Prescriptive grammars of Malay/Indonesian state that the agent in \textit{di}-passives should be third person and prohibit first and second person agents. Researchers are not unanimous as to whether this statement is descriptively accurate. This paper examines various texts in Formal and Colloquial Malay, and shows that the restriction exists as a strong tendency rather than an absolute syntactic rule. The paper accounts for the restriction in terms of information structure and syntax, by claiming that \textit{di}-passive agents cannot be salient/given (first and second person) because the eventuality described by a \textit{di}-verb is non-salient/new. We link the low salience of \textit{di}-verbs to the \textit{v}-to-\textit{V} movement involved in \textit{di}-passives. This movement does not occur in another kind of passives, i.e. bare passives, whose agents are not subject to the person restriction.

1. Introduction

Malay has two types of passives: \textit{di-} (morphological) passives (1a) and bare passives (1b). They are so called based on their surface morphological characteristics. The verb bears the overt passive voice marker \textit{di-} in the former whereas it bears no overt voice marker in the latter.\footnote{Bare passives are referred to by various names in the literature: ‘object-preposing construction’ (Chung 1976b; Willett 1993), ‘Passive Type 2’ (Dardjowidjojo 1978; Sneddon et al. 2010), ‘pasif semu’ [pseudo-passive] (Asmah 2009), ‘object(ive) voice’ (Arka & Manning 1998; Cole, Hermon & Yanti 2008), and so forth. See Nomoto (2006) for a summary of various existing terms.} Besides this morphological difference, the two passives also differ in the status of the agent. The agent in \textit{di-}passives appears to be optional, and hence an adjunct (but see section 4.3 for a different view). By contrast, the agent in bare passives is an obligatory argument and immediately precedes the verb.

\begin{enumerate}
\item \textit{Di-}passive
\begin{align*}
\text{Dokumen itu sudah \textit{di}-semak oleh mereka.} \\
\text{document that already \textit{PASS}-check by them}
\end{align*}
\begin{itemize}
\item ‘The document has already been checked by them.’
\end{itemize}
\item Bare passive
\begin{align*}
\text{Dokumen itu sudah *(merekA) semak.} \\
\text{document that already they \textit{check}}
\end{align*}
\begin{itemize}
\item ‘They have already checked the document.’
\end{itemize}
\end{enumerate}

Of these two passive types, this paper focusses on \textit{di-} passives, and in particular, on the person restriction on the agent. The agent of \textit{di-} passives are restricted to third person in

\footnote{This study was supported in part by the JSPS Grant-in-Aid for Young Scientists (B) (#23720199, #26770135) awarded to Nomoto. This paper is based on our presentation at the 23rd Annual Meeting of Southeast Asian Linguistics Society (SEALS), 29–31 May 2013, Chulalongkorn University, Bangkok, Thailand. We thank the audience at the conference for their helpful comments. We also thank the three NUSA reviewers for their valuable and constructive comments. Any remaining errors are ours.}
Di-passive agents are encoded in three ways, as shown in (2). In the ‘oleh type’ (2a), the agent is introduced by the preposition *oleh* ‘by’. In the ‘DP type’ (2b), the agent immediately follows the verb, with no preposition between them. Lastly, in the ‘pro type’ (2c), no overt agent occurs, though the presence of an agent is still entailed.

(2) a. *Oleh* type

_Surat itu sudah di-poskan oleh kerani._
letter that already PASS-post by clerk

b. *DP* type

_Surat itu sudah di-poskan kerani._
letter that already PASS-post clerk

c. *Pro* type

_Surat itu sudah di-poskan pro._
letter that already PASS-post

‘The letter was already posted (by the clerk).’

Prescriptive grammars (e.g. Nik Safiah et al. 2008) impose a person restriction on di-passive agents, which can be summarized as follows:

(3) Person restriction on di-passive agents

\[ \text{di-}V \text{ Agent}_1/2/3 \]

“The agent must be third person. No first and second person agents are allowed.”

Descriptive grammars are not unanimous with respect to this person restriction. Some consider that the restriction is descriptively accurate (e.g. Arka & Manning 1998; Donohue 2007; Sneddon et al. 2010), while others claim that the restriction is only a prescriptive rule and is descriptively inadequate (e.g. Asmah & Rama 1968; Chung 1976a; Lufti 1985; Mintz 2002; Abdullah 2006). Guilfoyle, Hung & Travis (1992) argue that speakers differ in the presence of the relevant restriction.

The questions that naturally arise are these: Is the person restriction in (3) descriptively accurate? If so, to what extent? These questions are important because the answers to them delimit the possible analyses of the syntax of di-passives. This paper attempts to answer these questions by examining various texts in Standard Formal and Standard Colloquial Malay as they are used in Malaysia.3 We argue that the restriction is descriptively accurate, but is not as strict as the formulation in (3). A natural question to ask then is: Why does such a restriction exist?

Our data consists of the following five sources:

(4) a. The magazine (*majalah*) subcorpus of the DBP Corpus4

b. Front page articles of *Utusan Malaysia* for the entire period of the year 2011

3 These Malay varieties should not be confused with the prescriptive variety, which we refer to as ‘Standardized Malay’ (*bahasa Melayu baku*) as opposed to ‘Standard Malay’ (*bahasa Melayu standard*).

c. A collection of folktales (*cerita rakyat*)

d. The Multilingual Corpora (Malay)

e. Miscellaneous (e.g. web pages)

The folktales in (4c) used to be available to download at the DBP website, but they are no longer available now. Table 1 gives information about the types and sizes of the four corpora. Utilizing multiple data sources ensures that our data is large and representative enough, and contains both Formal and Colloquial Malay sentences.

Table 1. The data for this study

<table>
<thead>
<tr>
<th>Texts Data type</th>
<th>Token (words)</th>
<th>Type (words)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBP (4a) formal</td>
<td>14,406,888</td>
<td>(unknown)7</td>
</tr>
<tr>
<td><em>Utusan Malaysia</em> (4b) formal</td>
<td>501,272</td>
<td>17,728</td>
</tr>
<tr>
<td>Folktales (4c) formal~colloquial</td>
<td>66,711</td>
<td>6,263</td>
</tr>
<tr>
<td>Multilingual Corpora (4d) colloquial</td>
<td>232,374</td>
<td>9,950</td>
</tr>
</tbody>
</table>

This paper is structured as follows. Section 2 discusses the first study, which deals with the *oleh* type. We show that *di-* passives with first and second person agents are attested in our data. Section 3 presents the second study, which deals with all types of *di-* passives. We show the frequencies of first and second person agents. Section 4 discusses the “why” question above. We propose an analysis of the person restriction based on information structure and its relation to syntax. Section 5 is the conclusion.

2. Study 1: Concordance

2.1 Method

In Study 1, we examined whether *di-* passives with first and second person agents actually exist. We only considered the *oleh* type, due to the first of the limitations discussed below. The data we used is the magazine subcorpus of the DBP corpus. We searched for the 14 different strings listed in (5), using the online concordance system. These strings differ from one another in the pronoun used.

(5)  
\[
\begin{array}{ll}
& \text{di oleh saya (1SG)} \\
& \text{di oleh aku (1SG)} \\
& \text{di oleh ku (1SG) [ku: the enclitic form of aku]} \\
& \text{di olehku (1SG)} \\
& \text{di oleh kami (1EXCL)} \\
\end{array}
\]

---


6 We thank Khazriyati Salehuddin for sharing her copy with us.


We are aware that this method has at least two limitations. First, the DBP’s online concordance system only allows one to handle the oleh type. The DP and pro types are practically impossible to handle automatically, as the corpus contains no morphosyntactic annotation and one has to conduct searches for individual verbs as many times as there are verbs in the corpus, e.g. dimakan saya, disemak saya, ditulis saya, dibaca saya, etc. Second, the search results were limited to 100 instances, which include many duplicates. We excluded all duplicates manually. This second limitation makes it difficult to conduct a reliable quantitative analysis, as 100 instances is often too small a number to make any generalization.

2.2 Results

Table 2 shows the results. It is clear that first and second person agents are present in di-passives. Saya (1SG), kami (1EXCL), kita (1INCL) and anda (2) are significantly more frequent than the others due to the genre and register of the corpus used. The corpus consists of magazine articles, which are written for the general public. The pronouns aku/ku/-ku (1SG) and awak, engkau/kau, saudara (2) are normally used in speech and do not appear often in magazine articles.

Representative examples are given below:

(6) Perkara yang paling di-ingati oleh saya ialah kalau apa-apa hal yang berlaku, emak akan memanggil saya .... ‘The thing that is remembered by me most is that if anything happened, my mum would call me ....’

(7) Usia tidak mengampun segala dosa yang di-buat oleh kita. ‘Age does not forgive all the sins that were committed by us.’

---

9 DBP has improved the online concordance system after we conducted the searchers for the present study, and this limitation is no longer present in the current system (02/09/2014). However, a more serious problem exists in the new system: it cannot deal with more than one word, as we did in Study 1.
Table 2. *Di-* passives with first and second person agents in the DBP Corpus

<table>
<thead>
<tr>
<th>Agent</th>
<th>Frequency</th>
<th>Subtotal</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SG</td>
<td>saya</td>
<td>27</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>aku</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ku</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-ku</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>EXCL</td>
<td>kami</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1 INCL</td>
<td>kita</td>
<td>31</td>
<td>149</td>
</tr>
<tr>
<td>2</td>
<td>awak</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>kamu</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mu</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-mu</td>
<td>2</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>engkau</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>kau</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>anda</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td></td>
<td>saudara</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

(8) … *Planta Soft* telah menyediakan segala yang *di-ingini oleh anda iaitu* margarine yang lembut dan sangat lazat. ‘… Planta Soft has prepared everything that is wanted by you, namely margarine that is soft and very tasty.’

The results of Study 1 show that *di-* passives with first and second person agents are actually used in texts. However, two questions remain unsolved. First, while the person restriction in (3) seems inadequate for the *oleh* type, it may be valid for the DP and *pro* types. Are first and second person agents also possible for the latter two types? Second, we should note that the 149 examples of first and second person agents reported above are perhaps negligible when the size of the entire data, comprising some 14 million words, is taken into account. It is therefore necessary to examine how frequent first and second person agents are compared to third person agents in the same text.

3. Study 2: Frequency

To answer these two questions, we examined the frequencies of first, second and third person agents in all three types of *di-* passives.

3.1 Method

We used the following three corpora: Front page articles of *Utusan Malaysia* in 2011 (4b), folktales (4c) and the Multilingual Corpora (Malay) (4d). For each corpus, the first 300 instances of *di-* verbs were picked out. We coded for each item the type of agent encoding ("*oleh*", "*DP*", "*pro*") and the person of the agent ("1st", "2nd", "3rd"). When more than one analysis was possible, the most likely code was assigned.
### 3.2 Results

The results are summarized in Tables 3–5. First and second person agents were attested in all three corpora, but their frequencies are relatively low: 3.7% (*Utusan Malaysia*), 2.0% (Folktale) and 4.7% (Multilingual Corpora). It is worthwhile to point out that all instances of first and second person agents occur in the *pro* type. Tables 3–5 also show that *di-* passives are most frequently used without an overt agent (*pro* type). *Di-* passives in Malay thus fit with Keenan & Dryer’s (2007) cross-linguistic generalization that passives with an overt agent are much less frequent than ones without an overt agent.

#### Table 3. *Di-* passive agents in *Utusan Malaysia* (4b)

<table>
<thead>
<tr>
<th></th>
<th><em>oleh</em></th>
<th><em>DP</em></th>
<th><em>pro</em></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>2nd</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3rd</td>
<td>52</td>
<td>75</td>
<td>162</td>
<td>289</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>75</td>
<td>173</td>
<td>300</td>
</tr>
</tbody>
</table>

#### Table 4. *Di-* passive agents in Folktale (4c)

<table>
<thead>
<tr>
<th></th>
<th><em>oleh</em></th>
<th><em>DP</em></th>
<th><em>pro</em></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2nd</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3rd</td>
<td>60</td>
<td>105</td>
<td>129</td>
<td>294</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>105</td>
<td>135</td>
<td>300</td>
</tr>
</tbody>
</table>

#### Table 5. *Di-* passive agents in Multilingual Corpora (4d)

<table>
<thead>
<tr>
<th></th>
<th><em>oleh</em></th>
<th><em>DP</em></th>
<th><em>pro</em></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>2nd</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3rd</td>
<td>30</td>
<td>17</td>
<td>239</td>
<td>286</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>17</td>
<td>253</td>
<td>300</td>
</tr>
</tbody>
</table>

Representative examples are given below, with necessary commentaries. The *pro* in (9) is the writer of the article in question or ‘the media’ including the writer, and hence is first person. The person who met beliau ‘him’ cannot be the reader (second person) or a third party excluding the writer/reader (third person).
(9) Beliau di-temui pro selepas merasmikan Seminar Pengurusan Sukan Institusi Pengajian Tinggi (IPT) 2010 di UiTM kampus Khazanah Alam Bandar Jengka di sini. ‘He was met by pro after he had officiated the 2010 Higher Academic Institution Sports Management Seminar at UiTM, Khazanah Alam Bandar Jengka campus here.’ (Utusan Malaysia, 01/01/2011)

The pro in (10) is most likely to be the writer (first person), as it is the writer who has seen the websites at issue. The writer may be assuming that his readers can confirm his statement if they also see the same website. Thus, the pro here could also be analysed as the readers (second person) or both the writer and readers (first person). However, it is unlikely that it refers exclusively to third person referents, that is, a third party including neither the writer nor the reader. We classified this example as “1st.”

(10) Jika di-lihat pro blog-blog dan laman sosial popular Facebook turut mendedahkan ‘keberanian’ golongan gay itu berkongsi pengalaman peribadi dan kehidupan mereka sebagai pengamal seks songsang. ‘If looked at by pro, blogs and the popular social site Facebook also expose the “bravery” with which the gay people share personal experiences and their lives as practitioners of sexual perversion.’ (Utusan Malaysia, 02/01/2011)

The sentence in (11) contains a clear example of a di- passive with a second person agent, as it is an imperative sentence. Passive imperatives, where the subject does not coincide with the agent, are not uncommon in literary works in Malay.

(11) Melur, tebu se-batang yang subur di tepi dapur jangan di-tebang pro. ‘Melur, the luxuriant sugar cane at the side of the kitchen mustn’t be cut down by pro. [= Melur, pro don’t cut down the luxuriant sugar cane at the side of the kitchen]’ (Folktale, #9 Kerana Tebu Sebatang)

Example (12) is taken from Multilingual Corpora (Malay), a corpus of Colloquial Malay. The pro here too is second person, because the speaker is asking the hearer to answer his question by doing the action expressed by a passive verb. It could also include the speaker, but cannot refer to only the speaker or a third party excluding the speaker and the hearer.
(12) Payah jugak kan kalau kita belajar bahasa lepas tu masih keliru, hard also right if we study language after that still confused tapi kan (Zu) [erm] kalau hendak di-bezakan pro antara bahasa but right Zu erm if will PASS-differentiate between language Melayu dengan bahasa Jepun /// rasanya mana yang lebih mudah Malai with language Japan supposedly which REL more easy di-pelajar /// di-pelajari?11 PASS-student PASS-learn

‘If you learn a language and still don’t get it, it looks like it’s not that easy. Don’t you think so? But, Zu, erm... if Malay is compared by pro with Japanese... which do you think is easier to learn?’

(Multilingual Corpora (Malay), 01Dec#3-B105)

3.3 Miscellaneous examples

Though not from our corpora, interesting examples of di- passives with first and second person agents were found on a sign prepared by KTM Berhad (Malayan Railway) (Figure 1). The relevant sentence is quoted in (13). The sentence contains two di- passive verbs. The agent is not expressed overtly for either verb, hence the pro type. Pro1 is most likely to be interpreted as first person. The relevant advice is given by the railway company which prepared the sign (first person), but not by the passengers who read it (second person). It is possible but unlikely that a third party that does not write or read the sign (e.g. the police) ordered the railway company to prepare the sign on their behalf (third person), given that the second sentence explicitly conveys the primary commitment of the railway company. Pro2, on the other hand, can only be interpreted as second person, because the items that need attention are the belongings of those for whom the sign is intended, normally the passengers (second person), but not of the railway company (first person) or others who have nothing to do with the sign (third person).

(13) Anda di-nasihatkan pro1 supaya menjaga keselamatan barang-barang yang you PASS-advice so.that take.care safety thing.PL REL di-bawa pro2 semasa berada di stesen dan di dalam tren. PASS-carry while be at station and at inside train

‘You are advised by pro1 to take care of the security of the belongings that are carried by pro2 while at the station and on board the train.’

All the examples of di- passives with first and second person agents given so far are those of the oleh type (Study 1) and the pro type (Study 2). No examples have been shown for the DP type. As the results of Study 2 suggest, it is difficult to find di- passives of the DP type with first and second person agents. However, that does not mean that first and second person agents are not attested at all for the DP type, as predicted by the restriction

11 In the fourth line, the speaker corrects his grammatical error after a pause (indicated by ‘///’). This error is not reflected in the free translation.
in (3). Examples from other sources are given in (14)–(16).

(14) *Ini yang paling comel dan di-sukai aku!*  
this REL most cute and PASS-like me  
‘This is the one that is cutest and *is liked by me!*’  
*(SHEILA & SHAHFIEKRY)*[^12]

(15) *Lirik-nya pula di-tulis saya sendiri.*  
lyric-3 on.the.other.hand PASS-write me own  
‘The lyrics on the other hand *were written by me myself.*’  
*(Harian Metro, 17/10/2012)*

(16) *Berapa buah lagu yang di-tulis anda dalam album kedua tersebut how.many CLF song REL PASS-write you in album second said  
dan apa-kah judul single pertama? — Terdapat 10 buah lagu yang and what-Q title single first be 10 CLF song REL  
kesemua-nya di-tulis oleh saya.*  
all-3 PASS-write by me  
‘How many songs are those *written by you* in the second album and what is the title of the first single?—There are 10 songs, all of which *were written by me.*’  
*(KOSMO!, 27/11/2012)*

Incidentally, a similar example is also found in Indonesian. The example in (17) is a part

---
of the lyrics of a pop song.

(17)  
Bila  
ku  
**di-terima-mu**  
when  
I  
PASS-accept-you  

Bintang-pun  
kau  
**berikan**  
star-even  
I  
give  

Bila  
kau  
**menerima-ku**  
when  
you  
accept-me  

**Ku**  
**berikan pelangi**  
I  
give  
rainbow  

‘When I’m accepted by you  
Even a star I’ll give to you  
When you accept me  
I’ll give you a rainbow’  

(GIGI Jomblo)

4. Discussion

4.1 Person restriction on **di-** passive agents: An information-based account

The preceding discussions have revealed that the person restriction on **di-** passive agents does exist. The results of Study 2 clearly show that most **di-** passives have a third person agent rather than a first or second person one. However, the restriction is not such an absolute syntactic constraint as formulated in (3), repeated below.

(3) Person restriction on **di-** passive agents

\[
**di-V Agent**_{1/2/3}  
“\text{The agent must be third person. No first and second person agents are allowed.}”
\]

We have demonstrated that instances of first and second person agents are found for all types of **di-** passives, i.e. **oleh** type, DP type and **pro** type. Their numbers are sufficiently large and cannot be ignored as errors or creative/rhetorical uses. We conclude that the restriction exists as a strong tendency whose nature is non-syntactic.

The results of this study are thus incompatible with analyses that rule out first and second person agents as totally ungrammatical. Such analyses include those proposed by Guilfoyle, Hung & Travis (1992) and Donohue (2007). Guilfoyle, Hung & Travis analyse the prefix **di-** as a realization of the third person-related features (for the conservative variety of Malay/Indonesian). Donohue, on the other hand, posits the Optimality Theoretic constraint \(*OBL/L/AGT\) as the undominated, highest ranked constraint. This constraint bans the grammatical function oblique from being assigned to participants which represent local (i.e. first or second) persons if they are agents.

Instead of strictly syntactic analyses, we propose an alternative account based on information structure: **di-** passive agents cannot be salient informationally. The salience here pertains to referential givenness in the spirit of Chafe (1976). Givenness indicates whether the denotation of an expression is present in the common ground, or the information shared by the speaker and hearer; in many theories (e.g. Prince 1992; Gundel, Hedberg & Zacharski 1993), givenness also indicates the degree to which the denotation in question is present in the immediate common ground (Krifka 2007). Thus, in terms of givenness, not being salient amounts to a low degree of givenness. The low salience/givenness of **di-**
passive agents accounts for the general rarity of first and second person agents as follows. Since the definition of givenness involves the common ground, which by definition is concerned with the mental states of the speaker and hearer, speech act participants, i.e. first and second person referents, are maximally given and inherently salient. Therefore, they are not suitable for di-passive agents, which are to be interpreted as non-salient.

Since the person restriction is informational in our analysis, the syntax can generate di-passives with first and second person agents freely, though most such sentences are filtered out by information structure. Our analysis allows for di-passives with first and second person agents in marked cases. Such marked cases include, for example, when the agent involves focus/contrast as in (18), or coordination of more than one noun phrase as in (19). The examples below are all from the DBP Corpus (4a). The word sendiri ‘own, alone’ in (18) involves focus semantics (Nomoto to appear). The entire agent phrase in (19) refers to a plural individual (plurality) including the speaker and can be substituted by the first person exclusive plural pronoun kami.

(18) a. Bayaran balik yang Allah akan berikan kepada anda hanya akan di-ketahui oleh anda sendiri.  
   ‘The rewards that Allah will give to you only will be noticed by you yourself.’

b. . . . buku ini di-tulis sendiri oleh al-Fara’, kemudian di-sesuaikan pula oleh al-Kasai dan di-baiki pula oleh saya sendiri”.
   ‘. . . this book was written by al-Fara’ himself, and later modified by al-Kasai and repaired by myself’.

(19) Malaysia di-wakili oleh saya dan tiga orang lagi rakan.  
   ‘Malaysia is represented by me and three other colleagues.’

It is worth noting that a long agent phrase like that in (19) sometimes sounds awkward in bare passive sentences (cf. Malaysia [saya dan tiga orang lagi rakan] wakili). The instability of the alternative passive construction also explains why a coordinated agent phrase is well-formed in di-passives.13 One of the reviewers of this paper suggests the possibility that dependent/subordinate clauses and relative clauses are other marked cases in which first and second person agents are tolerated by information structure.

---

13 Alwi et al. (1998:346) state that in Standard Indonesian, only di-passives but not bare passives are grammatical if the agent phrase consists of a pronoun and another pronoun or non-pronominal phrase:

(i) a. Tugas itu harus di-selesaikan oleh kamu dan saya.       (di-passive)
   task that should finish by you and me
   ‘The task should be finished by you and me.’

b.* Tugas itu harus kamu dan saya selesaikan.       (bare passive)
   task that should you and me finish
4.2 Prevalence of the pro type

We have seen in section 3.2 that, like passives in many other languages, di-passives in Malay are most frequently used without an overt agent (pro type). At first brush, this fact appears to run counter to our information-based analysis of di-passive agents presented above. This is because it is generally agreed upon in the literature of information structure that the level of salience/givenness inversely correlates with the amount of overt material, i.e. the more salient/given a denotation is, the less phonetic material the linguistic expression associated with it contains. Gundel, Hedberg & Zacharski (1993) thus identify “Ø (zero) NPs” as the form with the highest givenness status “in focus” in Mandarin Chinese, Japanese, Russian and Spanish. If what is represented as pro above were the same thing as their “Ø (zero) NPs,” pro should be more salient/given than the overt first and second person pronouns, and one would expect the pro type di-passive to be the least frequent of the three types, quite contrary to actual fact.

The generalization about the relationship between salience and the amount of overt material holds true in Malay. Its validity can be confirmed by the fact that inherently salient first and second person agents are more difficult to find in the DP type than in the oleh type. The oleh-less counterparts of the ‘di-V oleh DP’ patterns found in Study 1 are either absent or less frequent in the same corpus. This fact can be checked by Google searches as well. For example, we found 21 instances of dibuat oleh kita (“dibuat oleh kita”) (cf. (7)) in Malaysian domain (.my) pages, but we could find only 3 instances of its oleh-less counterpart dibuat kita (“dibuat kita”) (searches conducted on 15/01/2014). These facts are expected because the agent phrase of the DP type is more salient, being one word shorter than that of the oleh type.

The problem lies in the assumption that pro is a kind of “Ø (zero) NP.” We argue that pro is an unspecified pronoun with no person or number specification, and that due to its unspecified nature, pro is low in salience/givenness. If so, the prevalence of the pro type di-passive makes perfect sense.

Our argument is justified by the following points. First, the interpretation of pro is not always straightforward. In many cases, it seems most appropriate to analyse pro as “unspecified,” though its referent is obvious in some cases. In this respect, pro is similar to orang, which literally means ‘person/people’ and is usually considered to be third person or unspecified, but sometimes refers to the speaker or, less commonly, the hearer. Recall that that all instances of first and second person agents in Study 2 occur in the pro type. Since the existing grammars of Malay do not recognize a null unspecified pronoun as a part of the pronominal paradigm of the language, we classified such cases that are actually best coded as “unspecified” into the category that was most plausible in accordance with the linguistic and extra-linguistic contexts.

Secondly, the null unspecified pronoun pro is employed elsewhere in the language. Nomoto (to appear) shows that positing the null unspecified pro greatly simplifies the description and analysis of anaphoric expressions in Malay. Specifically, he claims that pro is a possessive pronoun comparable to one’s in English and occurs in the following

---

14 In fact, there were many more results for the phrase "dibuat kita" than for "dibuat oleh kita". However, the majority of the results for the former were irrelevant because kita belonged to a different clause than the verb and hence was not the verb’s agent. Such results were not counted.
two contexts: as a possessor argument of diri ‘(physical) self’ and with the intensifier sendiri ‘own, alone’, as shown in (20a) and (20b) respectively.

(20) a.  
  diri pro  ‘oneself’  b.  kereta pro sendiri  ‘one’s own car’
     diri-ku  ‘myself’                  kereta-ku sendiri  ‘my own car’
     diri-mu  ‘yourself’               kereta-mu sendiri  ‘your own car’
     diri-nya  ‘himself/herself’      kereta-nya sendiri  ‘his/her own car’

Pro in di- passives occurs in the same syntactic context, i.e. positions associated with genitive structural case. It can occur immediately after the verb, as in (21a), but cannot occur after the preposition oleh, as in (21b), because the object of a preposition is assigned accusative case. Note that genitive and accusative overt enclitics are identical in form.

(21) a.  
  di-semak pro  b.  *di-semak oleh pro  ‘be checked (by someone)’
     di-semak-ku  di-semak oleh-ku  ‘be checked by me’
     di-semak-mu  di-semak oleh-mu  ‘be checked by you’
     di-semak-nya  di-semak oleh-nya  ‘be checked by him/her’

Moreover, the fact that the agent of a bare passive (cf. (1b)) must be expressed overtly, which previous studies have stipulated in one way or another, now follows automatically from the Case condition on pro. The agent position of bare passives is thought to be associated with nominative (or ergative) case, but not with genitive case, because only proclitic but not enclitic forms of overt pronouns are licensed there, as shown by the contrast between (22a) and (22b). Hence, the covert pronoun pro, associated with genitive case, is not licensed in this position, as in (22c).

(22) a.  
  Dokumen itu sudah {ku-/kau-}semak.16  
  document that already {I-/you-} check
     ‘I/You have already checked the document.’

  b.  *Dokumen itu sudah {-ku-mu-nya} semak.  
     document that already{-my/-your/-his/her} check

  c.  *Dokumen itu sudah pro semak.  
     document that already one check

4.3 Why does the restriction exist only in di- passives?

So far, we have accounted for the person restriction on di- passive agents in terms of informational salience/givenness: first and second person agents are rarely used in di-passives because di- passive agents should not be salient whereas first and second person agents, being speech act participants, are inherently salient. It must be emphasized here that pointing out the relation of a constituent in a construction to a particular informational

---

16 Malay lacks a third person proclitic pronoun. One may be tempted to regard the passive marker di- as the missing proclitic, assuming that di- passives agents are always third person. However, as shown earlier, such an assumption is wrong. See Adelaar (2005) for a discussion of the historical origin of di-.
status is only a half of the explanation.\textsuperscript{17} The fundamental question that must be addressed to complete the explanation is why the relationship in question exists. In the present study, the question is: What mechanism is responsible for the low salience of \textit{di-} passive agents? Since no person restriction is present for bare passive agents, the answer to this question must also account for the difference between \textit{di-} and bare passives. It is therefore necessary to discuss the syntax of passives.

The trees in (23a–c) and (23d) schematically show the structures we propose for the three \textit{di-} passive types and bare passives, respectively.

\begin{enumerate}
\item \textit{Pro} type
\begin{itemize}
\item \textit{DP} \hspace{1cm} \textit{vP}
\item \textit{DP} \hspace{1cm} \textit{vP}
\item \textit{DP} \hspace{1cm} \textit{vP}
\end{itemize}
\item \textit{Oleh} type
\begin{itemize}
\item \textit{DP} \hspace{1cm} \textit{vP}
\item \textit{DP} \hspace{1cm} \textit{vP}
\item \textit{DP} \hspace{1cm} \textit{vP}
\end{itemize}
\item \textit{DP} type
\begin{itemize}
\item \textit{DP} \hspace{1cm} \textit{vP}
\item \textit{DP} \hspace{1cm} \textit{vP}
\item \textit{DP} \hspace{1cm} \textit{vP}
\end{itemize}
\end{enumerate}

\textsuperscript{17} This does not apply to functional approaches in which semantic/pragmatic functions are primitives and used to define syntactic constructions. For example, Givón (2001:94) defines the passive as the voice whereby “the agent is extremely non-topical (‘suppressed’, ‘demoted’), so that the patient is the surviving topical argument in the clause”, where ‘topicality’ is the extent to which a referent continues to be mentioned in a text. Since functions instead of structural characteristics are primitive, the range of passives is far broader in functional approaches than in structural approaches. Givón thus regards sentences such as \textit{One fires people occasionally} and \textit{They dance in the street there} also as passives (136).
d. Bare passives

\[
\text{vP} \\
\text{DP (overt agent)} \\
\text{v'} \\
\text{v} \\
\text{VP} \\
\text{Ø}
\]

Our analysis posits no morphosyntactic condition on the person of the agent DP. Being a passive voice marker, the prefix \textit{di-} is a v head lacking an accusative Case assigning ability, unlike an active voice marker, which has this ability. Adopting Legate’s (2012) analysis of Acehnese passives, we claim that \textit{di-} does not suppress the external argument (logical subject). The external argument position (Spec,vP) is occupied by \textit{pro} in the \textit{pro} and \textit{oleh} types (23a–b), and by an overt DP in the DP type (23c). The referent of the null unspecified agent \textit{pro} either is left unspecified or is specified by pragmatic inference in the \textit{pro} type (23a), whilst it is restricted by the overt agent DP in the \textit{oleh} type (23b). The overt agent DP in the DP type is an argument merged in Spec,vP, but not a result of “omitting” \textit{oleh} ‘by’ from the \textit{oleh} type in (23b).

One may wonder if (23c) is a correct structure, because the surface order of \textit{di-} passives is ‘\textit{di-}V (\textit{oleh}) \textit{DP}_\text{agent}’, not ‘\textit{DP}_\text{agent} \textit{di-}V’. We capture this surface word order by hypothesizing V-to-v and v-to-T movements, as shown in (24). These movements do not affect the surface order in the \textit{pro} and \textit{oleh} types, as \textit{pro} in Spec,vP, which is crossed by the verb, is phonologically deficient.

\[(24) \quad [\text{TP} \left[ T + [\text{v, } \textit{di-} + \text{V }] \right] [\text{vP} \text{ DP (overt agent)} [\text{v'} [\text{v \textit{di-}}] [\text{vP} \text{ X}]]]]]

\[\text{v-to-T movement} \quad \text{V-to-v movement}\]

Our analysis of \textit{di-} passives suggests that bare passives are only minimally different from \textit{di-} passives. Bare passives differ from \textit{di-} passives in two respects (apart from the verbal morphology). First, they do not allow \textit{pro} to occur in Spec,vP; the agent must be overt. Second, the verb does not move to T; it occurs immediately after the agent. Since, in disallowing \textit{pro} in Spec,vP, bare passives are the same as DP type \textit{di-} passives, one can regard bare passives as the fourth type of \textit{di-} passives, namely a DP type without v-to-T movement. Crucially, this means that \textit{di-} and bare passives are in fact not two substantially different types.

---

18 A similar situation holds for \textit{pro} in anaphoric expressions (cf. (20)). According to Nomoto (to appear), \textit{pro} in ‘\textit{diri pro}’ remains unspecified and gives rise to a generic interpretation, whereas \textit{pro} in ‘\textit{pro sendiri}’ is restricted by and specified as the antecedent of \textit{sendiri}. Incidentally, stipulating coreference between \textit{pro} and the overt agent phrase to determine the interpretation of \textit{pro} in the \textit{oleh} type (23b) creates a violation of Binding Condition B, as one of the reviewers points out.

19 According to Kroeger (this volume), Classical Malay has another type which can be situated between (23b) and (23c). In this type, the agent DP restricted by the overt agent phrase is not \textit{pro} but the third person enclitic -\textit{nya}, as in (i), taken from Sejarah Melayu. Such a structure is thought to have mediated the diachronic development of the \textit{oleh} type from the DP type.

(i) Maka \textit{oleh} segala mereka itu akan Raja Suran di-bawa-\textit{nya} kepada raja-\textit{nya}.

\[\text{so by all } \text{them that of Raja Suran PASS-carry-3 to king-3} \]

‘So, they all took Raja Suran to their king.’ (A. Samad 1979:15)
different voice systems, but two subtypes of the same voice. The verbal morphology difference between *di-* and bare passives is thought to reflect the presence/absence of v-to-T movement.

Let us now return to the issue of why the agent must not be salient in *di-* passives whereas no such restriction exists for bare passives. Given that the only syntactic difference between the two passive types is the presence or absence of v-to-T movement, we propose to relate the low salience of *di-* passive agents to the v-to-T movement.

The salience/givenness properties of noun phrases are usually regarded as lexically specified. For example, the definite article *the* in English encodes as part of its meaning the salience of ‘*the NP*’. A similar analysis, however, is not possible for *di-* passive agents. This is because the salience at issue does not depend on the form of a noun phrase, but on its relation to a verb, specifically agenthood. Extending the analysis of *the* to the prefix *di-* will not result in the salience of the individual denoted by the agent noun phrase. Rather, what one obtains instead is the salience of the eventuality described by the *di-* verb. It is thus necessary to somehow derive the low salience of *di-* passive agents from the salience of this eventuality.

We suggest that a v-to-T movement makes the eventuality described by the verb non-salient, and consequently the agent involved in this eventuality cannot be salient either. The theme argument is not affected by this informational effect of v-to-T movement, because the theme DP moves to the preverbal subject position (Spec,TP) and the informational status associated with the subject position overrides the effect of the verb movement.

Unlike noun phrases, the salience/givenness of verbs has been neglected in the literature of information structure (but see Schwarzschild 1999). Hence, it is unclear how to prove the low salience of the eventuality described by a *di-* verb. Fortunately, however, Hopper’s (1983) study of the discourse functions of three clause types in the Early Modern Malay text *Hikayat Abdullah*, in our understanding of it, lends support to our claim. Hopper maintains that a verb-initial clause such as those in boldface in (25) below “serves to FOREGROUND events” (72), “focuses purely on the event—the change—itself” and “narrates sequenced events which pertain to the main line of the discourse” (84). Verb-initial sentences are used in the same way in Modern Malay, though they are limited to the literary genre.

---

20 If so, the commonly used alternative name of bare passives ‘object(ive) voice’ (cf. footnote 2) is deemed a misnomer, as it implies the existence of a voice system completely distinct from *di-* passives. Moreover, the related view that Malay has the Philippine-type symmetric voice system is also problematic.

‘Now the fire came about because the crewmen were smoking cheroots, and they threw away the stubs into the boat, and the fire spread to the ropes and burned up the ship.’ (Hopper 1983:72)

Hopper states that this function of verb-initial clauses is obliterated by the positioning of a noun phrase before the verb, which he analyses as “a device for arresting the flow of the discourse and holding up the action by momentarily focusing attention away from ACTIONS TO PARTICIPANTS” (87). This quote indicates that by “foreground” Hopper means “require or draw attention of the addressee.” In terms of salience/givenness, it is a denotation which is not already salient/given enough in the common ground that requires special attention of the addressee. Hence, in verb-initial clauses, the verb is not salient. As schematically shown in (24) above, a verb-initial clause is created in $\textit{di}$- passives, as it involves a v-to-T movement; this movement does not occur in bare passives. One can thus think of the discourse function of verb-initial clauses pointed out by Hopper as reflecting the low salience status of the $\textit{di}$- verb. The low salience status of the $\textit{di}$- verb in turn will not arise without a v-to-T movement.

The remaining issue, of course, is what triggers this verb movement. At this stage of research, we can only conjecture that the movement occurs to meet the conditions related to the syntax-information structure interface. Specifically, the interface cannot read off the information status of the verb specified lexically by $\textit{di}$- independently from the verb’s arguments if the verb stays within vP. This is presumably because vP is the syntactic domain in which the verb’s core arguments are first merged.

5. Conclusion

This paper has shown based on various texts that the person restriction on $\textit{di}$- passive agents is descriptively accurate and claimed that it exists not as an absolute syntactic rule but as a strong tendency. We proposed that the restriction is ultimately ascribed to the lexical meaning of $\textit{di}$- and syntax. The prefix $\textit{di}$- encodes that the eventuality described by ‘$\textit{di}$- V’ is non-salient/new; the syntactic verb movement makes this informational meaning accessible for interpretation. $\textit{Di}$- passive agents also become non-salient as a consequence of the low salience of the eventualities with which they are associated. Since first and second person agents are speech act participants and inherently salient, they are not suitable for $\textit{di}$- passive agents. Unlike $\textit{di}$- passives, bare passives are not subject to the relevant person restriction, because they involve neither the prefix $\textit{di}$- nor v-to-T movement.
Abbreviations

1 first person  PP preposition phrase  
2 second person  PRF perfect  
3 third person  PL plural  
CLF classifier  Q question marker  
DP determiner phrase  REL relativizer  
EXCL exclusive  SG singular  
INCL inclusive  Spec specifier  
NP noun phrase  TP tense phrase  
P preposition  V verb  
PART particle  v little verb  
PASS passive  vP little verb phrase  

References


Kroeger, Paul. this volume. Passive agents in Malay: The binding properties and discourse functions of agentive =nya.


