



Article

Canonical Versus Non-Canonical Passives in Indonesian

*I Nyoman Udayana¹, I Nyoman Aryawibawa², I Made Madia³*¹⁻³Universitas Udayana, Denpasar, Bali, Indonesia**SUBMISSION TRACK**

Received: August 31, 2024

Final Revision: December 7, 2024

Accepted: December 10, 2024

Available Online: December 25, 2024

KEYWORDS

non-canonical passive constructions, unmarked form, oblique role, passive types, verb markers

CORRESPONDENCEE-mail: nyoman_udayana@unud.ac.id**A B S T R A C T**

While many studies have explored passive constructions in Indonesian, there is still a significant gap in research focusing on distinguishing passive clauses and categorizing them into canonical and non-canonical forms. This study aims to address this gap by identifying different types of passive clauses in Indonesian and classifying them into these two categories. The primary objectives are to establish clear criteria for validating the existence of canonical and non-canonical passive forms and to provide evidence supporting their distinction. To achieve these aims, we first analyze markers on Indonesian verbs, particularly the prefixes *di-*, *ter-*, and unmarked (zero) forms. The data used for this analysis consists of naturally occurring expressions and clauses sourced from the Leipzig Corpora. This data is examined through the lens of the (non)-canonical theory of passivization. The findings indicate that the *di-* marker on verbs predominantly signals canonical passive constructions. However, there are cases where the *di-* form shifts to non-canonical usage, especially when the agent is obligatorily present and cannot be syntactically demoted to an oblique role. In contrast, the *ter-* form and the unmarked form, which can denote either a bare active (BA) or a bare passive (BP) construction, consistently represent non-canonical passive constructions.

I. INTRODUCTION

In natural languages, passive clauses are generally regarded as derivatives of their active counterparts (Huddleston and Pullum 2005: 26). Fundamentally, active clauses serve as a universal type found in all natural languages worldwide. Nevertheless, distinctions arise between active and passive clauses concerning morphology and semantics.

Not all active sentences across languages exhibit morphological uniformity (Legate 2021; Keenan 2013; Shibatani 1988). In some languages, the active status of a clause is marked on the verb, while in others, the verb remains unmarked. This variability also extends to passive sentences. In some languages, passive constructions feature explicit markers on the verb, whereas in others, the verb remains unmarked. For example, Manggarai, a language spoken in East Nusa Tenggara, is considered to have passive constructions, yet the verbs in these clauses lack any passive marking

(Arka & Kosmas 2005). Another language of the type is Acehnese (Legate 2012, 2021). Semantically, passive clauses derived from active clauses typically involve valence reduction. Specifically, a trivalent verb in an active clause becomes divalent in its passive form, and a divalent verb in an active clause becomes monovalent in its passive form. However, given the variation in passive forms across languages, not all sentences classified as passive undergo changes in valence.

Our study focuses on Indonesian, a language in which passive formation is governed by verb morphology. This characteristic has led to its classification as having morphological passives (Keenan, 2013), setting it apart from the periphrastic passives commonly found in other languages. Canonical passives in Indonesian are defined by the presence of passive morphology and the application of the passive principle, particularly the demotion of the agent. However, not all passives adhere to these criteria; those that deviate are categorized as

non-canonical passives.

A significant study on Indonesian passives identified instances where passive clauses exhibit unique properties, emphasizing the importance of distinguishing between canonical and non-canonical forms. In the following, we review key studies on Indonesian passives to provide a comprehensive foundation for our analysis.

In her seminal work on the Object-Creating Rule in Indonesian, Chung (1976) identifies two distinct passive constructions in the language. First, she analyzes the *di-* passive, where the object of the active clause is promoted to the subject position, while the original subject is demoted to an adjunct role and is no longer considered an argument. Chung also examines another construction, which she classifies as a passive form, referred to as the object preposing construction. She argues that this construction aligns with passive derivation because the object of the active clause assumes the subject position, while the agent of the active clause is retained. Sneddon *et al.* (2010) support this classification, identifying the object preposing construction as a type 2 passive construction. However, Arka & Manning (1998) dispute this interpretation, arguing that the object preposing construction does not qualify as a true passive form. Similarly, Arka (1998, 2003) contends that a comparable construction in Balinese should also not be considered a passive construction. Additionally, Cole *et al.* (2010) describe this type of construction in (Jakarta) Indonesian as passive semu (“pseudo-passive”), further distinguishing it from canonical passive forms.

Alexiadou (2012) and Alexiadou and Schäfer (2013) categorize passives in English and other European languages into two types: canonical passives and non-canonical passives. Canonical passives correspond to the “be” passive, while non-canonical passives, often referred to as the “get” passive, deviate from the typical features of the “be” passive, as noted by Reed (2011). Non-canonical passives are so named because they do not fully conform to the standard characteristics of canonical passives. In some cases, they resemble anticausative constructions and share features with middle constructions. While these studies primarily focus on English, the distinction between canonical and non-canonical passives offers a useful framework for analyzing passives in Indonesian.

Udayana (2022) examines Indonesian passives and their discourse contexts, highlighting that they are derived from their active clause counterparts. This derivation process results in passive constructions that often exhibit structural differences. A key feature of this derivation is *theme promotion*, where the theme participant—originally the object in the active clause—is promoted to the subject position in the passive clause. Udayana further argues that the theme, if expressed as an indefinite noun phrase (NP) in the active clause, must be transformed into a definite noun phrase (DP) in the passive clause. This shift in definiteness ensures semantic continuity between the theme phrases in the active and passive forms, aligning with the principles of information structuring (Lyngfelt & Solstad, 2006). Retaining an indefinite NP object as an indefinite NP subject in the passive clause would result in an interpretation of two different entities, thereby violating these information-structuring principles. Another important observation in Udayana’s work (2022) concerns *long passives*, where the agent by-phrase is obligatory in the passive construction. As noted in Indonesian linguistics (e.g., Sneddon *et al.*, 2010), the use of first- and second-person agents in the by-phrase of long passives is prohibited. However, Udayana clarifies that this prohibition stems from pragmatic considerations rather than syntactic constraints.

None of the above studies addresses the canonical status of Indonesian passive clauses explicitly, which constitutes a gap that the present study aims to fill.

II. METHOD

This study is a descriptive analysis focusing on passive clauses in Indonesian, particularly examining whether the clauses under study belong to the canonical or non-canonical forms of passivization.

Data Collection

The data were gathered from the Leipzig Corpora Collection on Indonesian, ensuring the inclusion of naturally occurring expressions and providing authentic data for analysis (Hasko, 2013). The primary goal of this data collection is to analyze the natural usage of passive clauses in Indonesian, enabling an accurate classification of these clauses as either canonical or non-canonical.

However, for the purposes of grammaticality judgment diagnostics and tests, example sentences were fabricated or constructed. This is because the passive counterparts of active clauses found in the Leipzig Corpora so far did not include the relevant sentences needed for the analysis.

Theoretical Framework

The theoretical framework employed in this research is based on the concept of canonicity in passive structures, as theorized by Legate (2021). This framework offers a comprehensive approach to understanding the variations and usage of passive clauses in the Indonesian language. By applying Legate's theory, the study aims to contribute to the broader understanding of passive constructions and their canonical status in within linguistic research in general and specifically in the context of Indonesian.

III. RESULTS

Data collected from the Leipzig Corpora on Indonesian uncovers the presence of two distinct types of passive clauses in the language: the canonical passive and the non-canonical passive. This differentiation becomes apparent through the three markings on the verbs, as illustrated in Table 1.

Table 1 Active-passive Voice Markings on the Indonesian Verbs

No	Type of Verbs	Markings on the Verb
1	Active Verbs	<i>meN</i> -X and \emptyset -X
2	Passive Verb	<i>di</i> -X, \emptyset -X, and <i>ter</i> -X

(Where *meN*-, \emptyset -, *di*-, and *ter*- are diathesis prefixes while X is the verb root)

Table 1 shows that \emptyset -X lacks any markings, presenting it in two forms: active and passive. These forms are specifically denoted as bare active (BA) and bare passive (BP), respectively, according to Nomoto (2018, 2021). This is consistent with Voskuil's (2000) claim on the voice forms of Indonesian, though the glossing is different.

IV. DISCUSSION

Diathesis Forms: *meN*- and \emptyset - Forms

Before discussing the distinction between canonical and non-canonical passive clauses in Indonesian, it is essential to first examine the two active diathesis prefixes: *meN*- and \emptyset . The *meN*- prefix is designated as an AV (agentive-focus)

marker, consistent with a similar phenomenon observed in both Indonesian and Balinese. For example, Arka and Manning (1998) discuss this phenomenon in the context of Indonesian, while Arka (1998, 2003) and Udayana (2013) explore its application in Balinese.

meN- form (Leipzig corpora):

- (1) a. Dia men-cari jalan menuju
3SG AV-look.for way to
tangga eskalator
stair escalator
'He looked for the way to the escalator stairs'
- b. Hampir semua orang
3SG all person
meny(s)ukai puding
AV-like pudding
'Almost everyone likes pudding'
- c. Mereka men-(t)anam padi,
3PL AV-plant rice
'They planted rice,'

As implied by its name, the \emptyset - form is expressed as the empty counterpart of the *meN*-form, as illustrated in (2) with the verbs in bold. The zero form is glossed as Bare Active (BA) following Nomoto (2018, 2021). It is important to note that the \emptyset form shares the same structure as the Bare Passive (BV) form, but the distinction lies in the subject function it co-occurs with, specifically in relation to either an agent or a theme role. This distinction becomes evident in cases involving non-canonical passive constructions.

\emptyset - form (Leipzig Corpora):

- (2) a. istri cari anak muda
wife BA-look.for man young
'The wife was looking for a young man'
- b. Kita tidak **sukai**
1PL.INCL NEG BA.like
kemaksiatan=nya
immorality=3POSS
'We don't like his immorality'
- c. Makanya saya **tanam** di sini
so 1SG BA.plant here
pohon tin dan zaitun
tree fig and olive
'So I planted fig and olive tress here'

Canonical Passive

Passive constructions can be analyzed through verb morphology (Legate, 2021; Haspelmath, 1990; Haspelmath & Sims, 2021). In this context, Legate

(2021) identifies three properties as characteristic of canonical passives.

(3) Canonical Passive:

- (a) Agent demotion: The agent is present in meaning but not as a noun phrase in its typical syntactic position. Instead, the agent is either implied as ‘someone’ or included in a ‘by’-phrase.
- (b) Theme promotion: The theme is promoted from its lower syntactic position, where it is interpreted as a theme, to the grammatical subject position.
- (c) Morphological marking: The verb morphology is distinct from that of the active voice.

Di- passive

The *di-* passive meets the criteria required by the principles outlined above. First, the agent, is relegated to an oblique function and simultaneously serves as an adjunct (a non-argument function), which can be syntactically omitted. Second, the theme is promoted from an object function to a subject function. Third, the verbal morphology changes from the *meN-* form to the *di-* form. Thus, the passive counterparts of the sentences in (1a) can be rewritten as (4).

- (4) a. Jalan menuju tangga eskalator
way to stair escalator
cari oleh dia
PASS.look.for by 3SG
‘The way to the escalator stairs was searched for by him/her’
- b. Puding hampir disukai oleh
pudding almost PAS.like by
semua orang
all person
‘Pudding is liked by most people’
- c. Padi ditanam oleh mereka
rice PASS.plant by 3PL
‘Rice was planted by them.’

Sentence in (4a), for example, has been transformed into a passive clause. The verb marker *meN-* in the active clause is replaced by the marker *di-*. The subject of the active clause, *dia* ‘(s)he,’ which has the semantic role of agent, now occupies the oblique adjunct position. The object of the active clause, *jalan menuju tangga eskalator* ‘the way leading to the escalator stairs,’ which has the semantic role of theme, now occupies the grammatical function of the subject.

The agent remains an agent but is demoted and no longer written as a verb argument in the passive clause. Thus, sentence (4a) is a canonical passive. The remaining clauses (4b-c) also share this characteristic, being canonical passive clauses.

The fact that the agent in the passive clause no longer serves as an argument of the verbal predicate is evident because the *by*-phrase can be omitted. Thus, sentence (4a), for example, can be rewritten as in (5).

- (5) Jalan menuju tangga eskalator
way to stair escalator
di-cari
PASS.search
‘The way to the escalator stairs was searched’

Although the agent in a passive clause is no longer an argument, as indicated by the deletion of the *by*-phrase in the surface syntax, semantically it can still be interpreted existentially as ‘somebody’ (Legate, 2021). Additionally, the existence of the agent argument associated with the event is implied. Thus, when the short passive is embedded within a purposive clause, the agent of the purposive clause is semantically related to the agent of the matrix clause, which points to the agent of the event. In other words, the agent of the verbal predicate associated with the purposive clause is controlled by the implied agent of the passive clause. Sentence (5) can be represented as in (6) to show the embedded purposive clause.

- (6) Jalan_i menuju tangga eskalator
way to stair escalator
di-cari [PRO *i/j untuk
PASS.search to
men-capai lantai dua dengan
AV-reach floor two with
cepat]
quick
‘The way to the escalator stairs was searched to get to the second floor quickly’

In (6), the agent cannot be linked to the NP subject of the main clause, even though the subject is human. Instead, it must be connected to the implicit agent indexed as *j*, the subject of the active clause counterpart.

The absence of an agent in sentence (5) can be demonstrated by the insertion of an agent-oriented adverbial such as *dengan sengaja* ‘intentionally’. However, it becomes ungrammatical if the adverbial

is realized as a non-agent-oriented adverbial, as shown in (7).

- (7) Jalan menuju tangga eskalator
 way to stair escalator
 di-cari dengan sengaja/*
 PASS.search with intentional
 dengan tidak sengaja
 with NEG intentional
 'The way to the escalator stairs was
 intentionally/ *unintentionally
 searched'

Non-Canonical Passive

The characteristics of non-canonical passive clauses are automatically the opposite of the characteristics of canonical passive clauses. Legate claims natural language may exhibit a passive is categorized as non-canonical if it does not fulfill either one or two of the properties of the canonical passives postulated in (3).

Generally, we can say that if a language has two passives, one must be canonical and the other non-canonical. Additionally, if a passive form splits into two different types, one must be canonical and the other non-canonical. Indonesian broadly possesses three passive forms: the *di-* passive, the *ter-* passive, and the bare passive. The last two belong to the non-canonical category. As we will show, the *di-* passive splits into two subtypes: the *di-* passive with a demoted agent and the *di-* passive with a non-demoted agent. According to Legate's theory, the first subtype is canonical, while the latter is non-canonical. As depicted in Table 1, Indonesian ultimately has three forms of non-canonical passives: the *ter-* passive, the bare passive, and the non-demoted agent *di-* passive. In the following, let us examine each of these three non-canonical passives in turn.

Ter- passive

As shown earlier, passives are a derived construction. They are related to their active equivalents. Intuitively, the *ter-* passive may be derived from the active form or from the associated *meN-*. The sentence in (8) can then be transformed into either the *di-* passive or the *ter-* passive, depending on our requirements and the context. If our choice is the former, the resulting construction is a canonical passive. If the option depends on the latter, a non-canonical passive will result. Thus, the sentence in (8) has its passive counterpart as in (9)

- (8) Mereka men-(t)anam padi,
 3PL AV-plant rice
 'They planted rice,'

- (9) Padi ter-tanam oleh mereka
 rice PASS.plant by 3PL
 'Rice got planted by them'

Like the *di-* passive, the agent denoted by the by-phrase can be dropped in the *ter-* passive. However, we have argued that while the *di-* passive is canonical, the *ter-* passive is non-canonical. This raises the question of why the *ter-* passive is considered non-canonical.

According to its semantic criterion, the *ter-* passive is associated with unintentional actions, in contrast to its *di-* passive equivalent. Therefore, if the non-canonical passive clause in (10) is combined with an adverbial of intention, the resulting clause becomes ungrammatical (Reed 2011; Alexiadou 2012; Alexiadou & Schäfer 2013; Fox & Grodzinsky 1989).

- (10) * Padi ter-tanam oleh mereka
 rice PAS.plant by 3PL
 dengan sengaja
 with Intentional
 'Rice got planted by them
 intentionally'

Sentence (10) is ungrammatical because the action is performed unintentionally, and the agent is considered absent at a certain level of representation. Therefore, unlike the *di-* passive, the agent in the *ter-* passive is perceived as lacking control in a purposive clause, as shown in (11). This aligns with the *get* passive being a non canonical cannot control into a purposive clause (see Fox & Grodzinsky 1998: 327 as cited in Alexiadou & Schäfer 2013: 6).

- (11) Padi_i ter-tanam oleh mereka_j
 rice PAS.plant by 3PL
 [Pro _{*i/*j} untuk men-dapat.kan
 to AV-get.CAUS
 hasil yang banyak]
 'Rice got planted by them to achieve
 large yields'

With the agent unable to control the event named by the predicate in the purposive clause, it suggests that the purposive clause must be self-contained, as illustrated in (12).

- (12) Padi ter-tanam untuk orang
 rice PASS.plant by person
 desa itu men-dapat.kan
 village to AV-get.CAUS
 hasil yang banyak
 'The rice got planted for the
 villagers to achieve large yields.'

Another test to show that the agent is not intentionally involved in the action has to do with the inability for the subject of the clause to refer to overt reflexive anaphor (Fox & Grodzinsky 1989). Again this runs counter to the situation with the *di*-passive. This is illustrated in the following contrast.

- (13) a* Makanan ter-hidang-kan untuk
 food PAS-serve-CAUS for
 diri sendiri
 self oneself
 'The food got served for oneself.'
 b. Makanan di-hidang-kan untuk
 food PAS-serve-CAUS for
 diri sendiri
 self oneself
 'The food got served for oneself.'

The inability of the *ter*-passive to combine with reflexive pronouns clearly indicates that the agent of the action is not implicit and, in some sense, is either non-existent or absent. The intuition with this is that the *ter* passive often implies some level of agentivity or control on the part of the subject. For instance, in the sentence *John menyebabkan dirinya tertembak* "John got himself fired," *John* is perceived as having some role in causing the action. However, this agentivity can complicate the relationship between the subject and a reflexive anaphor, making it less compatible because the reflexive anaphor typically requires the subject to be purely the recipient of the action, not the instigator. In contrast, the *di*-passive typically lacks this agentive implication. It presents the subject more straightforwardly as the recipient of the action, which aligns well with the use of reflexive anaphors. For example, *John ditembak oleh dirinya sendiri* "John was fired by himself" is more semantically neutral and doesn't imply that John actively caused his firing, making it compatible with the reflexive.

A characteristic tied to the *ter*-passive, being a non-canonical passive, is that it pertains to achievement, not to accomplishment (which aligns with the *di*-passive). This is similar to the *get* passive in English (Alexiadou 2011, Alexiadou

& Schafer 2012, Reed 2011). Thus, compare the following contrast:

- (14) a* Padi itu ter-tanam selama
 rice DEM PASS-plant for
 limabelas menit
 fifteen minute
 'The rice got planted for fifteen
 minutes'
 b Padi itu di-tanam selama
 rice DEM PASS-plant for
 limabelas menit
 fifteen minutes
 'The rice was planted for fifteen
 minutes'

The *ter*-passive, which is linked to achievements, is incompatible with adverbials of duration, reflecting its deviation from the typical non-canonical passive.

The non-deviational nature of the *ter*-passive is evident in its interpretation as an adjectival form, as seen in (15a). However, in (15b), replacing it with the *di*-form would not yield a semantically acceptable result.

- (15) a. Sangat positif jika mereka
 very positive if 3PL
 tertarik kepada Solomon
 PAS.attract to name
 'Very positive if they are interested
 in Solomon'
 b.* Sangat positif jika mereka
 very positive if 3PL
 ditarik kepada Solomon
 PAS.attract to name

In Indonesian, there is a close relationship between the *ter*-passive and anticausatives. Consider the following examples:

- (16) a Pendukung Ali terpecah
 supporter name PASS.break
 oleh dua kubu
 by two camp
 'Ali's supporters were divided by
 two camps'
 b Pendukung Ali ter-pecah
 supporter name PASS-break
 menjadi dua Yaitu, Syi'ah
 become two that is name
 dan Khawarij
 and name
 'Ali supporters were divided into
 two: Syi'ah and Khawarij'

In many languages, anticausative and inchoative constructions often overlap. Both describe situations where a subject undergoes

a change of state, and in many cases, the same verb form can be used in both constructions. For example, “The ice melted” can be seen as both anticausative (no external agent is mentioned) and inchoative (the process of melting began). Inchoative is commonly expressed in Indonesian with *ber-* form. Thus (16b) can be rewritten as. *Pendukung Ali berpecah menjadi dua, yaitu: Syiah dan Khawarij*. The *ber-* form in turn associates more with the middle voice in Indonesian (see Beavers & Udayana (2023)). Another property worth mentioning here with respect to the *ter-* form or the *ter-* passive is that the *ter-* form has the same form as superlative adjective marker Indonesian.

This property is compatible with the perfective verb form in English in that the *ter-* form in Indonesian complies with adjectival passive in English (17).

- (17) a. Malam ini aku bisa
night DEM 1SG AUX
tuliskan baris-baris tersedih
because line_RED saddest
‘Tonight, I can write the saddest lines’
- b. Karina pun tertarik pada
name PART PAS.interest in
Rudi
name
‘Karina is also interested in Rudi’

The *ter-* form that is associated with the adjectival form status can manifest in the *ter-* form indicating different category of a modifier, a definite marker. There is the only one particular example that we find in Indonesian, the form involving the base verb *sebut* ‘mention’. (18a) *disebut* is a canonical passive making the use of the base verb grammatical. An attempt to replace the *di-* form with the *ter-* form results in an unacceptable clause (18b).

- (18) a. Itu semua zat kimia
that all substance chemical
karena ada rumus
because exist formula
kimia=nya maka disebut
chemistry=DEF so PASS.call
Biokimia
Biochemistry
‘These are all chemical substances, because they have chemical formulas, they are called biochemistry’

- b.* Itu semua zat kimia
that all substance chemical
karena ada rumus
because exist formula
kimia=nya maka ter-sebut
chemistry=DEF so PASS.call
biokimia
biochemistry
- (19) a. Orang yang di/tersebut
person REL PAS/PAS.call
kemudian adalah temannya
later COP friend.3POSS
‘The person called later is his friend’
- b. Orang *disebut /tersebut
person PAS.call/PAS.call
adalah temannya
COP friend.3POSS
‘The person mentioned is his friend’

In (19a), also refers to the fact that the verb base itself has no adjectival status, so both the *ter-* passive and the *di-* passive are allowable. However, (19b) has adjectival manifestation the use of the *di-* form is ungrammatical while the *ter-* form is impeccable. Importantly, the *ter-* form in (19b) is lexicalized as a definite article marker.

Bare Passive

Bare passive as the name suggests relates to the fact that the verb under discussion has no marking on the verb. and because of this characteristic the verb so used is glossed as bare passive BP verb (Nomoto 2021). The usage of this verb can be illustrated as in the following examples:

- (20) a. Sebagai gantinya, Yenny
as replacement name
saya berikan soto ayam
1SG BP.give soup chicken
yang ada di depan kantor
REL exist in front office
‘Instead, Yenny was given the chicken soup (that was sold) in front of the office by me’
- b. Impian berpolitik sudah
dream BER.politik PERF
Tuhan berikan lebih dari
God BP.give more than
yang Iman impikan
REL name dream.CAUS
‘The dream of being involved in politics has been granted by God beyond what Iman had ever dreamed of’

The debate over the bare form verb, as exemplified in (20), has become a significant

topic in Indonesian linguistics. Unlike in English, where topicalization does not alter the voice of the clause—as seen in sentences like ‘I like pizza’ and ‘Pizza I like’—the situation is different in Indonesian. Those who argue that the movement of the object in Indonesian is a form of topicalization, similar to English, should consider the following contrastive sentences, which illustrate what is known as a symmetrical voice system (See Riesberg 2014, Himmelmann & Riesberg 2013, Riesberg & Primus 2015 for more information on voice symmetry)

- (21) a. Yenny saya berikan soto
 Name 1SG BP.give soup
 ayam
 Chicken
 ‘I gave Yenny the chicken soup’
 b. Saya berikan Yenny soto
 1SG BA.give name Soup
 ayam
 chicken
 ‘I gave Yenny the chicken soup’

The examples provided in (21) do not clarify whether the same form is associated with both the bare active and bare passive voices. In English, the translation of these forms can often remain identical, typically resulting in an active interpretation. For instance, even though (21b) is marked with an AV (Actor Voice) prefix, as in *Saya memberikan Yenny soto ayam*, the structure in (21b) can still be interpreted semantically as an active clause. However, Sneddon *et al.* (2010), as noted, argues that a sentence like (21a) actually constitutes a passive construction in Indonesian, specifically what he terms ‘passive type 2.’ In contrast, the *di-* form, as in *Saya diberikan Yenny soto ayam*, is categorized as ‘passive type 1. Given this situation, the challenge now is to provide evidence that the bare form, as in (21a), is indeed passive. At first glance, it is clear that the bare form can function as either a bare active or bare passive, as noted by Nomoto (2021). When the bare form is combined with the AV form, it further reveals an ambiguity between an active and a passive interpretation. Consider the following contrast:

- (22) a. Jono dan Tono suka
 name and name like
 pukul-memukul
 BA.hit-AV.hit
 ‘Jono and Tono like hitting people
 (i.e. they are body-guards)’
 b. Jono dan Tono suka
 name like BA.hit-AV.hit
 pukul-memukul Ali
 BP.hit-AV.hit
 ‘Jono and Tono like hitting Ali’

- b. Jono dan Tono suka
 name like BA.hit-AV.hit
 pukul-memukul
 BP.hit-AV.hit
 ‘Jono and Tono like hitting each other’

The similarity between the two clauses is twofold. First, both appear to be intransitive. In (22a), the clause seems intransitive because the object is understood generically, implying that the associated object refers to people in general—specifically, Jono and Tono, who are both assumed to be bodyguards. In (22b), however, the intransitivity of the verb form is specifically tied to the reciprocal construction. Second, the distinction between the bare forms in each clause is clear: (22a) has a bare active construal, whereas (22b) has a bare passive construal. This raises the question of why the form *pukul* in (22b) belongs to the passive diathesis.

The evidence that ‘pukul’ in (22b) relates to a bare passive construction is that, if the same verb form were reduplicated in an active-active manner, the resulting sentences would not yield a reciprocal interpretation. This is demonstrated by the ungrammaticality of the resulting sentence.

- (23) a.* Jono dan Tono suka
 name and name like
 pukul-pukul
 BA.hit-AV.hit
 *‘Jono and Tono like hitting each other’
 b.* Jono dan Tono suka
 Name and name like
 memukul-mukul
 BP.hit-AV.hit
 *‘Jono and Tono like hitting each other’

If (23a-b) were associated with a dedicated reduplicated form to indicate the plurality of the action, the verbs in question would have to take an overt object noun phrase (NP). This means that the reduplication not only marks the plurality of the action but also necessitates the presence of an explicit object to maintain grammaticality.

- (24) a. Jono dan Tono suka
 name and name like
 pukul-pukul Ali
 BA.hit-AV.hit
 ‘Jono and Tono like hitting Ali’
 b. Jono dan Tono suka
 name like BA.hit-AV.hit
 memukul-mukul Ali
 BP.hit-AV.hit
 ‘Jono and Tono like hitting Ali’

That the bare form in (25) relates to passive diathesis (pertaining to reciprocation) is well supported by Old Javanese, as illustrated by the examples in (26).

- (26) a. Arek - arek iku tulung-
 name and name help
 t-in-ulung
 <PASS> help
 'The children helped each other'
- b. Bapak karo ibuk=ne
 father and mother=DEF
 sawang-sinawang
 BA.look-<PASS>look
 'Mother and father looked at each other'

The reduplication of the verb with the same root but different in voice marking (passive-active) is strategized by Indonesian to adopt a reciprocal construction. This seems to be a language-particular phenomenon. In other natural languages, a clause such as the one in (27) the reduplication of the same active form yields a reciprocal construction as exemplified by Godié (Kru, Niger-Kongo, Côte d'Ivoire) (by Marchese 1986:231) found in Maslova & Nedjalkov (2005).

- (27) wa wà-wà
 they love-love
 They love each other'

Di- passive + Undemoted Agent

As discussed earlier, verb forms marked with *di-* combined with an optional *by*-phrase are typically identified as canonical passives. However, the *di-* form can also indicate non-canonical passive constructions, particularly when the agent is not demoted to an oblique position. This deviation from the canonical pattern occurs when the *di-* form retains the agent in its original syntactic position, as illustrated in the following examples:

- (28) a. Seluruh hasil gambir,
 all product gambir
 dibeli pemerintah
 PASS.buy government
 'All the gambir products were bought by the government'
- b. Dalam isi pesan yang
 in content message REL
 dikirim Devi itu dia
 PASS.send name that 3SG
 merenungkan kesedihan=nya
 AV.pour.CAUS sadness=3POSS
 'In the message sent by Devi, she expressed her sadness'.

At first glance, the constructions in (28) resemble the bare passive in (20), which might lead us to assume they exhibit the effects of the symmetrical voice system, where the agent is retained in *di-* constructions, making them look like active clauses. However, both constructions share a key feature: they satisfy the requirements for a passive construction by foregrounding the patient of the previously associated active clauses. This foregrounding links both constructions to a passive operation, even though again the agent is retained. Consequently, Legate (2021) refers to these as non-canonical passives.

The second interpretation that may link the bare passive to the *di-* form with an undemoted agent is its information packaging effect, particularly with first and second person subjects. Recall that Sneddon *et al.* (2010) claims the *di-* passive is prohibited with first and second person agents, except when the agent is undemoted, as in *buku saya beli* and *buku dibeli oleh saya*. In these cases, the first construction is allowed, while the second is not, leading to what he calls passive type 2. Thus, *buku dibeli saya* is preferred over *buku itu dibeli oleh saya*. The same principle applies with second person subjects.

V. CONCLUSION

In Indonesian, the passive voice is categorized into different statuses. The *di-* passive, where the agent is expressed as an oblique argument, represents the canonical passive form. This form clearly indicates a typical passive construction. However, when the *di-* form is used with an agent that cannot be expressed as the object of a preposition (and thus cannot be omitted), it constitutes a non-canonical passive.

The non-canonical passive forms also include the *ter-* form and the bare passive form. The bare form appears in two ways: it can indicate an active verb construction where the object NP (noun phrase) occupies the post-verbal syntactic position, or it can involve the agent NP immediately preceding the verb. In the latter case, the verb remains unmarked by any prefix, reflecting a bare passive construction.

ETHICS STATEMENT

The authors have read and followed the ethical requirements for publication in Data in Brief and that the current work does not involve human

subjects, animal experiments, or any data collected from social media platforms.

CREDIT AUTHOR STATEMENT

I Nyoman Udayana: Draft the main sections of the manuscript, including introduction, methods, results, and discussion, incorporate feedback from co-authors and external reviewers, handle the submission process and respond to reviewers' comments; **I Nyoman Aryawibawa:** Conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, writing- original draft preparation, reviewing and editing; **I Made**

Madia: Contribute specific expertise having to do with theoretical framework, literature review, or specialized knowledge related to the study, write or enhance the background, theoretical framework, and conclusion). Review the manuscript for technical accuracy and provide additional references or resources.

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

- Appiah, K. R., Ankomah, C., Osei, H. Y., & Hattoh-Ahiaduvor, T. (2019). Structural organisation of research article titles: A comparative study of titles of business, gynaecology and law. *Advances in Language and Literary Studies*, 10(3), 145-154. <http://dx.doi.org/10.7575/aiall.v.10n.3p.145>
- Alexiadou, Artemis & Doron, Edit. (2012). The syntactic construction of two non-active Voices: passive and middle. *Journal of Linguistics* 48(1). 1–34.
- Alexiadou, A. (2012). 'Noncanonical passives revisited: Parameters of Nonactive Voice'. *Linguistics*: 50 (6) 1079-1110. DOI: 10.1515/ling-2012-0036
- Alexiadou, A. & Schäfer, F. (eds). (2013). *Non-canonical passives*. Amsterdam: John Benjamins Publishing Company.
- Arka, I Wayan. (1988). From Morphosyntax to Pragmatics in Balinese. Sydney: Sidney University Ph.D. Dissertation.
- Arka, I Wayan. (2003). *Balinese Morphosyntax: A Lexical Functional Approach*. Canberra: Pacific Linguistics.
- Arka, I Wayan & Manning, C. D. (1988). Voice and Grammatical Relations in Indonesian: A New Perspective in *Proceedings of International Lexical-Functional Grammar Conference (LFG 98)*. M. Butt & T. Holloaway King (eds.). Stanford, CA: Central Study of Language and Information.
- Arka, I Wayan & Kosmas, Jeladu. (2005). Passive without Passive Morphology? Evidence from Manggarai in *the Many Faces of Austronesian Voice Systems: Some new empirical studies*. Arka, I Wayan and Malcolm Ross (ed.), pp 87-117. Pacific Linguistics.
- Beavers, John & Udayana, I Nyoman. (2023). Middle voice as generalized argument suppression: The case from Indonesian. *Natural Language and Linguistic Theory*, 41(1): 1-52. DOI: 10.1007/s11049-022-09542-5
- Cole, Peter, Hermon, Gabriella, & Tjung, Yassir. (2006). Is there Pasif Semu in Jakarta Indonesian? *Oceanic Linguistics* 45 (1): 64-90. DOI:10.1353/ol.2006.0009
- Chung, Sandra. (1976). An Object Creating Rule in Bahasa Indonesia in *Linguistic Inquiry*, 7 (1): 41 -87.
- Fox, Danny & Grodzinsky, Yosef. (1998). Children's passive: A view from the *by*-phrase. *Linguistic Inquiry* 29(2). 311–332.
- Hasko, Victoria. (2013). Qualitative Corpus Analysis in the Encyclopedia of Applied Linguistics. Chapell, Carol A (ed). Oxford: Wiley-Blackwell.
- Haspelmath, Martin. (1990). The Grammaticization of Passive Morphology. *Studies in Language*, 14 (1): 25 – 72.
- Haspelmath, Martin & Sims, Andrea D. (2010) (2nd Edition). *Understanding Morphology*. London: Hodder Education, an Hachette UK Company.
- Huddleston, Rodney & Pullum, Geoffrey. (2005). *A Student's Introduction to English Grammar*. Cambridge: Cambridge University Press.
- Himmelmann, Nikolaus P. & Riesberg, Sonja. (2013). Symmetrical Voice and Applicative Alternations: Evidence

- from Totoli. *Oceanic Linguistics*. 52 (2): 396-422. DOI: 10.1353/ol.2013.0021
- Legate, J. A. (2012). Subjects in Acehnese and the nature of the passive. *Language* 88 (3): 495–525
- Legate, J. A. (2021). Noncanonical Passives: A Typology of Voices in an Impoverished Universal Grammar in *Annual Review of Linguistics*, 7: 7.1 – 7.20
- Lyngfelt, Benjamin & Solstad, Torgrim (eds.). (2006). *Demoting the Agent: Passive, Middle, and Other Voice Phenomena*. Amsterdam: John Benjamins Publishing Company. 17
- Maslova, Elena & Nedjalkov, Vladimir P. (2005). *Reciprocal Constructions* in The World Atlas of Language Structures, Haspelamth Marin, Matthew S. Dryer, David Gil, and Bernard Comrie (eds.). Oxford: Oxford University Press.
- Nomoto, Hiroki (unpublished). (2021). Overlooked bare voice constructions in Westren Austronesian in *NUSA*, 71: 1-18.
- Nomoto, Hiroki. (2018). The Development of the English-Type Passive in Balinese. *Wacana*, 19 (1): 149 – 167.
- Reed, L. A. (2011). Get-passives in *The Linguistic Review* (28): 41-78. DOI: 10.1515/tlir.2011.022
- Riesberg, Sonja. (2014). *Symmetrical voice and linking in Western Austronesian languages*. Berlin: Mouton De Gruyter.
- Riesberg, S. & Primus, B. (2015). Agent Prominence in Symmetrical Voice Language. *STUF - Language Typology and Universals* 68 (4): 551-564. <https://doi.org/10.1515/stuf-2015-0023>
- Shibatani, Masayoshi. (1988). *Passive and Voice*. Amsterdam: John Benjamins Publishing Company.
- Sneddon, James Neil, Adelaar, Alexander, Djenar, Dwi Noverni, & Ewing, Michael C. (2010). *Indonesian: A Comprehensive Grammar*. London: Routledge.
- Udayana, I Nyoman. (2013). *Voice and Reflexives in Balinese*. Ph.D. Dissertation the University of Texas at Austin.
- Udayana, I Nyoman (2022). The Indonesian *di-* passives and discourse context. *Journal of Language Teaching and Research*, 13 (6): 1304 - 1312.
- Voskuil, J.E. (2000). Indonesian Voice and A-Bar Movement. In Paul, Ileana, Vivianne Phillips, and Lisa Travis (eds.). *Formal Issues in Austronesian Linguistics*, 195-212. Dordrecht: Kluwer.