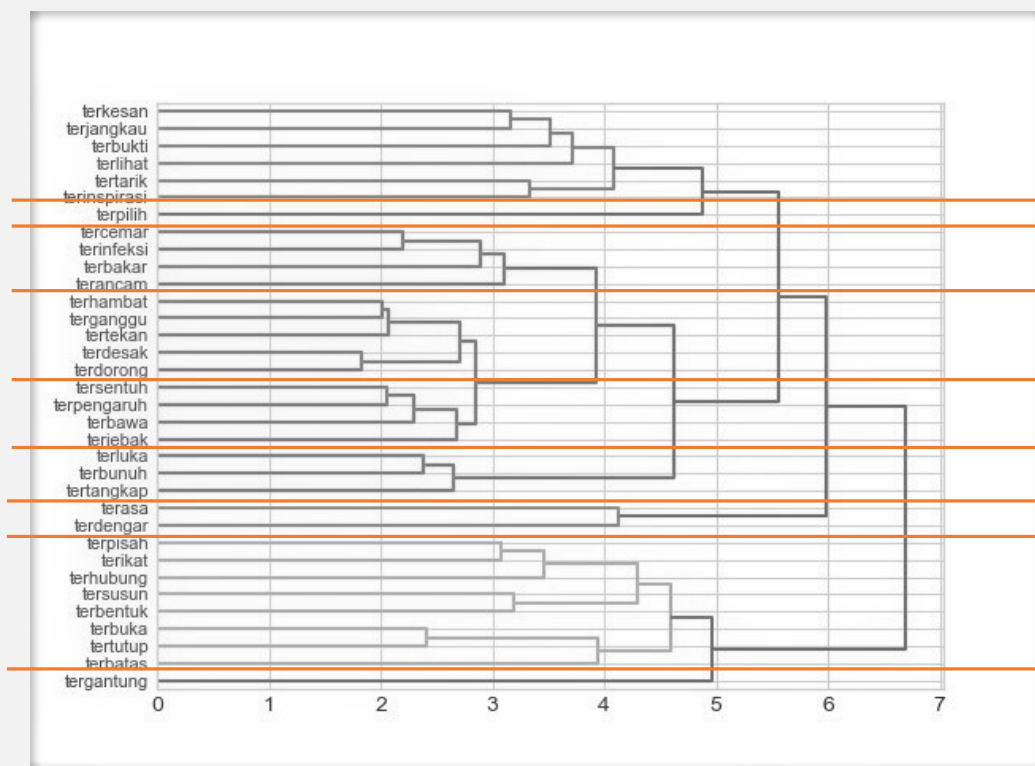
- sg = 1 [The training algorithm is Skip-gram , not CBOW]
- size = 200 [The dimensionality of the vector is 200.]
- min_count= 10 [Consider only words that occur more than 10 times.]
- window = 5 [Measure the five words before and after the word in question.]
- iter = 3 [Number of iterations (epochs) over the corpus is three]

○個体間距離計算: ユークリッド距離

○クラスター形成: Ward法

2-3. 調査方法

動詞の種類決め方



→全9タイプに分類

2-3. 調査方法

動詞の種類決め方

Type1: 「脅す」「感染する」「燃やす」「汚す」

Type2: 「邪魔する」「押す」「詰め寄る」「妨害する」

Type3: 「傷つける」「捕まえる」「殺す」

Type4: 「見る」「届く」「引く」「明らかにする」「印象を持つ」「感化する」

Type5: 「選ぶ」

Type6: 「聞く」「感じる」

Type7: 「閉める」「開ける」「結ぶ」「形成する」「分ける」「制限する」「関係づける」
「積む」

Type8: 「ぶら下がる」

Type9: 「影響する」「触れる」「持つ」「罠にかける」

2-3. 調査方法

動詞の種類決め方

Type1: terancam, terinfeksi, terbakar, tercemar

Type2: terganggu, terdorong, terdesak, terhambat

Type3: terluka, tertangkap, terbunuh

Type4: terlihat, terjangkau, tertarik, terbukti, terkesan, terinspirasi

Type5: terpilih

Type6: terdengar, terasa

Type7: tertutup, terikat, terbentuk, terpisah, terbatas, terhubung, tersusun, terbuka

Type8: tergabung

Type9: terpengaruh, tersentuh, terbawa, terjebak

2-3. 調査方法

分析手法

説明変数	カテゴリ変数	有生性・動詞の種類
	量的変数	語数
応答変数	二項分布	oleh/zero

→ ロジスティック回帰分析 (R 4.0.4を使用)
変数の選択は変数増加法

3. 結果

動作主の有生性

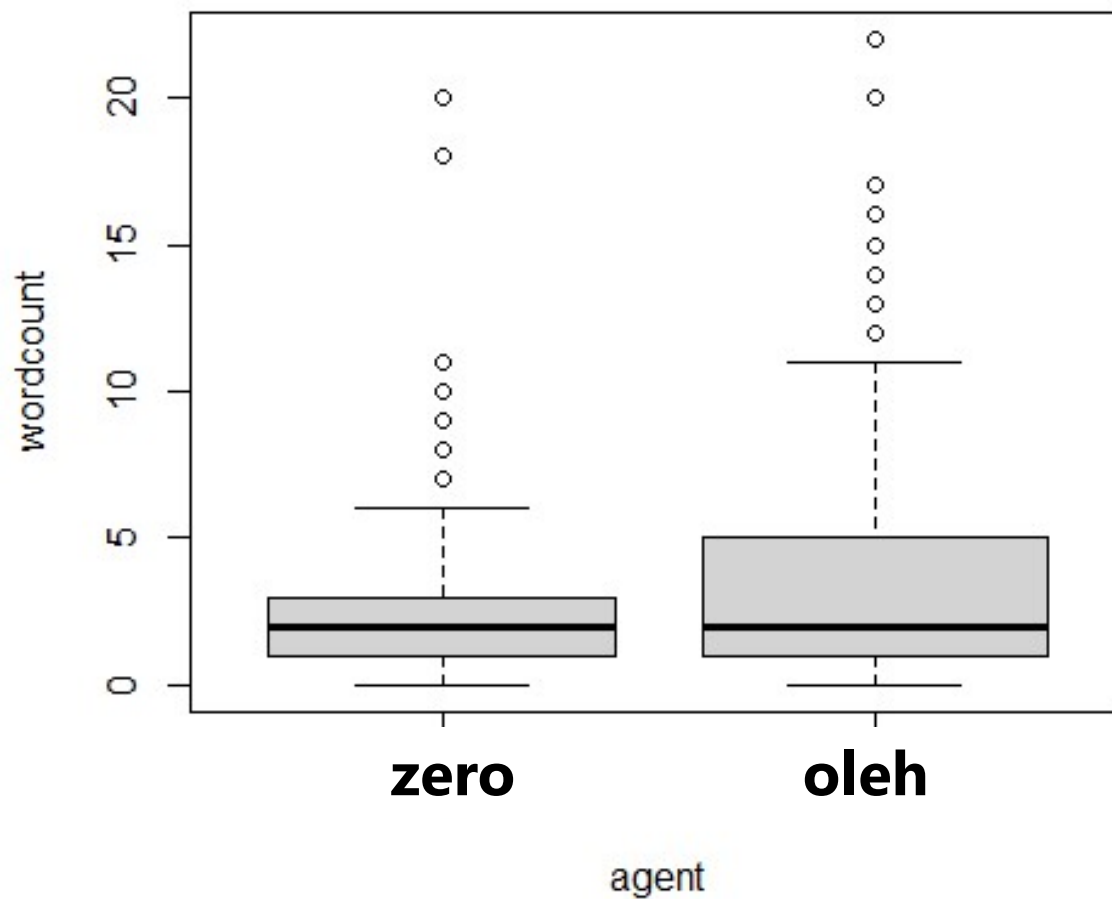
	oleh	zero	Total
animate	265	46	311
inanimate	779	712	1491
Total	1044	758	1802

p=1.968E-29<0.01
RR: 1.631 (1.525, 1.744)

3. 結果

動作主の語数

$t = -9.478$
 $df = 1788.7$
 $p\text{-value} < 2.2e-16$



3. 結果

動詞の種類

	Type1	Type2	Type3	Type4	Type5	Type6	Type7	Type8	Type9
Oleh	73 (30.2%)	149 (80.5%)	104 (74.3%)	265 (81.0%)	9 (52.9%)	59 (90.8%)	192 (48.1%)	11 (100%)	182 (43.7%)
Zero	169 (69.8%)	36 (19.5%)	36 (25.7%)	62 (19.0%)	8 (47.1%)	6 (9.2%)	207 (51.9%)	0 (0.0%)	234 (56.3%)

3. 結果

動詞の種類

	Type1	Type2	Type3	Type4	Type5	Type6	Type7	Type8	Type9
Oleh	73 (30.2%)	149 (80.5%)	104 (74.3%)	265 (81.0%)	9 (52.9%)	59 (90.8%)	192 (48.1%)	11 (100%)	182 (43.7%)
Zero	169 (69.8%)	36 (19.5%)	36 (25.7%)	62 (19.0%)	8 (47.1%)	6 (9.2%)	207 (51.9%)	0 (0.0%)	234 (56.3%)

3. 結果

動詞の種類

	Type1	Type2	Type3	Type4	Type5	Type6	Type7	Type8	Type9
Oleh	73 (30.2%)	149 (80.5%)	104 (74.3%)	265 (81.0%)	9 (52.9%)	59 (90.8%)	192 (48.1%)	11 (100%)	182 (43.7%)
Zero	169 (69.8%)	36 (19.5%)	36 (25.7%)	62 (19.0%)	8 (47.1%)	6 (9.2%)	207 (51.9%)	0 (0.0%)	234 (56.3%)

3. 結果

ロジスティック回帰分析

統計的に有意

	Estimate	Std. Error	z value	Pr(> z)		オッズ比	信頼区間(95%CI)	
(Intercept)	-0.81035	0.10136	-7.995	1.30E-15	***	0.4447	0.36386408	0.5414509
animate	1.22043	0.19289	6.327	2.50E-10	***	3.38863	2.33783700	4.9870043
wordcount	0.20385	0.02482	8.213	< 2e-16	***	1.22611	1.16918521	1.2886983
verbtype1	-0.69725	0.16274	-4.285	1.83E-05	***	0.49795	0.36031134	0.6824029
verbtype2	1.41869	0.20277	6.997	2.62E-12	***	4.13172	2.80508812	6.2235200
verbtype3	0.98226	0.21828	4.5	6.80E-06	***	2.67049	1.75328691	4.1346101
verbtype4	1.15979	0.16889	6.867	6.54E-12	***	3.18926	2.30106832	4.4651451
verbtype6	1.70877	0.45954	3.718	0.0002	***	5.52215	2.40276369	15.0066757
verbtype8	14.47403	252.98017	0.057	0.9544		NA	0.01405253	NA

OR < 1 はzeroが、OR > 1はolehが出やすい

3. 結果

ロジスティック回帰分析

Oter受身文においてoleh/zeroの選択は動作主の有生性・語数・動詞のタイプによって左右される

- 有生物・語数の長いものはolehをとりやすい
- type1はzeroをとりやすい
- その他のtypeはolehをとりやすい

4. 考察

有生性について

相対的に無生物はzeroの方を好む

(7) **Ilmuwan nuklir itu ter-bunuh {?oleh / Ø} senapan.**
scientist nuclear that TER-kill by / zero firearm

「その研究者は銃器で殺された (直訳：その研究者は銃器によって殺された)」

参考 suara.com

<https://www.suara.com/news/2020/12/09/060054/iran-ilmuwan-nuklir-terbunuh-dengan-senapan-mesin-dikendalikan-satelit?page=all>

4. 考察

動詞のタイプについて

Olehをとりやすい動詞群

Type2: 「邪魔される」「押される」「追い詰められる」「妨害される」

Type3: 「傷つけられる」「捕まる」「殺される」

Type4: 「見られる」「手を差し伸べられる」「引かれる」「証明される」
「感じられる(terkesan)」 「感化される」

Type6: 「聞かれる」「感じられる(terasa)」

→働きかけが感じられるものはolehをとりやすい

zeroをとりやすい動詞群

Type1: 「脅かされる」「感染させられる」「燃やされる」「汚される」

→通常自然と起こるものはzeroをとりやすい

4. 考察

動詞のタイプについて

対立を支持する例

(8) **Matahari ter-tutup {?oleh / Ø} awan.**

sun TER-close by / zero cloud

「太陽は雲に隠れた (直訳：太陽は雲によって隠された)」

(9) **Pintu itu ter-tutup {oleh / *Ø} angin.**

door that TER-close by / zero wind

「そのドアは風で閉まった (直訳：そのドアは風によって閉まった)」

4. 考察

動詞のタイプについて

Olehをとりやすい動詞群

Type2: 「邪魔される」「押される」「追い詰められる」「妨害される」

Type3: 「傷つけられる」「捕まる」「殺される」

Type4: 「**見られる**」「手を差し伸べられる」「引かれる」「**証明される**」
「**感じられる(terkesan)**」「感化される」

Type6: 「聞かれる」「**感じられる(terasa)**」

zeroをとりやすい動詞群

Type1: 「脅かされる」「感染させられる」「燃やされる」「汚される」

4. 考察

動詞のタイプについて

- ・ 知覚動詞について

意味面からの一般化に当てはまらない→意味の曖昧性の排除

(8) **Suara saya ter-dengar {oleh / *Ø} kucing.**

voice

1SG

TER-hear

by / zero

cat

「私の声は猫に聞かれた」

→私の声は猫に聞こえる (My voice sounds like a cat's.)のような解釈を避けるため

5. 結論

接頭辞ter-受動文におけるoleh/zeroの関係

選択の要因は？ →

(接頭辞di-のような統語的要請ではなく)

oleh/zeroは動作主の有生性・語数・動詞のタイプ

- ・ zeroは通常自然と起こる事象を表す場合に用いられ、olehは他動性の高い事象を表す場合に用いられやすい表現
- ・ 意味的な曖昧性がある場合は片方が排除される

5. 今後の課題

独立変数の追加について（他の要因について）

- ・まだ正確さをあげるべき→追加するべき要因がある

言語外要因：レジスター、年代

言語内要因：統語環境(主節/従属節)、否定、定性、Givenness

5. 今後の課題

文字数について

zeroにはすでに定型表現になっているものがある可能性
→だとすれば将来的にこれらをデータから排除することが必要

terbawa suasana	「雰囲気流される」
tertangkap mata	「目撃される (目に捕まる)」
terpisah waktu	「時間に隔てられた」
terjebak macet	「渋滞に捕まる」

参考文献

- Goldhahn, Dirk. & Thomas, Eckart. & Uwe, Quasthoff. 2012. "Building large monolingual dictionaries at the Leipzig Corpora Collection: From 100 to 200 languages." *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. 759–765.
- Jeoung, Helen. 2020. "P-Drop across Languages of Java: A Field Report." *NUSA: Linguistic Studies of Languages in and around Indonesia* 69. 27–41.
- Jeoung, Helen, and Alison Biggs. 2017. "Variants of Indonesian Prepositions as Intra-Speaker Variability at PF." *University of Pennsylvania Working Papers in Linguistics* 23 (1). 83–92.
- Kroch, Anthony S. 1989. "Reflexes of grammar in patterns of language change". *Language Variation and Change* 1. 199-244.
- Landauer, Thomas K. & Susan T. Dumais. 1997. "A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge." *Psychological Review* 104. 211–240.
- Labov, William. 1969. "Contraction, deletion, and inherent variability of the English copula". *Language* 45. 715-762.
- Mikolov, Tomáš, Wen-tau Yih, & Geoffrey Zweig. 2013. "Linguistic regularities in continuous space word representations." In Proceedings of the 2013 conference of the north american chapter of the association for computational linguistics. Human language technologies. 746–751.
- Mikolov, Tomáš, Kai Chen, Greg Corrado, & Jeffrey Dean. 2013. "Efficient estimation of word representations in vector space." arXiv preprint arXiv:1301.3781.
- Mikolov, Tomáš, Ilya Sutskever, Kai Chen, Greg Corrado, & Jeffrey Dean. 2013. "Distributed representations of words and phrases and their compositionality." arXiv preprint arXiv:1310.4546.
- Moeliono, Anton M., Hans Lapoliwa, Hasan Alwi, Sry Strya Tjatur Wisnu Sasangka, and Sugiyono. 2017. *Tata bahasa baku bahasa Indonesia 4th edition*. Jakarta: Balai Pustaka.
- Nomoto, Hiroki. 2021. "Bare passive agent hierarchy". In Henrison Hsieh and Keely New (eds.) *Proceedings of the Twenty-Seventh Meeting of the Austronesian Formal Linguistics Association (AFLA)*. 57-70. Ontario: University of Western Ontario.
- Nomoto, Hiroki, Shiro Akasegawa and Asako Shiohara, 2018. "Reclassification of the Leipzig Corpora Collection for Malay and Indonesian". *NUSA: Linguistic Studies of Languages in and around Indonesia* 65. 47–66.
- Oshima, David Y. 2003. "Out of Control: A Unified Analysis of Japanese Passive Constructions". In Jong-Bok Kim and Stephen Wechsler (eds.), *The Proceedings of the 9th International Conference on Head-Driven Phrase Structure Grammar*. 245–265.
- Sneddon, J. N., Adelaar, A., Djenar, D. N. & Ewing, M. 2010. *Indonesian: A Comprehensive Grammar 2nd edition*. Abingdon: Routledge.
- Zúñiga, Fernando, and Seppo Kittilä. 2019. *Grammatical voice*. Cambridge: Cambridge University Press.