Preliminary observations on the stylistic change in direct speech encoding strategies in Malay and Indonesian

Asako SHIOHARA[#] & Yuta SAKON[*]

#Research Institute for Languages and Cultures of Asia and Africa,
Tokyo University of Foreign Studies

*Tokyo University of Foreign Studies

Classical Malay (CM) narratives feature a style with highly schematized structures, with the direct speech (DS) encoding strategy as a key example; the pre-DS frame, comprising a root and speaker constituent (RF), is predominant. In contrast, modern Indonesian (MI) employs a more diverse range of strategies, including the post-DS RF, pre-DS speech verbs, and frameless DS (or free DS). This study investigates stylistic change by examining texts from the transitional period, specifically the late 19th to early 20th century. The rise of post-DS RF is traced back to atypical CM texts from the late 19th century. Concurrently, Low Malay novels, influenced by Western literature, embraced post-DS frames and frameless DS as novelistic techniques. In the early 20th century, Balai Pustaka established a new standard, adopting strategies that continues to be the norm today. This stylistic change can be connected to broader contexts, such as the transition from oral to written media or from aural to visual experiences.

1. Introduction[1]

This is a diachronic study of the styles observed in strategies for coding direct speech (DS) in narratives from classical Malay (CM) to modern Indonesian (MI). Direct speech is a type of reported speech characterized by delivering the speaker's words as they were spoken, instead of from the viewpoint of the narrator, as in indirect speech (IS) (Leech and Short 1981).

CM, represented by traditional narratives often labelled as *hikayat*, *cerita* (both mean "a story"), or *sejarah* "history," is known to use highly schematized expressions, both lexically and structurally (Sweeney 1987). This also applies to the structure to encode DS; the norm of the DS frame in CM is characterized by the employment of root forms followed by the speaker noun phrase (the speaker NP).

Excerpt (1) is from *Hikayat Raja Pasai*. Here the root *kata* "words" frames the DS followed by the speaker NP *orang dalam jong itu* "the person in the ship" in the first and third clause, while the root *sahut* "reply" frames the DS followed by the speaker NP *orang mengail itu* "the person who is fishing" in the second clause.[2] In all clauses, DS follows

---

[2] Judgement about the word-class attributes of individual roots may vary across speakers. With regard to this point, we followed the KBBI (Kamus Besar Bahasa Indonesia) https://kbbi.kemdikbud.go.id/.

---

the frames. Hereafter, these types of frames will be labelled as "root frame" (RF), which refers to a root followed by the speaker constituent.

(1) (a) *Maka **kata** orang dalam jong itu,* "*Apa ada khabar dalam negeri ini?*"
     (b) *Maka **sahut** orang mengail itu,* "*Adapun khabar dalam negeri ini badak makan anaknya.*"
     (c) *Maka **kata** orang dalam jong itu,* "*Tiada kami tahu apa ertinya katamu itu.*"

     (a) So, the person in the ship said, "Is there any news in this country?"
     (b) Then the person who was fishing answered, "As for the news in this country, rhinos eat their young."
     (c) The person on the ship said, "We do not know what you mean by that word." (119:18, in Jones 1987:63)

The word class attributes of the roots may be either a noun (e.g., *kata* "word"), a verb (e.g., *sahut* "answer"), or pre-categorial (Himmelmann 2005:129; e.g., *-panggil*)[3]. As a DS frame, RF is opposed to frames headed by a verb with derivational morphemes, such as *berkata* "say," as shown in Excerpt (2) below. This type of frame will be labeled as the "verb-frame" (VF). In the following example sentences, the DSs and frames are marked by square brackets with the indications DS and RF or VF, respectively.

In contrast to CM, modern Indonesian standards (MI) employ different strategies to encode DS. First, the RF follows or is inserted in the DS, and the VF replaces the RF in the pre-DS position. Second, DS may occur without frames, especially in novels (Cumming 1991:66, Wijaya 2016:3–4, Djenar 2018:9–10); this type of DS will be labelled as frameless DS in what follows. Excerpts (2) and (3) were cited from a modern Indonesian novel, *Bumi Manusia*, published in 1980. In Excerpts (2)(a)–(f), the DS in (a) is framed by the inserted RF, while the DS in (e) and (f) is framed by the post-DS RF. In (b) to (d), the DSs are presented as frameless.

(2) (a) ["*Minke,*"]$_{DS}$ [***panggil Nyai***]$_{RF}$, ["*benarkah orang sudah mulai bisa bikin es? Es yang benar-benar dingin seperti dalam buku-buku itu? Seperti yang membeku di musim salju di Eropa?*"]$_{DS}$
     (b) ["*Betul, Mama, setidak-tidaknya menurut suratkabar.*"]$_{DS}$
     (c) [*Suurhof menelan sambal mendelik padakku.*] $_{DS}$
     (d) ["*Aku hanya mau tahu apa berita koran itu benar.*"]$_{DS}$
     (e) ["*Nampaknya semua akan bisa dibikin oleh manusia, Mama,*"]$_{DS}$ [***jawabku***]$_{RF}$, *tapi dalam hati aku heran ada orang bisa meragukan berita koran.*
     (f) ["*Semua? Tidak mungkin,*"]$_{DS}$ [***bantahnya***]$_{RF}$. (p. 42)

     (a) "Minke," Nyai called out, "Is it true, people can now make ice? Ice that is really cold, as the books say?"
     (b) "It's true, mama, at least according to the newspapers."
     (c) "Suurhof swallowed while glaring at me."
     (d) "I want to know if the newspaper reports are true."
     (e) "It seems everything will be able to be made by man, madam," I answered, though, in my heart, I was more amazed that somebody could doubt a newspaper report.
     (f) "Everything? Impossible," she replied.

---

[3] Pre-categorial roots are denoted by a hyphen (-) at the beginning.

In Excerpt (3), the first DS clause is framed by pre-DS VF, while the next three DS clauses are frameless.

(3) (a) *Di atasnya berdiri jambang bunga dari tembikar bikinan Eropa. Bunga-bungaan bersembulan dari dalamnya dalam karangan yang serasi.*

   (b) *Annelies mengikuti pandangku,* [***berkata***]VF: [*"Aku sendiri yang merangkai."*]DS

   (c) [*"Siapa gurunya?"*]DS

   (d) [*"Mama, Mama sendiri."*]DS

   (e) [*"Bagus sekali."*]DS

   (a) Above it, is the floral crest of European-made pottery. Flowers bloomed from it when matching the bouquets.

   (b) Annelies followed my gaze and said, "I made the arrangements myself."

   (c) "Who was the teacher?"

   (d) "Mama, Mama herself."

   (e) "Very good." (p. 31)

Strategies employed in the MI novel observed in Excerpts (2) and (3), that is, pre-DS VF and post-DS RF, are also dominant in other MI genres, such as newspaper articles. Excerpts (4) and (5) were cited from articles in the newspaper KOMPAS in 2018. In Sentence (4), the pre-DS VF is, while in Excerpts (5), the post-DS RF is employed.

(4) *Sebelum mengakhiri kata sambutannya, Jokowi* [***berkata***]VF, [*"Oh iya, ada yang tahu enggak nama ibu dan bapak menteri yang ada di sana?"*]DS *sambil menunjuk sejumlah menterinya yang duduk berderet di bagian kanan.*

Before ending his remarks, Jokowi said, "Oh yes, does anyone know the names of the ministers who were there?" while pointing to a number of his ministers who sat in a row on the right.

(https://nasional.kompas.com/read/2018/05/04/18162831/siapa-nama-ibu-menko-pmk-jawaban-anak-perempuan-ini-bikin-jokowi-ngakak)

(5) [*"Mereka bersyukur karena sudah bisa lolos, bisa ikut pemilu. Intinya itu,"*]DS [***kata Jokowi***]RF.

"They are grateful because they have been able to pass, they can participate in the election. That is the point," said Jokowi.

(https://nasional.kompas.com/read/2018/03/02/14102441/kata-presiden-soal-pertemuan-90-menit-dengan-psi-di-istana?page=all.)

Thus, the preferred DS-encoding strategies in CM and MI are summarized in Table 1.

Table 1. Preferred DS encoding strategies in each text type

	CM	MI
Pre-DS RF	+	−
Pre-DS VF	−	+
Post-DS/ inserted RF	−	+
Post-DS/ inserted VF	−	−
Frameless DS	−	+

In sum, the transition from CM to MI can be encapsulated in three main aspects: (i) the relocation of RF from the pre-DS position to the post-DS position, (ii) VF taking over the pre-DS position previously occupied by RF, and (iii) the emergence of frameless DS.

This study presents a preliminary observations and analysis of the aforementioned shifts in DS encoding strategies in Malay and Indonesian languages. The structure of the rest of this paper is as follows: Section 2 provides a quantitative analysis of the strategies used in CM and MI, confirming the change. In Section 3, we investigate the transition using late 19th to early 20th-century literary sources when the new standard emerged. Finally, in Section 4, we summarize the discussion.

In the remainder of the Introduction, we address two fundamental premises of this study. The first concerns the idiosyncratic syntactic status of RF within both CM and MI grammar. Based on Cumming's (1991:67) analysis, the most plausible approach would be to view RF as an NP forming an equational sentence with DS. However, with pre-categorial roots such as *-panggil* "call", the structure cannot be considered an equational sentence. Additionally, some root forms that are frequently used in RF, such as *sahut* "answer" and *jawab* "answer", although defined as noun roots in KBBI, idiosyncratically occur only in this specific position; in all other syntactic environments where nouns are required, the *-an* suffixed forms, such as *sahut-an* "answer" and *jawab-an* "answer," are employed. Thus, instead of considering this structure as an equational sentence, it would be more appropriate to view it as a distinct syntactic construction, either derived from the equational sentence or not, specific to DS encoding. For the MI usage of RF inserted in DS, as illustrated in Sentence (1), the interpretation of an equational sentence is more difficult to hold. RF in this position, and presumably in the Post-DS position, is more appropriately analyzed as a sentence adjunct, analogous to reported speech markers observed in a number of other languages (cf. Aikehnvald 2004: 132, 284).

The second point concerns the potential indeterminacy between DS and IS, particularly in CM. CM texts, which are written in *Jawi* scripts, lack the quotation markers available in MI to indicate DS, which can cause indeterminacy between DS and IS when interpreting the original texts. This study follows the judgment of the transliterators who put quotation marks or colons for DS when interpreting DS in the texts.

2. Frequency of pre-DS frames and post-DS frames in CM and MI

In this section, we quantitatively demonstrate the degree to which the preferred DS encoding structures, summarized in Table 1 in the Introduction, appear in CM and MI respectively. We focus on sentences that include the four most frequently used and semantically neutral speech roots, namely (i) *kata* "say," (ii) *ujar* "say," (iii) *jawab* "answer," and (iv) *sahut* "answer," which serve as the DS frame head.[4] Frameless DS, prevalent in modern novels, is excluded from quantitative analysis due to difficulties in automatic corpus identification.

---

[4] The following list shows all the extracted verbs.

ujar: RF; *ujar, ujarnya,ujarku,* VF; *berujar*
jawab: RF; *jawab, jawabnya, jawabku,* VF; *dijawab, menjawab*
sahut: RF; *sahut, sahutnya, sahutku,* VF; *disahut, menyahut, disahuti, menyahuti*
kata: RF; *kata, katanya, katamu, kataku,* VF; *berkata, berkata-kata, mengatakan, dikata, dikatakan, mengata, katakan*

First, we examine CM texts, which feature the RF in the pre-DS position as the unmarked schema. We selected 17 titles, shown in Table 2, dated between the 1300s and the 1800s from the Malay Concordance Project (MCP) corpus (http://mcp.anu.edu.au/Q/mcp.html). In accordance with MCP's adoption of modern spelling, the excerpts presented in the subsequent sections of this article will also adhere to this convention.

Table 2. CM texts targeted in this study

	Period	Text title
1	1300s	Hikayat Bayan Budiman
2	1300s	Hikayat Raja Pasai
3	1500s	Hikayat Inderaputera
4	1600s	Sejarah Melayu
5	1700s	Hikayat Patani
6	1700s	Hikayat Hang Tuah
7	1800s	Hikayat Raja Bikrama Sakti
8	1810s	Hikayat Negeri Johor
9	1850s	Hikayat Siak
10	1860s	Tuhfat al-Nafis
11	1860s	Asal Keturunan Raja Barus
12	1860s	Salasilah Melayu dan Bugis
13	1860s	Hikayat Raja Damsyik
14	1870s	Cerita Patani dan Kelantan
15	1870s	Hikayat Indera Nata
16	1870s	Hikayat Nakhoda Asik
17	1880s	Hikayat Pahang

First, we present the RF and VF frequencies. Table 3 shows the overall occurrences of RF and VF, with RF constituting 88.2% of the cases.

Table 3. Frequencies of RF and VF in the CM texts

	Text title	RF	VF	Total
1	Hikayat Bayan Budiman	865	30	895
2	Hikayat Raja Pasai	97	7	104
3	Hikayat Inderaputera	592	84	676
4	Sejarah Melayu	521	46	567
5	Hikayat Patani	46	6	52
6	Hikayat Hang Tuah	1,316	183	1,499
7	Hikayat Raja Bikrama Sakti	355	34	389
8	Hikayat Negeri Johor	4	0	4
9	Hikayat Siak	91	80	171
10	Tuhfat al-Nafis	306	63	369
11	Asal Keturunan Raja Barus	96	39	135
12	Salasilah Melayu dan Bugis	200	89	289
13	Hikayat Raja Damsyik	260	20	280
14	Cerita Patani dan Kelantan	21	3	24
15	Hikayat Indera Nata	469	38	507
16	Hikayat Nakhoda Asik	216	4	220
17	Hikayat Pahang	82	12	94
	total	5,537 (88.2%)	738 (11.8%)	6,275

Table 4 shows the breakdown of the RF and VF frequencies in Table 3 by the root. We note that, first, of all four roots, lexemes from the ubiquitous root *kata* predominate, and second, in general, RF is consistently dominant, regardless of the root.

Table 4. Frequencies of RF and VF for each root in the CM texts

	Text title	*kata* RF	*kata* VF	*ujar* RF	*ujar* VF	*jawab* RF	*jawab* VF	*sahut* RF	*sahut* VF
1	Hikayat Bayan Budiman	719	27	9	0	10	0	127	3
2	Hikayat Raja Pasai	58	6	20	0	2	0	17	1
3	Hikayat Inderaputera	334	83	193	0	0	0	65	1
4	Sejarah Melayu	438	41	2	0	1	1	80	4
5	Hikayat Patani	46	6	0	0	0	0	0	0
6	Hikayat Hang Tuah	1,177	182	0	0	0	0	139	1
7	Hikayat Raja Bikrama Sakti	266	33	0	0	0	0	89	1
8	Hikayat Negeri Johor	0	0	0	0	4	0	0	0
9	Hikayat Siak	52	70	0	0	38	9	1	1
10	Tuhfat al-Nafis	154	54	0	0	152	9	0	0
11	Asal Keturunan Raja Barus	96	3	0	0	0	36	0	0
12	Salasilah Melayu dan Bugis	104	60	0	0	95	21	1	8
13	Hikayat Raja Damsyik	243	20	0	0	5	0	12	0
14	Cerita Patani dan Kelantan	15	3	0	0	5	0	1	0
15	Hikayat Indera Nata	454	37	0	0	0	0	15	1
16	Hikayat Nakhoda Asik	159	3	0	0	0	1	57	0
17	Hikayat Pahang	37	2	0	0	45	7	0	3
	Total	4,352	630	224	0	357	84	604	24

Table 5 presents the frequencies of the three potential RF positions: pre-DS frame, post-DS, and inserted. The pre-DS RF accounts for 99.4% of RF frames, resulting in 87.7% of all DS frames in the examined CM texts, indicating a high frequency. Concerning the VF position, all 738 instances were found in the pre-DS position.

Table 5. The frequencies of the RF positions relative to DS

	Text title	Pre-DS	inserted	Post-DS
1	Hikayat Bayan Budiman	865	0	0
2	Hikayat Raja Pasai	96	0	1
3	Hikayat Inderaputera	592	0	0
4	Sejarah Melayu	519	1	1
5	Hikayat Patani	46	0	0
6	Hikayat Hang Tuah	1316	0	0
7	Hikayat Raja Bikrama Sakti	353	0	2
8	Hikayat Negeri Johor	3	0	1
9	Hikayat Siak	91	0	0
10	Tuhfat al-Nafis	304	1	1
11	Asal Keturunan Raja Barus	76	0	20
12	Salasilah Melayu dan Bugis	198	1	1
13	Hikayat Raja Damsyik	259	0	1
14	Cerita Patani dan Kelantan	21	0	0
15	Hikayat Indera Nata	469	0	0
16	Hikayat Nakhoda Asik	216	0	0
17	Hikayat Pahang	81	1	0
	Total	5,505(99.4%)	4(0.1%)	28(0.5%)

Notwithstanding the general predominance of pre-DS RF, some texts deviate from this tendency. They include *Hikayat Siak*, *Tuhfat al Nafis*, and *Salasilah Melayu dan Bugis* in which the ratio of VF is high, and *Asal Keturunan Raja Barus* "The Descent of King Barus," in which the ratio of post-DS to pre-DS RF is relatively high. They are all works written in the late 19th century. We return to this point in Section 3.

Next, we examined DS encoding structures in MI based on text data from the web corpus Malindo Conc (https://malindoconc.lagoinst.info/concordance/ind/; Nomoto et al. 2018). Among the three sets of modern Indonesian data in the Malindo Conc, we employed the one titled web-ind, which includes the widest range of genres. The same set of lexemes was searched: *kata*, *ujar*, *jawab*, and *sahut*.

Table 6 shows the frequencies of VF and RF, including these lexemes. The ratio of VF to RF is significantly higher in MI than in CM.

Table 6. Frequencies of RF and VF in MI texts

RF	VF	Total
2,402 (68.4%)	1,110 (31.6%)	3,512

Table 7 shows the breakdown of the RF and VF frequencies listed in Table 6 by the root. The figures in the table show that among the four roots, *kata, ujar,* and *sahut* especially the latter, exhibit a high RF ratio, whereas the root *jawab* exhibits similar RF and VF ratios.

Table 7. Ratio of RF to VF

kata		*ujar*		*jawab*		*sahut*	
RF	VF	RF	VF	RF	VF	RF	VF
1,590	866	517	7	250	228	45	9

Table 8 shows the frequencies of the three positional options of RFs: pre-DS, inserted, and post-DS. More than 90% of the RFs occur in the post-DS position.

Table 8. Frequencies of the three positional options of RF

Pre-DS	Inserted	Post-DS	Total
134 (5.6%)	40 (1.7%)	2,228 (92.7%)	2,402

In contrast, almost all cases (97.7%) of VF occurred in the pre-DS position.

The quantitative analysis of DS encoding strategies in CM and MI conducted in this section reveals that (i) CM displays a strong predominance of pre-DS RF, accounting for 87.7% of all DS frames in the examined CM texts, and that MI exhibits two prevalent patterns of DS encoding structures: pre-DS VF (30.9%) and post-DS RF (63.4%), with the latter being more dominant. This confirms the stylistic changes mentioned in the Introduction: (i) the positional shift of RF from the pre-DS position to the post-DS position, and (ii) VF replacing RF in the pre-DS position.

3. The process of stylistic change: How did the change occur?

3.1 CM texts in the late 19th century

In this section, we investigate the process of stylistic change observed in the DS encoding strategies summarized in Section 2 based on specimens of text written and published in the transitional period from the late 19th to the early 20th century. The old orthography has been revised to a new one by the authors in all the cited texts. First, we will explore the atypical CM titles from the late 19th century mentioned in Section 2, then investigate Low-Malay novels from the late 19th century, and lastly delve into the so-called Balai Pustaka literature published in the early 20th century.

As mentioned in Section 2, some late 19th CM texts, namely *Hikayat Siak*, *Tuhfat al Nafis*, and *Salasilah Melayu dan Bugis* already exhibit deviations from the CM canonical pre-DS RF strategy, showing that the norm had collapsed in the late 19th century, at least to some extent. The common deviation observed in the three titles is the high ratio of VF, instead of RF. This is exemplified in Excerpt (6) cited from *Hikayat Siak*; the combination of VF and RF observed here is typical in *Hikayat Siak*.

(6)   *Maka Tengku Simbuk* [***berkata***]_{VF} *kepada baginda,* [*"Patik hendak keluar, tiadalah tertahan, lapar sangat."*]_{DS} *Maka* [***titah baginda***]_{RF}, [*"Baiklah, aku serahkan kepada Allah taala."*]_{DS}

So Tengku Simbuk said to his majesty, "Patik wanted to go out, he couldn't hold back, he was very hungry." So, the king said, "Alright, I leave it to Allah Almighty."

Asal Keturunan Raja Barus ("*Barus*" hereafter), written in 1866, exhibits another type of deviation, that is, a relatively high ratio of post-DS RF to pre-DS RF. In most cases, post-DS RF co-occurs with the pre-DS frame, which can be either VF or RF. Excerpt (7) illustrates this. Here, the DS is surrounded by two RFs.

(7)   [***Kata Sutan Ibrahim***]_{RF} [*"Tidak akan sekarang karena aku hendak melihat-lihat negeri orang,"*]_{DS} [***katanya***]_{RF}.

Sultan Ibrahim said, "Not now because I want to see the country." (Drakard 1988:181)

Excerpt (8) is another passage from *Barus*, where DS is encircled by RF frames ((a) and (c)), and both VF and RF occupy the pre-DS position within a single sentence ((b)).

(8)   (a)   *Dalam pada itu maka* [***bertanya***]_{VF} *orang punya dusun itu kepada Sutan Ibrahim* [***katanya***]_{RF}, [*"hai tuan-tuan yang lalu dari mana tuan-tuan datang kemari atau orang dari mana tuan?"*]_{DS} and [***katanya***]_{RF}.
      (b)   *Dalam pada itu maka* [***menjawab***]_{VF} *Sutan Ibrahim* [***katanya***]_{RF}, [*"Kamu orang mana?"*]_{DS}
      (c)   [***Kata orang dusun itu***]_{RF}, [*"Aku orang negeri di sini juga"*,]_{DS} [***katanya***]_{RF}.

      (a)   During that, asked the person who had the hamlet to Sultan Ibrahim saying, "Oh, gentlemen, so where did you come here or where are you originally from?" he said.
      (b)   Then Sultan Ibrahim answered saying, "Where are you from?"
      (c)   The villager said, "I am a countryman here too." (Drakard 1988:145)

In Excerpts (8)(a) and (b), the speech verbs *bertanya* "ask" and *menjawab* "answer" precede the pre-DS RF *katanya*. In this instance, the pre-DS and post-DS RF function similarly to quotation markers (" ") as punctuation symbols, marking the beginning and end of DS. Their primary role is to demarcate the DS domain, while the task of indicating the protagonists' actions is left to the preceding speech verbs. This suggests a type of semantic bleaching might be occurring.

The use of 'semantically bleached' RF, specifically for pre-DS RFs that signal the initiation of DS, is commonly observed in earlier CM texts. For instance, in Excerpt (9) taken from *Sejarah Melayu*, the verb *berseru* "shout" precedes the RF *katanya*.

(9) *Maka Sultan Ahmad pun tersenyum, lalu baginda pun kembalilah ke istana. Maka Peringgi pun* [***berseru***]$_{VF}$ *dari kapal,* [***katanya***]$_{RF}$, [*"Hei orang Melaka! Ingat-ingat kamu sekalian! Demi Dios, esok hari kami naik ke darat."*]$_{DS}$

Sultan Ahmad smiled and returned to the palace. Peringgi then called out from the ship, saying, "Hey Melaka people! Remember you all! By Dios, tomorrow, we go ashore." (SM 253:11)

This usage of the pre-DS RF, which is observed in Excerpts (9) above, could have been the origin of the post-DS RF, such as in Excerpts (8), (a), and (c). After the establishment of semantically bleached RFs marking the starting points, it is conceivable that similar-functioning RFs evolved to indicate DS endpoints. Consequently, two RF markers functioning akin to a pair of quotation marks delimiting the DS domain can emerge.

Nonetheless, the presence of this post-DS RF is confined to *Barus* in the context of late 19th-century CM texts. The idiosyncrasy of *Barus* could be associated with the geographical context in which the text was composed, specifically *Barus* in northwest Sumatra. While maintaining close relations, the region remained somewhat distant from the central hub of Malay literature (Drakard 1988:14). Such regional conditions could contribute to a minor deviation from the norm, resulting from the amalgamation of a newly emerging style.

The post-DS RF usage observed in *Barus* may have begun to gain widespread adoption in the late 19th century. This perspective is corroborated by another High Malay text from this period, Haji Ibrahim's letters and diaries. A renowned High Malay writer in Riau, Haji Ibrahim, maintained a close relationship with Dutch linguist Von de Wall. Their correspondence, as well as a diary entry discussing Von de Wall's passing, has been published (Van der Putten 2001). The style predominantly exhibits CM characteristics with a noticeable deviation in the post-DS RF. Excerpt (10), taken from Haji Ibrahim's 1867 letter, demonstrates the co-occurrence of post- and pre-DS RFs.

(10) *Maka serta sahaya baca di hadapan paduka Tuan Residen maka sahaya pun serahkan dibaca sendiri oleh sri Tuan Residen, lalu* [***katanya***]$_{RF}$ [*'baik Datuk balas surat tuan Von de Wall yang sahaya tiada menerima surat wissel dengan surat kepada paduka Engku Haji itu maka bungkusan baju hujan itupun tiada kita menerima,'*]$_{DS}$ [***katanya***]$_{RF}$.

Then I read (the letter) in front of the resident and (the resident) told me to read it (to him), then he said, "OK, you should answer to Von de Wall that I did not receive the letter and the money order for Raja Ali Haji, nor his rainwear," he said. (Van Der Putten 2001:119).

Excerpt (11) is cited from an entry in Haji Ibrahim's diary dated May 2, 1873. Here, the RF occurs only in the post-DS position.

(11) *Anaknya Tuan Adolf sampai dengan mil dua belas hari lalu, maka malam Sabtu kelemarin dia beri aku teraan cakap2 dan pantun, dan lagi kitab teraan keluar dari Arab.* [*'Ayahanda tak ada guna lagi'*]$_{DS}$, [***katanya***]$_{RF}$.

His son Mr. Adolf came by mail boat two weeks ago, and last Friday night he gave me the published conversations and pantuns, also some printed books from the Middle East. "Father has no need for them anymore," he said. (Van Der Putten 2001:78).

As will be observed in the subsequent section, this post-DS RF usage is commonly observed in late 19th-century Low Malay literature and eventually becomes the standard in the newly established Indonesian language of the 20th century.

3.2 Low Malay novels in the late 19th century

Around the same time that the rise of post-DS RF began to spread, the publication of literature in Low Malay, a language exhibiting the features of vernacular Malay, began, mainly in Batavia (Watson 1971, Adelaar 2000:235, Mahdi 2016).

Frequent use of VF is a characteristic of Low Malay novels, which would reflect colloquial language of the time. Excerpts (12) and (13) are cited from a Low Malay novel titled *Lawah-lawah Merah* "A read spider" published in 1871. The Pre-DS VF observed in Excerpt (12) was dominant throughout the novel.

(12) *Tuan Arthur lantas* [**berkata**]$_{VF}$: [*coba angkau pergi pangil Ming sopaiya dia juga bole turut liat disini sama-sama*!]$_{DS}$

Mr. Arthur then said, "let's try to go call Ming so that he can also see here together." (p. 67)

Many of early Low Malay narratives are the translated novel, and *Lawah-lawah Merah* is one of them (Lombard-Salmon 1979). Here, various novel techniques were adopted from their foreign counterparts. A prominent technique is the utilization of pre-DS VF and VF inserted within the DS. Excerpt (13) demonstrates post-DS VF and inserted VF structures. This strategy is unique to Low Malay and is absent in both CM and MI.

(13) ["*Apa ankau tiada kenal ini kipas*?"]$_{DS}$ [**bertanya**]$_{VF}$ *Fo-Hap, dengan paksa suru perempuan itu angkat mukanya*. ["*Tiada,*"]$_{DS}$ [**menyaut**]$_{VF}$ *Liou Siu,* [*tiada! Apakah ini huruf yang ada di kipas*.]$_{DS}$

"Do you not know this fan?" Fo-Hap asked and forced the woman to raise her face. "No," replied Liou Siu, "No! Are these letters on the fan?" (p. 8)

A slightly different type of variation in DS encoding strategy is observed in Low Malay narratives of a later period that were originally written in Malay. For example, in *Tjerita Njonja Kong Hong Nio* (Kommer 1900), both VF and RF that are either in the post-DS position or inserted are observed.[5] In Excerpt (14)(a) and (d), two post-DS frames are observed; the former is VF and the latter is RF, whereas in (b) and (c), frames inserted within the DS are observed; the former is RF and the latter is VF.

(14) (a) ["*Sudah pulang, cutak*?"]$_{DS}$ [**menanya**]$_{VF}$ *Kong Hong Mio*.
      (b) ["*Saya Nya*!"]$_{DS}$ [**sahut si Cutak**]$_{RF}$ *dengan suara rendah*. ["*Saya tau nyonya mau urusken lekas itu perkara, sebab itu saya sudah jalan terus tiada mampir di mana-mana*"]$_{DS}$

---

[5] *Tjerita Njonja Kong Hong Nio* was originally published by A. Veit & Co dan W. P. Vasques in 1900, but again, the authors employed the version included in Toer (1980).

(c) [*"Abis?"*]DS [***menanya***]VF *Kong Hong Nio*, [*"apa angkau sudah ketemu orangnya?"*]DS

(d) [*"Orangnya tiada, Nya cumah iparnya saya sudah ketemu,"*] [***jawab si Cutak***]RF.

(a) "Are you home yet, Cutak?" Asked Kong Hong Mio.
(b) "Yes," replied Cutak in a low voice, "I know that the lady wants to take care of the matter quickly, that's why I've been walking without stopping anywhere."
(c) "Then, have you found the person?" asks Kong Hong Nio.
(d) "The person is gone, but I have met her brother-in-law," replied Cutak.

Excerpt (15), also cited from *Kong Hong Nio*, exhibits another novel technique: successions of frameless DS, which continues to be a common strategy in novels until today.

(15) (a) *Sasudahnya marika itu duduk di satu batang kayu, maka berkata Liem Hok Kan pada bapa Udin.*
(b) *"Bapa Udin, kombali aku musti minta tutunganmu."*
(c) *"Baiklah, baba, tapi inget baba ampunya perjanjian iang dulu baba belon juga sampeiken."*
(d) *"Kau jangan kuwatir, itu perkata saya tiada lupa, ambil dulu ini sedikit uwang. Sekarang aku mau minta lagi pertulunganmu."*
(e) *"Aku ada perlu mau dapet racun, maka angkau musti cariken itu buat aku."*
(f) *"Nanti saya cari, baba," kata bapa Udin.*

(a) After they sat down on a log, Liem Hok Kan said to Mr. Udin :
(b) "Mr. Udin, I must ask for your help again."
(c) "Alright, but remember that you have a promise that you haven't fulfilled yet."
(d) "Don't you worry, I haven't forgotten the word, take this little money, first. Now, I want to ask for your help again."
(e) "I need to get poison, so you must find it for me."
(f) "I'll find it," said father Udin.

3.3 Balai Pustaka literature from the early 20th century

The rise of Low Malay publications, examples of which are captured above, was considered undesirable by the Dutch government. Adelaar (2000:289) summarizes this policy.

> In the 20th century, the Dutch East Indian government became concerned about the popularity of the Indonesian press. It disapproved of its non-standard language and found its content sometimes subversive. The government's efforts to standardize Malay culminated in the foundation, in 1908, of a language planning bureau later called *Balai Pustaka* ("Literature Bureau"). (Adelaar 2000:235)

Balai Pustaka did contribute to standardization, however, DS encoding did not simply apply to CM norms, it developed a new standard. The authors of *Balai Pustaka* novels actively employed modern novelistic techniques, such as the post-posed frame, that had

been adopted in Low-Malay novels. The use of the VF in the pre-DS position is also a characteristic of the new standard. Excerpts (16) and (17) are cited from Merari Siregar's 1920 novel *Azab dan Sengsara* "Torment and Misery", which is considered the first *Balai Pustaka* novel. In (16), the post-DS RF is employed, whereas in (17), the pre-DS VF and inserted RF are employed.

(16) (a) ["*Mak, ampun Mak!*"]$_{DS}$ [**kata si anak**]$_{RF}$ *seraya melompat memeluk ibunya itu. Akan tetapi air matanya makin lebat bercucuran ke pipinya.*

   (b) ["*Apakah yang anakku tangiskan, sedang jauh malam begini? Pikirku Riam sudah tertidur,*"]$_{DS}$ [**kata si ibu**]$_{RF}$ *dengan suara perlahan-lahan.*

   (a) "Mak, forgive Mak!" said the child jumping up to hug her mother. But the tears were pouring down her cheeks.

   (b) "What is my child crying about, this far into the night? I thought your Riam was asleep," said the mother in a slow voice. (p. 22)

(17) (a) *Ia pun* [**berkata**]$_{VF}$ *seraya menghampiri anak perempuan itu,* ["Sebenarnya saya sudah letih, Riam, tengoklah beratnya pekerjaan itu, tetapi …"]$_{DS}$

   (b) ["*Saya pun tahu juga, angkang sudah payah; itulah sebabnya angkang kusuruh pergi berhenti ke pondok itu*"]$_{DS}$, [**kata Mariamin**]$_{RF}$ *dengan lekas mendahului perkataan Aminu'ddin,* ["*Tetapi kalau laki-laki biasa juga letih oleh pekerjaan, bagaimanakah perempuan, manusia yang lebih lemah, yang tiada mempunyai daya dan kekuatan sebagai laki-laki?*"]$_{DS}$

   (a) He said as he approached the girl, "Actually I am tired, Riam, look at the weight of the work, but ..."

   (b) "I know too, Angkang is tired; that's why I told Angkang to stop at the hut," Mariamin said quickly, preceding Aminu'ddin's words, "But if ordinary men are also exhausted by work, what about women, weaker human beings, who do not have the power and strength of men?" (p. 37)

Excerpts (18) and (19) are cited from *Sengsara Membawa Nikmat* (St. Sati 1936), another representative novel in what is called *Balai Pustaka* literature. Again, the pre-DS VF and inserted RF were observed.

(18) *Ia membawa sebuah surat kabar mingguan pada tangannya. Maka anak itu pun* [**berkata**]$_{VF}$ ["*Papa apa ini?*"]$_{DS}$

   He carried a weekly newspaper in his hand. Then, the child said, "Dad, what is that?" (p. 184)

(19) ["*Tidak, Bapak!*"]$_{DS}$ [**jawab Midun**]$_{RF}$ *dengan ketakutan.* ["*Pikiran saya masih sehat; ayah dan Bapak Pendekar ada di surau kecil di sebelah.*"]$_{DS}$

   "No," Midun answered fearfully, "My thought is still sound, the father and Mr. Pendekar are in the small prayer house next door." (p. 26)

Through Balai Pustaka publications, Minangkabau writers and editors played a crucial role in establishing a new standard (Teeuw 1972:120, Adelaar 2000:236). The adopted DS encoding strategy is thought to reflect the variety of formal registers used in Sumatra at that time. As mentioned above, the frequency of pre-DS VF was high in some CM texts from the late 19th century, and post-DS RF appears in late 19th century CM texts; the latter is considered to have been common up to that point. In addition, we can speculate that at that time, the pre-DS RF, which was canonical in the CM narrative, was considered too classical or archaic to be employed as the standard style, especially after many readers had gained experience with Low Malay novels in the colloquial style. As the *Balai Pustaka* standard language became more widespread, one of the markers of the Low Malay style disappeared from the written language, namely, the post-DS VF observed in Excerpt (13).

With the establishment of the norm where the VF appears before DS and the RF in other positions, the idiosyncrasy of the DS encoding strategy observed in CM grammar, as mentioned in Introduction, has become less pronounced. This is because the Pre-DS VF allows the structure to be consistent with overall grammar at the basic clause structure level, while the RF, which constitutes an idiosyncratic construction, functions as an adjunct that does not play a core role in the clause structure. It is unclear whether the authors who established the norms for Balai Pustaka were consciously aware of this. However, setting a more unmarked structure pattern by diverging from classical expressions may have been one reason for adopting the pre-DS VF as a new norm.

The change in DS encoding strategies from CM to MI can be related, in a larger context, to the change of medium from oral or aural to written or visual. It is widely accepted that CM narratives have their origins in oral literature passed down from generation to generation (Sweeney 1987, Derks 1996). In the context of oral literature, the CM narrative style is constrained by linearity, necessitating preceding frames for DS to avoid confusion. MI texts designed for reading are linear and utilize visually recognizable punctuation. Oral literature's mnemonic patterns facilitate storytelling (Ong 1982, Sweeney 1987), while MI authors are liberated from these patterns and can creatively convey characters' speech through the choice from pre-DS VF, post-DS or inserted RF, and frameless DS observed in novels.

4. Summary

In this paper, we have investigated the stylistic change in the DS encoding strategy from CM to MI. The DS frame can be divided into two structural types: RF, comprising a root followed by the speaker NP, and VF, led by derived verbs. As quantitatively demonstrated in Section 2, RF is prevalent in both CM and MI; however, its position differs. In CM, RF precedes DS, whereas, in MI, it follows DS or is inserted within DS. The change, albeit more incidental in terms of frequency, was accompanied by VF replacing the RF in the pre-DS position. In addition, although not quantitively verified, frameless DS is frequently observed in MI novels. Thus, overall, the change from CM to MI can be seen as the collapse of the highly schematized CM norm of pre-DS RF and the diversification of the DS encoding strategy. The post-DS RF observed in late 19th-century atypical CM texts, such as *Asal Keturunan Raja Barus* and Haji Ibrahim's letter and diary, suggests that this strategy was becoming common in some types of written language at this time. Additionally, the frequent use of VF in the pre-DS position appears in late 19th-century CM texts such as *Hikayat Siak*, *Tuhfat al Nafis*, and *Salasilah Melayu dan Bugis*, indicating a similar trend. The diversification of DS encoding strategies is also evident in contemporaneous low-Malay novels. Dominated by pre-DS VF, likely originating from

Batavian colloquial variations, Low Malay authors also employed novelistic techniques such as post-DS or inserted VF/RF and frameless DS, reflecting Western novel influences. The Balai Pustaka, established by the Dutch government in the early 20th century, adopted a Malay variety that incorporated modern novelistic techniques from Low Malay novels. Consequently, the new standard embraced post-DS or inserted RF, pre-DS VF, and frameless DS, which persist today, while post-DS VFs from Low Malay novels became obsolete. This change might have lessened the syntactically idiosyncratic status of DS encoding constructions observed in CM. Table 9 summarizes the potential frame types in CM, Low Malay novels, Balai Pustaka novels, and Modern Indonesian.

Table 9. Preferred DS encoding strategies in each text type

	CM	Late 19th Low Malay novels	Early 20th Balai Pustaka novels	MI
Pre-DS RF	+	−	−	−
Pre-DS VF	−	+	+	+
Post-DS/ inserted RF	−	+	+	+
Post-DS/ inserted VF	−	+	−	−
Frameless DS	−	+	+	+

In a broader context, these differences may be attributed to the distinct media through which the texts were transmitted when their respective structures were developed: oral and aural for one and written and visual for the other. CM texts, rooted in oral literature passed down through generations, require a highly schematized style with easily memorable mnemonic patterns. Conversely, MI authors benefit from diverse DS encoding strategies, including frameless DS, facilitated by the visual assistance of punctuation and quotation marks.

This study has solely concentrated on the structural aspects of DS-encoding strategies and excluded the stylistic effects resulting from the authors' choice of strategy in the pre-modern and modern literature. As Djenar (2018) elaborates, in contemporary literature, authors utilize frames not merely for identifying the DS speaker but also as a device to evoke specific emotions in readers. An in-depth examination of such effects will be the focus of our future research.

Abbreviations

CM	classical Malay	NP	noun phrase
DS	direct speech	RF	root frame
IS	indirect speech	VF	verb frame
MI	modern Indonesian		

References

Adelaar, K. Alexander. 2000. Malay: A short history. *Oriente Moderno* 19(80)(2). 225–242.

Aikhenvald, Alexandra Y. 2006. *Evidentiality*. Oxford: Oxford University Press.

Cumming, Susanna. 1991. *Functional change: The case of Malay constituent order*. Berlin: Mouton de Gruyter.

Derks, Will. 1996. "If not to Anything Else," Some Reflections on Modern Indonesian Literature. *Bijdragen tot de Taal-, Land- en Volkenkund* 152. 341–352.

Djenar, Dwi Noverini. 2018. The pragmatics of reported speech. Paper presented at the Konferense Linguistik Tahunan Atma Jaya 16, 10–12 April 2018.

Drakard, Jane. 1988. *Sejarah Raja-Raja Barus: Dua Naskah dari Barus*. Jakarta: Penerbit Angkasa and Ecole française d'Extrême-Orient.

Hinmmelmann, P. Nikolaus. 2005. The Austronesian languages of Asia and Madagascar: typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. London and New York: Routledge.

Leech, G. N. & M. H. Short. 1981. *Style in fiction*. London: Longman.

Lombard-Salmon, Claudine. 1979. Lawah-lawah merah ou "Araignée rouge": un "roman sino-malais" qui s'avère être la traduction d'un roman français. *Archipel* 17. 63–69. doi: https://doi.org/10.3406/arch.1979.1457

Mahdi, Waruno. 2016. Linguistic variety in later nineteenth-century Dutch-edited Malay publications. *Nusa* 60. 107–185 [Permanent URL: http://hdl.handle.net/10108/87446]

Nomoto, Hiroki, Shiro Akasegara & Asako Shiohara. 2018. Reclassification of the Leipzig Corpora Collection for Malay and Indonesian. *NUSA* 65. 47–66.

Ong, Walter J. 1982. *Orality and literacy: The technologizing of the word*. London/ New York: Methuen.

Sweeney, Amin. 1987. *A full hearing: Orality and literacy in the Malay world*. Berkeley: University of California Press.

Teeuw, A. 1972. The impact of Balai Pustaka on modern Indonesian literature. *Bulletin of the School of Oriental and African Studies, University of London* 35(1). 111–127.

van der Putten, Jan. 2001. *His word is the truth: Haji Ibrahim's letters and other writings*. Leiden: Research School of Asian, African, and Amerindian Studies, Universiteit Leiden.

Watson, C. W. 1971. Some preliminary remarks on the antecedents of modern Indonesian literature. *Bijdragen Tot de Taal-, Land- En Volkenkunde* 127(4). 417–433.

Wijaya, Juliana. 2016. Direct reported speech in spoken Indonesian: Verbal markers and discourse practices. *Jakarta: Linguistik Indonesia*, February 2016. 1–14.

Source Materials

Author unknown. 1875. *Lawah-lawah Merah,* Batavia: Ogilvie & Co.

Francis, G. 1896. *Njai Dasima*. (Mistress Dasima) Pramoedya Ananta Toer (ed.) 1982. *Tempo Doeloe* (The old days). 225–247.

Jones, Russel. 1987. *Hikayat Raja Pasai*. Selangor: Fajar Bakti Petaling Jaya.

Kommer. H. 1900. *Cjerita Njonja Kong Hong Nio*. (*The story of Ms. Kong Hong Nio*.) Pramoedya Ananta Toer (ed.) 1982. *Tempo Doeloe* (The old days). 251–313.

Siregar, Merari. 1920. *Azab dan Sengsara* (Torment and Misery). Jakarta: Balai Pustaka. (Cited from the 30th version published in 2011)

Sati, Tulis Sutan. 1929. *Sengsara Membawa Nikmat* (Misery Brings Favor). Jakarta: Balai Pustaka. (Cited from the 7th version published in 2000)

Toer, Pramoedya Ananta. 1980. *Bumi Manusia*. Jakarta: Hasta Mitra. (Cited from the 20th version published from Lentera Dipantara, Jakarta published in 2015.)