1. Introduction

1.1. Background

(1) Two types of passives in Malay/Indonesian

a. *Di-* passive: overt passive marker *di-*, agent = adjunct
   Dokumen itu sudah *di*-semak (oleh *meraka*),
   document that already *PASS*-check by 3PL
   ‘The document has already been checked by them.’

b. Bare passive: no overt voice marker, agent = argument (Spec,vP)
   Dokumen itu sudah *(meraka) semak,
   document that already 3PL check
   ‘They have already checked the document.’

• The topic of this paper: the person restriction on the agent in *di-* passives (1a)

(2) Three patterns to encode agents

a. *Oleh“ type: introduced by the preposition *oleh* ‘by’
   Surat itu sudah *di*-poskan *oleh* kerani,
   letter that already *PASS*-post by clerk

b. *DP“ type: immediately after V, no preposition
   Surat itu sudah *di*-poskan *kerani,
   letter that already *PASS*-post clerk

c. *pro“ type: no overt agent
   Surat itu sudah *di*-poskan *pro*,
   letter that already *PASS*-post
   ‘The letter was already posted (by the clerk).’

• Prescriptive grammars impose a person restriction on the agent (e.g. Nik Safiah et al. 2008):

   (3) Person restriction on the *di-* passive agent
      *di-V Agent$_{1/2/3}$
      “The agent must be 3rd person. No 1st and 2nd person agents are allowed.”

• Descriptive grammars are not uniform with respect to the person restriction.

(4) a. The restriction is descriptively real.
   e.g. Arka and Manning (1998), Donohue (2007), Sneddon et al. (2010)

b. The restriction is only a prescriptive rule and descriptively inadequate.

c. Two groups of speakers exist that differ in the presence of the restriction.
   Guilfoyle, Hung and Travis (1992)

1.2. The issue

Q: Is the person restriction in (3) descriptively real? If so, to what extent?

• The answers to these questions delimit the possible analyses of the syntax of *di-* passives.

• This paper attempts to answer these questions by examining various naturally occurring texts in Standard Formal and Standard Colloquial Malay (Malaysia). ¹

Our answer: Yes, but not in the form as strict as (3).

1.3. Organization

§1 Introduction

§2 The data

§3 Study 1: Concordance
   — deals with the “oleh” type; shows the existence of *di-* passives with 1st and 2nd person agents

§4 Study 2: Frequency
   — deals with all types; shows the frequencies of 1st and 2nd person agents

§5 Conclusion
   — discusses implications for the syntax of *di-* passives

¹ These Malay varieties should not be confused with the prescriptive variety, which we refer to as ‘Standardized Malay’ (bahasa Melayu baku).

— This study was supported in part by the JSPS Grant-in-Aid for Young Scientists (B) (#23720199). We would like to thank our assistants Ayano Osaka, Miki Osuga and Nahoko Yamashita for helping us with the data collection and coding.
2. The data

(5) a. The magazine (majalah) subcorpus of the DBP Corpus
b. Front page articles of Utusan Malaysia for the entire period of year 2011
c. A collection of folktales (cerita rakyat): used to be available at the DBP website

d. The Multilingual Corpora (Malay)

e. Miscellaneous (e.g. web pages)

Table 1 The data for this study

<table>
<thead>
<tr>
<th>Texts</th>
<th>Data type</th>
<th>Token</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBP (5a)</td>
<td>formal</td>
<td>14,406,888</td>
<td>(unknown)</td>
</tr>
<tr>
<td>Utusan Malaysia (5b)</td>
<td>formal</td>
<td>501,272</td>
<td>17,728</td>
</tr>
<tr>
<td>Folktales (5c)</td>
<td>formal ~ colloquial</td>
<td>66,711</td>
<td>6,263</td>
</tr>
<tr>
<td>Multilingual Corpora (5d)</td>
<td>colloquial</td>
<td>232,310</td>
<td>9,974</td>
</tr>
</tbody>
</table>

Utilizing multiple data sources ensure that our data is large and representative enough, and contains both Formal and Colloquial Malay sentences.

3. Study 1: Concordance

3.1. Method

(6) a. Data: The magazine subcorpus of the DBP Corpus (5a)
b. Procedure: Searched for the following patterns using the online concordance system:

\[
\begin{align*}
\text{di (PASS) oleh 'by' saya (1SG)} \\
\text{di oleh aku (1SG)} \\
\text{di oleh ku (1SG) [ku: the enclitic form of aku]} \\
\text{di olehku (1SG)} \\
\text{di oleh kami (1EXCL)} \\
\text{di oleh kita (1INCL)} \\
\text{di oleh awak (2)} \\
\text{di oleh kamu (2)}
\end{align*}
\]

c. Limitations

(i) The system only allows one to handle the “oleh” type. The “DP” and “pro” types are (almost) impossible to find automatically.

(ii) The search results are limited to 100 instances, which include many duplicates. Unsuitable for a quantitative analysis.

3.2. Results

- Many di- passives were found with 1st and 2nd person agents.

Table 2 Di- passives with 1st and 2nd 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>1SG</td>
<td>saya</td>
<td>27</td>
<td></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>1EXCL</td>
<td>kami</td>
<td>16</td>
<td>77</td>
</tr>
<tr>
<td>1INCL</td>
<td>kita</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>awak</td>
<td>3</td>
<td>149</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>

3.3. Examples

(7) Perkara yang paling diingati 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 …’
Usia tidak mengampunkan segala dosa yang dibuat oleh kita. ‘Age does not forgive all the sins that we committed by us.’

Planta Soft telah menyediakan segala yang diingini oleh anda iaitu majerin yang lembut dan sangat lazat. ‘… Planta Soft has prepared everything that is wanted by you, namely margarine that is soft and very tasty.’

3.4. Unsolved questions
- While the restriction in (3) seems inadequate for the “oleh” type, it may be valid for the “DP” and “pro” types.

Q1: Are 1st and 2nd person agents also possible for the “DP” and “pro” types?
- Prescriptive grammars tend to ban a pattern when it is rarely used. Indeed, one seldom encounters di-passives with 1st and 2nd person agents.

Q2: How frequent are 1st and 2nd person agents?

4. Study 2: Frequency

4.1. Method

a. Data
(i) Front page articles of Utusan Malaysia in 2011 (5b)
(ii) Folktales (5c)
(iii) The Multilingual Corpora (Malay) (5d)

b. Procedure
(i) Took the first 100 instances of di- verbs.
(ii) For each item, coded
   - the type of agent encoding: “oleh”, “DP”, “pro”
   - the person of the agent: “1st”, “2nd”, “3rd”

- When more than one analysis is possible, the most likely code was assigned.

4.2. Results
- 1st and 2nd person agents were attested, but very infrequent (0–3%).
- All instances of them occurred in the “pro” type.

4.3. Examples

4.3.1. Utusan Malaysia

Beliau ditemui pro selepas merasmikan Seminar Pengurusan Sukan Institusi Pengajian Tinggi (IPT) 2010 di UiTM kampus Khazanah Alam Bandar Jengka di sini. (01/01/2011)

‘He was met by pro after he had officiated the 2010 Higher Academic Institution Sports Management Seminar at UiTM, Khazanah Alam Bandar campus here.’

- pro = writer, *reader, *a third party excluding writer/reader

Jika dilihat pro blog-blog dan laman sosial popular Facebook turut mendedahkan ‘keberanian’ golongan gay itu berkongsi pengalaman peribadi dan kehidupan mereka sebagai pengamal seks songsang. (02/01/2011)

‘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.’

- pro = writer, *writer & reader, *a third party excluding

4.3.2. Folktales

Jika dilihat pro blog-blog dan laman sosial popular Facebook turut mendedahkan ‘keberanian’ golongan gay itu berkongsi pengalaman peribadi dan kehidupan mereka sebagai pengamal seks songsang. (02/01/2011)

‘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.’

- pro = writer, *writer & reader, *a third party excluding

4.3.3. Multilingual Corpora

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Di-passive agents in Utusan Malaysia (5b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>oleh</td>
<td>DP</td>
</tr>
<tr>
<td>1st</td>
<td>0</td>
</tr>
<tr>
<td>2nd</td>
<td>0</td>
</tr>
<tr>
<td>3rd</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Di-passive agents in Folktales (5c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>oleh</td>
<td>DP</td>
</tr>
<tr>
<td>1st</td>
<td>0</td>
</tr>
<tr>
<td>2nd</td>
<td>0</td>
</tr>
<tr>
<td>3rd</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Di-passive agents in Multilingual Corpora (5d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>oleh</td>
<td>DP</td>
</tr>
<tr>
<td>1st</td>
<td>0</td>
</tr>
<tr>
<td>2nd</td>
<td>0</td>
</tr>
<tr>
<td>3rd</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
</tr>
</tbody>
</table>
4.3.2. Multilingual Corpora (Malay)

(14) [Hmm] yang aku paling best tu, [eh] yang bab aku paling best pulak yang rasenye macam kalau nak dibandingkan pro dengan drama Jepun dengan drama Korea, aku le // aku lebih suka tengok drama Jepun. (01Dec#3-B4) ‘Hmm… what’s I’m best is, eh, my best favourite is, I guess, like, if Japanese dramas are compared by pro with Korean ones, for me… I like to watch Japanese dramas better.’

(15) Payah jugak kan kalau kita belajar bahasa lepas tu masih keliru, tapi kan (Zu) [erm] kalau hendak dibezakan pro antara bahasa Melayu dengan bahasa Jepun // rasanya mana yang lebih mudah dipelajar // dipelajari? (01Dec#3-B105) ‘If you learn a language and still don’t get it, it looks 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 learned… learn?’

4.3.3. Miscellaneous

(16) KTM Berhad (Malayan Railway) signboard (Figure 1) Anda dinasihatkan pro1 supaya menjaga keselamatan barang-barang yang dibawa pro2 semasa berada di stesen dan di dalam tren. ‘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.’

Figure 1  Di- passives with 1st and 2nd person agents in a signboard (picture taken 22/05/2013)

- 1st and 2nd person agents are difficult to find for the “DP” type.

(17) Ini yang paling comel dan disukai aku! ‘This is the one that is cutest and is liked by me!’ (SHEILA & SHAHFIKRY)

(18) Liriknya pula ditulis saya sendiri. ‘The lyrics on the other hand were written by me myself.’ (Harian Metro, 17/10/2012)

(19) Berapa buah lagu yang ditulis anda dalam album kedua tersebut dan apakah judul single pertama? — Terdapat 10 buah lagu yang kesemuanya ditulis oleh saya. ‘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)

6 http://www.sheilashahfiekry.my/2012/08/ini-entri-takda-makna-pun.html (accessed 26/05/2013)
5. Conclusion

5.1. Summary

(3) Person restriction on the di-passive agent

di-V Agent

“The agent must be 3rd person. No 1st and 2nd person agents are allowed.”

- The person restriction on di-passive agents does exist. Most di-passives have a 3rd person agent rather than a 1st or 2nd person one.
- However, the restriction is not an absolute syntactic constraint as formulated in (3). Instances of 1st and 2nd person agents are found for all three types of di-passives: “oleh”, “DP” and “pro”.
- Instead, the restriction is a strong tendency whose nature is largely non-syntactic.
- Besides the existence of the person restriction, this study has also revealed:

5.2. Implications for the syntax of di-passives

- The generalizations in (21) are in accordance with Kartini and Nomoto’s (2012) analysis of di-passives.

(22) Kartini and Nomoto (2012)

a. No morphosyntactic condition on the agent DP.
b. di- = v that specifies Restrict (Chung and Ladusaw 2004) as the semantic operation to apply next (= Di-passives are an external argument “incorporation” construction.)
c. Spec,vP (logical subject) of the “pro” (2c) and “oleh” (2a) types = pro

d. Spec,vP (logical subject) of the “DP” type (2b) = overt DP

The “DP” type is not a kind of the “oleh” type with a null oleg ‘by’.

(23) a. Buku itu di-baca lelaki itu.

b. λx[read’(x, book)](φ)(man) = [read’(man, book) ∧ φ(man)]

The agent DP occurs after the verb as a result of V-to-v and v-to-T movements.

While the passive markers in Malay and English impose no restriction on the agent, the related language Acehnese employs different passive markers that impose different restrictions on the agent in terms of person, number and familiarity. See Legate (2012) for a Restrict-based analysis of Acehnese passives.

---

7 Di-passives in Malay thus fit with the following characterization by Keenan and Dryer (2006: 332):

“And text counts for various languages (e.g. English by Svartvik (1966), Dutch by Kirsner (1976), Chamorro by Cooreman (1987), Modern Greek by Roland (1994)) show that agented passives are much less frequent than agentless ones.”

8 The agent DP occurs after the verb as a result of V-to-v and v-to-T movements.

9 While the passive markers in Malay and English impose no restriction on the agent, the related language Acehnese employs different passive markers that impose different restrictions on the agent in terms of person, number and familiarity. See Legate (2012) for a Restrict-based analysis of Acehnese passives.
Di-passives are typically used without an overt agent (“pro” type).

1st and 2nd person agents become more difficult in the following order:

“pro” type > “oleh” type > “DP” type

Why (21a)?
Agents are informationally backgrounded in di-passives (22b).

Why (21b)?
- Assumption: Speech act participants (1st and 2nd) are more prominent than non-participants (3rd).
  → 1st and 2nd person agents are very rare.
- Overt materials are more prominent than covert ones.
  → “pro” > “oleh”/“DP”
- Arguments are more prominent than adjuncts.
  → “oleh” > “DP”

cf. Prediction based on the amount of overt materials: “DP” > “oleh”

The results of this study are incompatible with the analyses by Guilfoyle, Hung and Travis (1992) and Donohue (2007). These studies predict that 1st and 2nd person agents are totally ungrammatical.

Guilfoyle, Hung and Travis (1992)

di- = a realization of the 3rd person-related features (in the conservative variety).

Donohue (2007)

*Onm/LI/Agt (Do not assign the grammatical function oblique to participants which represent local (1st or 2nd person) persons if they are agents) is the undominated, highest ranked constraint.

References