Résumé

À travers les siècles certaines langues ont fonctionné comme des langues savantes, c’est-à-dire des langues enseignées dans les écoles pour leur valeur culturelle intrinsèque, mais sans pour autant constituer une langue de communication. Dans le monde islamique, l’arabe a longtemps servi comme langue savante, dont l’acquisition se faisait sous la forme de mémorisation de textes joints à leur traduction dans la langue vernaculaire, qui servait comme langue d’instruction dans les écoles. Dans le présent article, nous prenons comme point de départ la position de l’arabe en Asie du Sud-Est, où le malais, tout en fonctionnant comme langue d’instruction dans les écoles, fut choisi par l’administration coloniale comme langue intermédiaire dans sa communication avec la population indigène. Par conséquent, ce furent les Anglais et les Néerlandais qui publièrent les premières grammaires de cette langue vernaculaire, composées dans le cadre de la linguistique européenne. Les premières descriptions du malais fondées sur un modèle linguistique arabe n’apparurent qu’à la fin du xixe siècle. Le représentant principal de cette tradition linguistique est Raja Ali Haji (m. probablement en 1873). Dans son traité Bustān al-kātibīn, il emprunta le modèle de la tradition grammaticale arabe afin de composer une esquisse de la structure du malais, dans laquelle il se servait en partie de la terminologie grammaticale malaise qui avait été développée dans le système scolaire traditionnel pour l’étude de la grammaire arabe et l’exégèse coranique.

Mots-clés

ˀĀjurrūmiyya, calque, emprunt lexical, grammaire, grammaire étendue, Indonésie, langue savante, malais, Malaisie, Raja Ali Haji, terminologie, tradition grammaticale arabe.

Abstract

Throughout history, a number of languages have achieved the status of learned language, i.e., a language included in the curriculum of an educational system without yielding any communicational benefits. In large parts of the Islamic world, Arabic was (and still is) such a learned language. Acquisition of the learned language took place through the memorization of texts, with instruction and/or translation in vernacular languages. The vernacular languages themselves were not deemed to be in need of grammatical description, which explains why grammars for them were late to be developed. The present paper focuses on Malay, the lingua franca of choice in Southeast Asia for both Muslim missionaries and British and Dutch colonial administrators, while serving as the auxiliary language in the Islamic curriculum. The first grammars of Malay were published by the British and Dutch. Malay grammars written by native speakers did not make their appearance until the nineteenth century. Their main representative is Raja Ali Haji (d. probably 1873). In his Bustān al-kātibīn, he used the grammatical framework of Arabic grammar for a grammatical sketch of Malay, using in part the Malay terminology that had been developed in traditional education for the study of Arabic grammar and Qur’anic exegesis.

Keywords

ˀĀjurrūmiyya, Arabic grammatical tradition, extended grammar, grammar, Indonesia, learned language, loan translation, loanwords, Malay, Malaysia, Raja Ali Haji, terminology.
1. Learning a Learned Language

The reasons for choosing a learned language and the mechanisms by which this learned language becomes entrenched in an educational system have been studied extensively, for instance by Ostler (2005) and Hamel (2005). Learned languages may be useful for any number of religious, political, or scholarly reasons, but are typically not used for everyday communication. Historical examples include Sumerian in Akkad, Latin in Medieval Europe, Greek in Classical Rome and in the Hellenistic world, Sanskrit in Southeast Asia, Chinese in East Asia, Persian in the Mughal and Ottoman empires, and Arabic in large parts of the Islamic world.

Those who know a learned language become members of an elite, who can afford to send their children to the right schools where the learned language occupies a central place. For practical reasons, a vernacular language often serves as the medium of instruction. In the Arabic-speaking world, for instance, Berber was used in schools in the Sous region as a language of instruction for beginning students. Referring to the Qurʾānic verse (Q. 14/4) that God has sent messengers to each community in their own language, the Berber poet Muḥammad Aẓnag (16th century CE) expresses this as follows: “Arabic is the language of the Seal of the Prophets/A Berber will understand it in Berber” (lʼarabi iluqa n ḥātimu lanbiyya/yans igan amaziğ ifhm l s lmazğiy-i) (van den Boogert 1997: 49).

Teaching a learned language takes place largely on the basis of texts, which are memorized together with their translation in the vernacular language. The translation is often word-by-word, creating a linguistic variety that Riddell (2002) calls “translationese”. Eickelman (1978) has analyzed how such a system of learning affects the selection of texts and the establishment of a canon in all disciplines, including grammar. For most students, linguistic knowledge was determined by a few grammatical treatises learnt by heart together with their translation.

The translating tradition created a stock of technical terms in the vernacular language. The teaching of Latin in the abbey of Saint Gall, for instance, took place with Old High German as an auxiliary language. As a result, a host of Old High German loan translations of Latin terms entered the language (Grotans 2006). In other cases, the Latin terminology was taken over wholesale in the form of loanwords. In the Irish tradition, two sets of terminology were used, one with mostly Irish loan translations, intended for an audience of students and learners, the other with mostly Latin loanwords, used by scholars (Poppe 1999). In much of the Islamic world, the Arabic tradition was followed faithfully and its technical

---

1 I wish to thank David Bertaina (University of Illinois Springfield) for kindly sending me a copy of his article on Elias of Nisibis, and Jan van der Putten (Universität Hamburg) for his help in obtaining some of the Indonesian material. Thanks are also due to an anonymous reviewer for some corrections and a useful reference to the Kāfiya as a possible source.
lexicon predominated, so that Arabic loanwords constituted the bulk of new technical terminology, with only a relatively small number of loan translations, as in Persian and Turkic grammar. In this respect, the Malay tradition was an exception, since, along with Arabic loanwords, it frequently used loan translations.

The coexistence of two languages in education is likely to provoke a comparison of their structure and qualities. The vernacular language is usually seen as a poorer variety, more lively, perhaps, and better suited to the expression of feelings, but not as rich or rational as the learned language, and certainly not deserving a grammatical analysis of its own. Most people speak the vernacular language as their mother tongue anyway and do not feel the need for a grammar. Yet, in some cases, scholars took the initiative to write a grammar of their vernacular language, for which they depended on the grammatical framework of the learned language. In Europe, Latin, the language in which all religious and scientific literature was written (Waquet 1998), provided the framework for all grammar study. The universality of its structure was taken for granted, even when people were aware of the differences with their own vernacular.

The full emancipation of the vernacular language became possible only when the socio-political context changed. In Europe, the triggering factor was the establishment of political entities that focused on their identity as a nation with a language of its own. After the Reconquista had been completed, the Spanish kings strove at founding an empire with one ruler, one religion, and one language. The publication of Nebrija’s grammar of Castilian in 1492, the first of its kind in Europe, showed that this new imperial language was just as capable of expressing anything as the prestige language, Latin. Nonetheless, the linguistic framework in which this newly emancipated language was described remained the Latin one.

In the Arabic-speaking world, too, vernacular languages were generally seen as inferior to Arabic and not deserving a grammar of their own. Exceptions to this view are few and far between. In a discussion about the Syriac language that took place in 364/1026 between the Christian bishop Elias of Nisibis and the vizier ʿAbū l-Qāsim al-Ḥusayn ibn ʿAlī, the Syrian bishop vigorously defended the superior quality of the Syriac language (Bertaina 2011): Not only is it possible to say in Syriac whatever you can say in Arabic, but it is even superior. When Elias is asked by the vizier whether Syriac distinguishes between the agent and the object by means of case endings (ʾiʿrāb), he explains that in Syriac a particle, l-, is used for the object, at least in those cases where confusion is possible. This is much better than the Arabic system, he says, which breaks down in words with virtual endings, like ʿdaraba ʿisā mūsā “Īsā hit Mūsā’. Elias states that the ʾiʿrāb fails in its main function, i.e. removing ambiguity (labs, ititbās), not only between fāʿiḥl (agent) and mafʿiḥl (object), but also between sentence types, such as assertion
and question (*Majālīs*: 116-118). In spoken Arabic and Syriac, such distinctions are expressed by intonation (nağamāt al-ṣawt; *ibid.*: 113.8), but while in written Arabic there is no way to indicate them, Syriac scholars have devised a system of interpunction (*ʿalāmāt*) for this purpose (*ibid.*: 122).

The pride Elias takes in the Syriac language may have had something to do with the fact that before the emergence of the Arabic tradition Syriac possessed a grammatical tradition of its own. In other cases, language communities developed their own tradition in Islamic times. The revival of Coptic and the emergence of Coptic grammars in the thirteenth century CE, for instance, came at a time of social and cultural emancipation of the Copts, when Coptic scholars aimed to demonstrate the rich texture of their language (Sidarus 2010). A similar apologetic aim is found in the dictionary of the Turkic languages by al-Kāšgarī (11th century CE). In his *Dīwān luġāt al-Turk*, he intended to show that the Turkic languages were just as capable as Arabic to express complex thoughts, with an equally rich lexicon (Ermers 1999: 17).

While Syriac could not qualify for ḵāriĝhood because it does not have endings, some languages do have declensional endings. In Persian, the direct object suffix -rā was interpreted by Persian grammarians as a *harf-i taḥṣīṣ* ‘particle of specification’, until later grammarians named it ʿalāmat-i mafʿūl ‘marker of the object’ (Jeremiás 1997: 174 & 180, n. 20). In Turkic, the ending -nī of the direct object was called by some grammarians ʿalāmat al-naṣb ‘marker of the accusative’ (Ermers 1999: 163-284). The Persian and Turkic endings received the Arabic name of the function they expressed (*naṣb, mafʿūl*), but were not analyzed as instances of ḵāriĝ: they expressed the same syntactic categories as the ones marked by ḵāriĝ in Arabic, but used different tools to mark them. This approach parallels the description of European vernacular languages within a Latin framework: their lack of declensional endings was compensated by applying Latin case names to the syntactic functions of words.

Since the pioneers who started to focus on the vernacular language had been trained in the powerful linguistic model of the learned grammatical tradition, which had the status of a universal grammar, valid for all languages, they had to find a way to accommodate the structure of their own language to this model. In the next section, we shall see how this played out in the description of Malay in Southeast Asia.

### 2. MALAY AS AN AUXILIARY LANGUAGE

Arabic was introduced in Southeast Asia in the wake of Islamization through the thirteenth-fifteenth centuries CE as the written language of the Qurʾān and the Islamic sciences. The language of instruction in the educational system was
Malay, the lingua franca of the Malaccan peninsula and the Indonesian islands, for which Arabic script (Jawi) was introduced. Arabic texts were memorized, supported by Malay translations. Grammar as a topic in the schools focused on Arabic, while there was no formal teaching of Malay. During the colonial period from the seventeenth century CE onward, Malay was the language of choice of the British and Dutch administrators, who used it for communication with the indigenous population. They were the first to write grammars of the language for the purpose of training colonial officers, within a Western linguistic framework and terminology (Kaptein 2000).

When at the end of the nineteenth century a few Malay scholars started to describe their own language, they did not employ the colonial model, but that of Arabic grammar, which until then had been used in Southeast Asia exclusively for the study of Arabic. Yet, since the teaching method in all Islamic disciplines, including grammar and Qurʾānic exegesis, was based on Malay translations, an indigenous Malay terminology had gradually been developed (Riddell 2017). Thus, when these pioneers started to write a grammar of their own language, they did not need to develop all of its terminology from scratch, but could draw on an existing stock of Malay terms (Versteegh 2019).

The first and foremost Malay-written grammar of Malay was that of Raja Ali Haji (probably d. 1873), a scholar from the Sultanate of Johor-Riau. At the sultan’s court, a form of Classical Malay had become the cultural language shared by courtiers and Dutch colonial administrators alike. In his Bustān al-kāṭībīn [Garden of Writers] (1857), Raja Ali Haji provides a grammatical analysis of this Malay,2 while his incomplete dictionary Kitab pengetahuan bahasa [Book of the Knowledge of the Language], written in the 1850s, aims at exploring the lexicon of the language. It opens with a compendium of Arabic grammar (Pengetahuan: 1-27) and features a large number of Arabic grammatical terms (van der Putten 2002).

One of the factors motivating Raja Ali Haji to engage in this project was the deficient performance of many speakers, for whom Malay was a second language, used for interlanguage contact and for dealing with the colonial administration. He explains that many people make mistakes while writing Malay:

I have met many people, either Malays from Johor, or other Malays, not to mention people from other nations, who did not have Malay as their own language and

---

2 See Kridalaksana (1991); Kaptein (2000); Hidayatullah (2012); Mustari (1998). In quotations from the Bustān, I use, with slight changes in the transliteration of the Arabic words, the transliterated edition by Hashim bin Musa (2005), which also contains a facsimile of the Leiden ms. Kl 107. The manuscript used by van Ronkel (1901) for his Dutch translation of the text, Leiden no. 218, is identical to the lithograph edition of Pulau Penyengat (Riau) [1851], but it has one additional paragraph about separate writing of words and one about ḥārāb; both are missing in the other manuscripts (see van Ronkel 1901: 533 & 550).
whose faulty practice even reached the point where they started to teach each other, while believing themselves to be experts in this science (Telah banyaklah aku dapat orang Melayu Johor atau Melayu yang lainnya, apalagi orang yang lain-lain bangsa daripada bangsa yang bukan bahasa Melayu itu bahasa dirinya, hingga berpanjanganlah amalnya yang palsu itu hingga mengajarlah setengahnya akan setengahnya padahal ia menyangkakan akan dirinya ahli pada ilmu itu; Bustān: 6).

This is a recurrent theme in his treatise: as people make mistakes in writing or speaking Malay, their work becomes worthless. What is worse, they are not even aware of the fact that their Malay is faulty. Since a number of letters of the Arabic alphabet are not used in spoken Malay, people err by adding or omitting letters (ibid.: 13), or by writing the wrong ligatures (ibid.: 14). This shows that they possess only “stolen knowledge” (ilmu yang dapat dicurinya; ibid.: 14.18f.). The concluding chapter of the treatise (ibid.: 48-51) touches upon yet another aspect of the author’s didactic ambition in writing his grammar. It is dedicated to the art of epistolography and includes precise instructions on the proper way to address people of different rank, urging the reader to use a polite style of Malay (ibid.: 49).

Like the Coptic, Turkic, and Syriac grammarians, Raja Ali Haji describes his own language on the premiss that anything said in Arabic can also be said in Malay. Both languages share an underlying structure that is identical with the model of the Arabic grammarians. He does not aim at a full-scale grammatical description of the Malay language, but only at a short instruction for those wishing to improve the quality of their writing and speaking. Yet, some rules are necessary: Nonetheless I have, where necessary, borrowed explanations, analogies, and methods from the science of the Arabs, such as šarf ‘morphology’, nahw ‘syntax’, luğa ‘lexicography’ and others from the scientific literature, to compose this book... (Syahadan sungguhpun yang demikian itu tempatnya aku mengambil petua dan kias dan jalan memperbuat kitab ini, iaitu ilmu Arab juga seperti šarf dan nahw dan luğa dan lainnya daripada kitab ilmu ...; Bustān: 24.33-25.1).

And, indeed, he freely uses Arabic grammatical terminology and, in some sections, snippets of linguistic theory taken from the Arabic grammatical literature. The main source for his description is the ‘Ajurrūmiyya and its commentaries, as is clear from the parallels in treatment. In addition, he must have been familiar with the ‘Alfiyya tradition. Since the elementary grammatical treatises and commentaries are highly interdependent, it is not always possible to pinpoint the

3 In the East Indies, the most popular commentaries on the ‘Ajurrūmiyya were those by al-‘Azharī (d. 905/1499), al-Kafrawī (d. 1207/1787), and ‘Aḥmad ibn Zaynī Dahlān (d. 1304/1886), a contemporary Meccan scholar (van Bruinessen 1990). I do not know whether al-Šīrīnī’s (d. 977/1569) extensive commentary Nūr al-sajīyya was used in Southeast Asia. An important commentary on the ‘Alfiyya is that by Ibn ‘Aqīl (d. 769/1367). Another popular treatise was Qatr al-nadā wa-ball al-sadā with autocommentary by Ibn Hišām (d. 761/1360).
exact source. Viain’s (2014) detailed comparison of the structure of the ʿAlfiyya (2014: 228-253 504-508) and the ʿĀjurrāmiyya (2014: 254-259, 509-511) points out the following differences between them:

i. Both texts have a rough division into nouns and verbs, but the ʿĀjurrāmiyya first treats general rules of ʿirāb, while the ʿAlfiyya first deals with a number of other constructions (relatives, numerals, ḥikāya).

ii. The ʿĀjurrāmiyya starts with the verb, the ʿAlfiyya with the noun; for the noun, both traditions follow a division into marfūʿāt, manṣūbāt, mafṣūdat.

iii. Within the nominative constructions, the ʿĀjurrāmiyya starts with the ʿāl, the ʿAlfiyya with the mubtadaʿ.

iv. Within the accusative constructions, the ʿĀjurrāmiyya begins with the mafʿūl bihi, followed by the mafʿūl muṭlaq, while the ʿAlfiyya has the reverse order.

v. Within the satellite constructions, the order in the ʿĀjurrāmiyya is ʿāl - taʿkīd - ʿatf - badal; in the ʿAlfiyya tradition taʿkīd and ʿatf switch places.

A look at the chapters of the Bustān (Table 1) shows that sometimes their order reflects that of the ʿĀjurrāmiyya, for instance by treating first the mafʿūl bihi and then the mafʿūl muṭlaq, but at other times it follows more closely that of the ʿAlfiyya tradition, for instance in treating the noun before the verb, and the mubtadaʿ before the ʿāl. Within the chapters, the relationship with the ʿĀjurrāmiyya commentaries is evident, yet it appears that the author inserted remarks from an amalgam of different traditions. Obviously, as a nineteenth-century Muslim intellectual he must have been familiar with commentaries from both traditions.

### Table 1

#### Chapters of the Bustān

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>§1</td>
<td>introduction</td>
<td>3-11</td>
</tr>
<tr>
<td>§2</td>
<td>letters</td>
<td>12</td>
</tr>
<tr>
<td>§3</td>
<td>alphabet</td>
<td>12f.</td>
</tr>
<tr>
<td>§4-9</td>
<td>Jawi spelling</td>
<td>13</td>
</tr>
<tr>
<td>§10</td>
<td>orthography</td>
<td>13-25</td>
</tr>
<tr>
<td>§11</td>
<td>parts of speech</td>
<td>25f.</td>
</tr>
<tr>
<td>§12</td>
<td>ʿāl</td>
<td>29-32</td>
</tr>
<tr>
<td>§13</td>
<td>ḫarf</td>
<td>32-38</td>
</tr>
<tr>
<td>§14</td>
<td>kalām/kalima</td>
<td>38f.</td>
</tr>
<tr>
<td>§15</td>
<td>mubtadaʿ/ḥabar</td>
<td>39-42</td>
</tr>
<tr>
<td>§16</td>
<td>types of sentence</td>
<td>42</td>
</tr>
<tr>
<td>§17</td>
<td>ʿāl/ʿāl</td>
<td>42f.</td>
</tr>
<tr>
<td>§18</td>
<td>mafʿūl bihi</td>
<td>43</td>
</tr>
</tbody>
</table>
There is one aspect in which Raja Ali Haji’s work differs from the commentaries in the Arabic tradition: his preoccupation with spelling and orthography. Spelling rules did not form part of Arabic grammar, but were crucial in Southeast Asia because of the idiosyncrasies of the Jawi script. Raja Ali Haji’s discussion of ˁiˁrāb (Bustān: 20f.) starts with the observation that it consists of a stroke (baris) above or underneath the letter, indicating a vowel, in combination with a letter wāw, yāʾ, or ʾalif. This clearly addresses the representation of vowels in Jawi script, rather than declension, which is inexistent in Malay. But then, confusingly, he continues by defining ˁiˁrāb as “changes in the endings of speech due to the difference in governing words, overt or implicit” (mā taqayyara ʿawāhir al-kalām li-ḥtiilāf al-ʾawāmil al-dāḥila ʿalayhā lafẓan ʾaw taqādiran; ibid.: 20.14f.). However, he says, the rules of ˁiˁrāb are very difficult and “it is not my purpose to translate the ʾilm al-naḥw, but only to set up rules for writing and speaking Malay” (bukan maksudku hendak menterjemahkan ilmu nahu, hanyalah maksudku hendak mengatur peraturan tertib surat-suratan dan perkataan bahasa Melayu juga; ibid.: 24.29-31). Therefore, he returns to the rules for correctly spelling the vowels in Malay, which are dealt with in the first nine sections of his treatise after the introduction (ibid.: 12-25). The main problem is that in Jawi script vowels are sometimes marked with and sometimes without a glide consonant. A word like kepada ‘for’ is spelt correctly <kpd>, but di dalam ‘within’ is spelt <ddʾlam>. According to Raja Ali Haji, people without an education are confused by this orthography and do not grasp the underlying rules (ibid.: 21). The main rule, he says, is that in words that are free from ambiguity,

4 This definition is almost identical with the one in the ʾĀjurrūmiyya (Carter 1981: 34), which has taqīyir ʾahwāl ʿawāhir al-kalām.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>§19</td>
<td>mafʿūl muṭlaq</td>
<td>p. 43f.</td>
</tr>
<tr>
<td>§20</td>
<td>mafʿūl lahu</td>
<td>p. 44</td>
</tr>
<tr>
<td>§21</td>
<td>mafʿūl fihi</td>
<td>p. 44</td>
</tr>
<tr>
<td>§22</td>
<td>mafʿūl maʾahu</td>
<td>p. 44</td>
</tr>
<tr>
<td>§23</td>
<td>ḥāl</td>
<td>p. 44f.</td>
</tr>
<tr>
<td>§24</td>
<td>tamyīz</td>
<td>p. 45</td>
</tr>
<tr>
<td>§25</td>
<td>taʾkid</td>
<td>p. 45</td>
</tr>
<tr>
<td>§26</td>
<td>badal</td>
<td>p. 45f.</td>
</tr>
<tr>
<td>§27</td>
<td>šifa</td>
<td>p. 46</td>
</tr>
<tr>
<td>§28</td>
<td>ʾidāfa</td>
<td>p. 47</td>
</tr>
<tr>
<td>§29-30</td>
<td>epistolography</td>
<td>p. 47-51</td>
</tr>
<tr>
<td></td>
<td>epilogue</td>
<td>p. 51-53</td>
</tr>
</tbody>
</table>
there is no need for any additional letters to indicate the pronunciation: <drpd> is immediately recognized as daripada ‘from’, so there is no need for clarification. But in words like kota ‘city’, kata ‘word’, kita ‘we [incl.]’, writing the glide <k>, <kwt>, and <kyt> is necessary in order to distinguish them from each other (ibid.: 23).²

It is not clear whether Raja Ali Haji regards Arabic as a superior language. He seems to believe firmly in the richness of his own language, at least for those who know how to speak it well. Occasionally, however, he refers to a qualitative difference in the linguistic inventory, for instance with respect to the negations. After explaining that the Malay negation tiada is used for both generic and specific negations, he remarks:

Malay is extremely poor compared to Arabic. In Arabic there are two, three, four different negative particles, each with its own rules (Sangatlah miskinnya bahasa Melayu ini jika dibangsakan dengan bahasa ‘Arab. Adapun pada bahasa ‘Arab harf tiada itu ada dua, tiga, empat jenis, masing-masing dengan hukumnya; ibid.: 35.23-26).

Differences between the two languages are noted throughout the Bustān, for instance in the verbal system. Raja Ali Haji explains that there are three classes of verbs, fiʿl māḍī (perfect), fiʿl muḍārī (imperfect), and fiʿl ʿamr (imperative). In Arabic, these three classes have different forms, but in Malay, the same meanings can be expressed by adding a particle (ibid.: 30.2f.): “As the sign of the fiʿl māḍī telah may be used, e.g. telah memukul ‘he hit’” (Adapun tanda fiʿl māḍī itu, boleh ditanggungkan telah seperti telah memukul). The term tanda is used here as an equivalent of ‘alāma ‘sign, marker’, which occurs elsewhere. Likewise, the fiʿl muḍārī is expressed by adding lagi akan, e.g. lagi akan memukul ‘he’ll hit’ (ibid.: 31.7).

Another example occurs in the discussion about those particles that govern the genitive (jarr), where it is stated explicitly that the two languages express the same meaning with different syntactic means:

When this particle occurs in Arabic, it needs its majrūr [i.e. genitive] which is governed by it.⁶ In Malay there is no mention of its majrūr, only of its intention and its meaning (Jika pada bahasa ‘Arab datang harf itu berkehendaklah majrūrnya yang dihabarkan. Jika pada bahasa Melayu tiadalah dibicarakan majrūrnya itu melainkan kehendaknya dan maknanya jua, adanya; ibid.: 32.12-15).⁷

---

² The discussion is complicated by Raja Ali Haji’s use of technical terms, for instance when he states that it is sometimes obligatory “to use ʾi irrāb with a letter” (wajib diiʾrābkan dengan huruf; Bustān: 23.21) in cases like kumbang ‘bumblebee’ <kwmbŋ> and kambing ‘goat’ <kmbyŋ>. He seems to suggest that there are two ways of indicating vowels (mengiʾrābkan), with haris and with huruf, but the text is not very clear.

⁶ It is not clear to me why dikhabarkan‘y to be predicated by it’ is used here; presumably, he means that the particle is the operator of the majrūr.

⁷ As an example, the author refers to the famous issue of ʿakaltu l-samaka ḥattā raʾs-i/a/u-hu ‘I ate the fish, up to/including/even its head’, but without quoting the Arabic sentence (Bustān: 33; van Ronkel 1901: 551). Here, too, the problem of ambiguity exists in both languages, but they have different tools to solve them: while Arabic has declensional endings, Malay uses particles (Versteegh 2019).
In the next section, we shall follow the order of the chapters of the Bustān to see how Raja Ali Haji extends the rules of Arabic grammar to Malay.

3. ARABIC GRAMMAR EXTENDED TO MALAY

Raja Ali Haji’s indebtedness to the Arabic grammatical tradition is evident right away from his classification of the parts of speech, in which he follows the Arabic tripartition:

Speech that is produced has to be one of three things. The first is in Arabic ism, which means ‘name’, the second is fi‘l, which means ‘action’, and the third harf. By this we mean here the harf that has a meaning" (Bermula yang diperbuat perkataannya itu tiadalalah sunyi ia daripada tiga perkara. Pertama, pada bahasa ‘Arab ism yakni nama, dan kedua, fi‘l yakni perbuatan, dan ketiga harf. Maka dikehendak harf di sini harf yang ada baginya makna; ibid.: 25f.).

After having introduced the three parts of speech with their Arabic names and the Malay equivalent for the noun (nama) and the verb (perbuatan), Raja Ali Haji deals with each part in a separate section under its Arabic name (ism, ibid.: 26-29; fi‘l, ibid.: 29-32; harf, ibid.: 32-38). The order of treatment, first the noun, then the verb, follows that of the ‘Alfiyya, as the ‘Ājurramiyya begins with the verb.

Within the sections about the noun and the verb, the pattern of the ‘Ājurramiyya is followed more or less closely. For the nouns, the division into indefinite (nakira) and definite (ma‘rifā) nouns follows right after the classification of the parts of speech (ibid.: 26-29). The definite nouns are divided into five categories, damūr (personal pronoun), ‘alam (proper noun), ‘išāra (demonstrative), mawsūl (relative), ‘idafa (annexion), omitting the noun with the definite article, but splitting the category of mubham (a generic term for demonstrative and relative pronouns) of the ‘Ājurramiyya into demonstratives and relatives.

For the analysis of the verbal forms in Malay, we need to go to the Kitab pengetahuan bahasa. According to van der Putten (2002: 426), two verbal forms are analyzed by Raja Ali Haji as nominals: the meN-form, which focuses on the agent, is called ism fā’il and translated as orang yang... ‘someone who...’, whereas the di- form, which focuses on the patient, is called ism maf‘ūl and translated as orang yang kena... ‘someone who is affected by...’. A typical example is found in the lemma ubat (obat) ‘medicine’ (Pengetahuan: 87f.), with its derived forms (Table 2). The term ism fi‘l is used to indicate the root word ubat, which can only

---

8 This expression refers to the description right at the beginning of Sibawayhi’s Kitāb where the term harf is specified as harf fā’ā li-ma‘nā ‘particle that brings a meaning’, presumably in order to distinguish it from harf in the sense of ‘letter’.
9 In the ‘Ājurramiyya, this is dealt with in the chapter on the na‘īr.
10 The later commentaries (Kafrāwī, Šarḥ: 209; Širbīnī, Nūr: 266) take Ibn ‘Ājurrum to task for having omitted the relatives.
be used nominally.\textsuperscript{11} While the explanation of the derived forms does not impose a nominal interpretation per se, the terminology used in classifying them does not seem to leave room for any alternative interpretation.

**Table 2**
The Lemma *ubat* in the *Pengetahuan*: 87f.

<table>
<thead>
<tr>
<th>The Lemma</th>
<th>Explanation</th>
<th>Arabic Equivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ubat</em> ‘medicine’</td>
<td>[explanation of what a medicine is]</td>
<td>ism fiʿil</td>
</tr>
<tr>
<td><em>mengubat</em> ‘to heal’</td>
<td>“i.e., someone who heals someone” (<em>iaitu orang yang mengubatkan orang</em>)</td>
<td>ism fāʿil</td>
</tr>
<tr>
<td><em>diubatkan</em> ‘to be treated’</td>
<td>“i.e., someone who has undergone treatment” (<em>iaitu orang yang kena ubat</em>)</td>
<td>ism mafʿūl</td>
</tr>
<tr>
<td><em>berubat</em> ‘to take a medicine’</td>
<td>“i.e., someone who has fallen ill takes a medicine” (<em>iaitu orang yang kena penyakit berubat ia</em>)</td>
<td>ḥāl</td>
</tr>
<tr>
<td><em>ubatkanlah</em> ‘heal!’</td>
<td>“i.e., someone orders someone to treat a patient” (<em>iaitu seseorang menyuruh kepada seseorang mengubatkan penyakit</em>)</td>
<td>fiʿl ‘amr</td>
</tr>
<tr>
<td><em>perubatan</em> ‘taking medicine’</td>
<td>“i.e., people are discussing the act of taking medicine” (<em>iaitu orang yang berkhabar- khabar akan pekerjaan berubat-ubatan</em>)</td>
<td>taʿwīl maṣdar</td>
</tr>
<tr>
<td><em>ubatan</em> ‘medicine’</td>
<td>“i.e., some components” (<em>iaitu beberapa juzuk</em>)</td>
<td>ism maṣdar</td>
</tr>
</tbody>
</table>

Within the category of particles (*ḥurūf*), Raja Ali Haji distinguishes between the *ḥurūf jarr* (prepositions) and those particles that serve to add a meaning to the sentence. The criterion by which the two categories are distinguished is the category of the Arabic equivalent of the Malay particle. To the former category belong those particles that correspond to Arabic particles governing the genitive, such as *dengan* (Arabic *bi-* ‘with’, ‘by [in oaths]’, etc.), *daripada* (Arabic *min* ‘from’), *kepada* (Arabic *li-* ‘for’), *seperti* (Arabic *ka-* ‘as’), etc. To the latter category belong particles used to address someone, like *hai*, *wai*, *weh*, corresponding to the *ḥurūf al-munādāt* (vocative particles); *melainkan*, introducing exceptions; and *tetapi* ‘but’, a *ḥarf al-istūdrāk* (Bustān: 34).\textsuperscript{12}

There is yet a third category of particles, discussed by Raja Ali Haji under the label of *ḥarf Melayu* ‘Malay particles’ (*ibid.*: 37.5), apparently because they have no direct Arabic equivalent. He shows himself to be an acute observer of the fine pragmatic nuances that some of these particles convey, for instance when he states

\textsuperscript{11} Elsewhere Raja Ali Haji uses *ism fiʿil* in the sense of ‘interjection’ (*Bustān*: 28).
\textsuperscript{12} In the case of *melainkan* ‘except’, a distinction is made between the positive and the negative exception. In Arabic, this distinction has a function since it determines the case ending in the excepted noun after *ʿillā*, whereas in Malay it does not.
that the particle *amboi* ‘gosh!’ may indicate both amazement and disappointment, depending on the intonation (*ibid.*: 37.30-34).

The next chapter (*ibid.*: 38f.) introduces the distinction between *perkataan* ‘speech’ and *kata kata* ‘words’, corresponding to that between *kalām* ‘speech’ and *kalima* ‘word’ in Arabic linguistic theory. In the Ḍjurūmiyya tradition, the four distinctive features of *kalām* are the following (*Carter* 1981: 8-10): it is a formal utterance (*lafẓ*), composite (*murakkab*), informative (*muʃūd*), and conventional (*bi-l-waḍˁ*). This excludes, among other things, speech by someone sleeping, and self-evident statements of the type *al-samaˀu fawqanā* ‘the sky is above us’ and *al-ˀardu taḥtanā* ‘the earth is beneath us’.

Two of these criteria are repeated almost verbatim by Raja Ali Haji, who defines *perkataan* as “an utterance that conveys useful information” (*lafaz yang memberi faedah*; *ibid.*: 38.26f.). He then gives the Malay translation of the two Arabic sentences quoted above, *langit di atas kita* and *bumi di bawah kita*, and declares that these fall outside the definition of speech (*perkataan*), because they do not convey new information.\(^\text{13}\)

In his discussion of syntax, Raja Ali Haji parses Malay sentences in the same way as in Arabic grammar. A sentence may begin with a noun serving as *mubtadaʾ* (topic), on which a ʿḫabar (predicate) is ‘made to lean’, in other words, ‘which supports a ʿḫabar’ (*ibid.*: 39). The terms for the sentence constituents are introduced with their Arabic form, but also with a Malay translation, *permulaan* (mubtadaʾ), *berita* (ḫabar); the verb *bersandar* ‘to lean on’/disandarkan ‘to be made to lean on’ apparently reflects Arabic musnad/ˀisnād, which is not used in this text.\(^\text{14}\) As an example of a nominal sentence he cites *zaydun qāʾimun* ‘Zayd is standing’, which is translated as *si zaid yang berdiri*, where the relative sentence *yang berdiri* ‘who is standing’ translates Arabic qāʾimun.\(^\text{15}\)

The rules (*hukuman*) for the nominal sentence are taken directly from Arabic syntax, for instance, the restriction that the *mubtadaʾ* cannot be an indefinite noun. Their dependence on the rules of Arabic grammar makes some of these restrictions difficult to understand when applied to Malay. One set of rules concerns the placement of the predicate, which in principle follows its topic. In some cases, this position is compulsory. Al-Šīrbindī (*Nūr*: 202-204) mentions three of them: i) when there might be confusion because both topic and predicate are definite, as in *zaydun ʿaḥūka* ‘Zaydisyourbrother’, unless the context makes clear which is the predicate, as in *ʿabū yūsufa* ʿabū ḥanīfata ʿAbū Yūsuf is ‘Abū Ḥanīfa’; ii) when the topic might become confused with the agent, so that the nominal sentence turns into a verbal

---

\(^{13}\) According to al-Kāfrāwī (*Šarḥ*: 13) ‘conventional’ (*bi-l-waḍˁ*) excludes foreign languages from the definition of ‘speech’, so this is not the most appropriate criterion to be used in a grammar of Malay. The shorter form of the definition in the *Bustān* corresponds to the one in Ibn Hišām (*Qatr*: 54) “speech is a meaningful utterance” (*al-kalām lafẓ muʃūd*).

\(^{14}\) Note that *bersandar* and disandarkan are also used as translation for ʿidāfa (*Bustān*: 47.2).

\(^{15}\) For the interpretation of *berdiri* as a relative sentence, see above.
sentence, as in qāma zaydun ‘Zayd stood up’; and iii) when the predicate is accompanied by ʾillā or ʾinnamā.

All three cases are mentioned by Raja Ali Haji, even though they are hardly valid for Malay. For the first case, he mentions the sentence si zaid itu saudaramu ‘Zayd is your friend’ as an example of a sentence in which the predicate needs to be placed last. This restriction does not apply, he says, when the context indicates the predicate, as in ʾabū yūsufa ʾabū ḥānīfata ‘Abū Yūsuf is ʾAbū Ḥanīfa’ (Bustān: 41.8f.).

The second restriction deals with the relationship between nominal and verbal sentences:

Don’t say telah berdiri si zaid ‘Zayd stood up’! If something like that is said, i.e. telah berdiri si zaid, it does not fall under the heading of topic/predicate, but under the heading of verb/agent (Maka jangan dikata telah berdiri si zaid, dan jika dikata juga seperti itu ya’ni telah berdiri si zaid, maka bukanlah pada bab mubtadaʾ ʾhabar ini, tetapi adalah ia pada bab fiʾl dengan fāʾil; ibid.: 20-23).

As for the third restriction, he simply states that in a sentence with hanya sungguhnya ‘only’ or melainkan ‘except’ the predicate always follows (ibid.: 41.31f.). The rather unnatural sounding example he cites for this restriction, tiada ada si zaid itu melainkan yang berdiri ia, probably means something like ‘It is not the case that Zayd is [anything else] than that he is standing’.

The author also deals with the elision of the topic or the predicate (cf. Carter 1981: 204). One of the cases he mentions is that of the predicate after the Malay equivalent of lawlā ‘if it weren’t for’. The example sentence in Malay, jikalau tiada si zaid binasalah si ʿummar ‘If it weren’t for Zayd, ʿUmar would be lost’ (Bustān: 42.7), looks constructed. This particular case derives from the commentaries on the Ṣafyya (v. 138), for instance by Ibn ʿAqīl (Ṣahr 1: 248; see also Zamahšari, Muṭṣṣal: 14.6f.; Ibn Hišām, Qaṭr: 121).

Some of the rules are incomprehensible outside the context of Arabic grammar, for instance the rule that with a predicate consisting of a single word no pronoun is expressed. This is explained in the following passage, one of the few instances of an Arabic grammarian being mentioned by name:

In a ḥabar mufrad, such as si zaid saudaramu ‘Zayd is your friend’, the rule of the ḥabar is not to receive a pronoun, i.e. ia ‘he’. But according to al-Kisāʾī [d. 189/805] and al-Rummānī [d. 384/994], and a group of others, a pronoun must be understood here, as when you say si zaid saudaramu yaʾni ia ‘Zayd is your friend’, i.e. ‘he’ (adapun misal ḥabar yang mufrad itu ‘si zaid saudaramu’ adalah hukum ḥabar ini tiadalah diberi damīr yakni ‘ia’; akan tetapi pada Syeikh Kisaʾi dan Syeikh Rumani dan beberapa jamāʾah, harus ditanggungkan damīr, seperti katamu ‘si zaid saudaramu’ yaʾni ‘ia’; Bustān: 40).

The issue concerning the presence of a pronoun in a nominal sentence, which is irrelevant for Malay, derives from a well-known controversy in Arabic grammar (see Ibn al-ʾAnbārī, ʾInṣāf: 30f.) about the question of whether a single noun as
Amar ḥabar in a sentence like ‘amr ḡulāmuka ḍāʾ Amr is your slave’ contains a pronoun referring to the mubtada’. According to the Kufans (i.e. al-Kisāʾī) and al-Rumnānī, this predicate contains a pronoun, since it is equivalent to a participle ḥādimuka ‘your servant [lit. your serving one]’, which according to all grammarians contains a pronoun. This particular statement probably derives from the commentaries on the Ṣafīyya: Ibn ʿAqīl (Ṣarḥ I: 205) mentions the opinion of al-Kisāʾī wa-l-Rumnānī wa-jamāʾa, which is similar to the passage in the Bustān.

After the topic/predicate, our text discusses the agent (fīʿil) and the verb (fiʿl), which is the order of treatment in the Ṣafīyya, though not in the ʿĀjurrāmīyya.16 Nothing much is said about this sentence type, but mention is made of the fact that the agent may be overt (ẓāhir) or hidden (tersembunyi); the latter term translates Arabic muḍmār. This is followed by brief sections on the different types of object (maʃūl): maʃūl bihi, maʃūl muṭlaq, maʃūl lahu, maʃūl fihi, and maʃūl maʿahu, followed by other accusative constructions, ḥāl, tamyīz. In general, the order of treatment of these sections follows that of the ʿĀjurrāmīyya in putting the direct object first, followed by the maʃūl muṭlaq, whereas the Ṣafīyya tradition starts with the verbal noun. Note that our author uses maʃūl muṭlaq instead of maṣdar; the latter is the term used by Ibn ʿĀjurrūm, for which he is criticized by al-Šīrbīnī (Carter 1981: 344; cf. Kasher 2019: 206). Raja Ali Haji also uses maʃūl fihi where the ʿĀjurrāmīyya has zarf; this is mentioned as an alternative term by the commentators (ʿAzharī, ʿṢarḥ: 100.6; ʿKafrawī, ʿṢarḥ: 261.5; ʿŠīrbīnī, Nūr: 352).

In the ʿĀjurrāmīyya tradition, the next sections deal with the tawbīʿi, satellite constituents that derive their case ending from the word they follow, i.e. attribute (naʿi), corroboration (taʾkīd), coordination (ʿaff), and apposition (badal). The author of the Bustān leaves out coordination and deals with the adjective (ṣifa) at the end. Concerning taʾkīd, he repeats almost literally the explanation in Daḥlān’s commentary (Ṣarḥ: 428-429; absent in al-Šīrbīnī): corroboration is needed to discard any ambiguity. When you say ‘Zayd came’, someone might assume that a letter or a messenger from him has arrived, but this possibility vanishes when you add nafsuḥu/dirinya ‘himself’ (Bustān: 45).

The section on badal (ibid.: 45f.) follows the ʿĀjurrāmīyya and its commentaries (Šīrbīnī, Nūr: 312ff., 318-320) closely. Four kinds of substitution are mentioned:

i. badal al-kull min al-kull: this is not the term used in the ʿĀjurrāmīyya, which has badal al-šayʾ min al-šayʾ:17 lalu aku dengan saudara engkau si zaid ‘I passed your friend, Zayd’, Arabic jāʾa zaydun ʿaḥūka ‘Zayd came, your brother’.

16 In his commentary on the ʿĀjurrāmīyya, al-Šīrbīnī (Nūr: 150ff.) mentions the alternative sequence of starting with the topic, which he attributes to Ibn Mālik and Ibn Hišām.
17 Al-Šīrbīnī mentions the former, but states that it is less appropriate; it is mentioned as an alternative name by Daḥlān (Ṣarḥ: 437.1) and al-Kafrawī (Ṣarḥ: 232.15). Ibn Hišām (Qaṭr: 288) has badal kull min kull.
The third type is mentioned also by some of the commentators on the genitive construction in Ibn al-Širbīnī’s commentary. This mirrors a remark in al-Širbīnī’s commentary about reading a verse as an example of an adjective used for the purpose of praise, to blame, to express commiseration, and to convey a sixth function, that of clarification, which is mentioned indeed in the commentaries on Ibn Mālik’s Alfiyya (v. 386), for instance by Ibn ʿAqīl (Ṣahih II: 43), and by Ibn Hišām (Ṣaḥīḥ: 237). The Alfiyya tradition discusses the ʿidāfa construction before the satellite constructions, rather than after them.

18 Presumably, Raja Ali Haji has taken over from an Arabic commentary a remark about ṭāḥīṣī, leaving out commiseration, replacing it with clariﬁcation. This mirrors a remark in al-Širbīnī’s commentary about ṭāḥīṣī, replacing it with clarification. Ibn al-Širbīnī discusses the third type of genitive construction, defined as the case of the adjective having the function of ṭāḥīṣī, not found in the commentaries on Ibn Mālik’s Alfiyya (v. 386), for instance by Ibn ʿAqīl (Ṣahih II: 43), and by Ibn Hišām (Ṣaḥīḥ: 237). The Alfiyya tradition discusses the ʿidāfa construction before the satellite constructions, rather than after them.

19 These categories of adjectives are mentioned by al-Zamaḵšarī (Muваṣṣal: 46), who also quotes the Qur’ānic verse as an example of an adjective used for the purpose of ʿuḍāḥ. The same categories are distinguished by Ibn Hišām (Ṣaḥīḥ: 267), who cites the same example for the function of ʿuḍāḥ and adds a sixth function, that of clarification (tawdīḥ). Ibn al-Ḥājjīb’s Kāfiya (p. 108).
is used for purposes of ta'rīf, the latter for purposes of tahṣīṣ. The two constructions are explained in the Malay text as faedah kepada pengenalan ‘[conveying] the meaning of making known’ for the former, and faedah kepada menentukan ‘[conveying] the meaning of determining’ for the latter. These two phrases are offered as a translation of the Arabic terms ma'rīfa and nakira, but it is doubtful they could be understood by an average Malay reader without knowledge of Arabic grammar.

4. CONCLUSION

In compiling his grammar of Malay, Raja Ali Haji followed the model of the grammatical literature in which he was raised. It is hard to ascertain exactly which treatises were his main authority, but it seems that he mostly followed the ʿĀjurrimiyya tradition, although he borrowed rules and remarks from other commentaries as well. In a fair number of cases, the grammatical rules and examples cited in the Bustān are irrelevant for the structure of Malay. His desire to find Malay analogues for Arabic constructions is evident, even to the point where he constructs Malay sentences rather than deriving them from living speech. The Malay practice of learning Arabic texts by heart together with a Malay translation may have increased the acceptability of “translationese”, and may even have led to the introduction of new constructions into the language in much the same way as European languages introduced constructions taken over by Latin, often on the basis of Bible translations from Latin or Hebrew.

With respect to lexicography, Raja Ali Haji himself noted the difference between his own approach and that of the Western scholars with whom he was in contact, as van der Putten (1995) reports. Presumably, he felt that his own discursive approach was more in line with the Arabic dictionaries, which he believed to be better suited to the Islamic language that was Malay.

Since Raja Ali Haji was a pioneer in writing about Malay grammar in Malay, the creation of a technical lexicon must in large part be attributed to him, although some Malay terminology in Qurʾānic exegesis and Arabic grammar was available. He provides Malay equivalents for a considerable number of Arabic grammatical terms in the form of loan translations, and in some cases even replaces an Arabic term with a Malay one, for instance in the ʿIsnād terminology. As the text is written in Jawi, it is not always possible to see whether an Arabic term is used in the original form or as an integrated loanword. The written word lafẓ, for instance, was probably pronounced lafaz, and fiʿl was pronounced fiʿil. The integration of Arabic loanwords is evident in a number of derivations, such as disifatkan ‘to be described’ (Bustān: 42.25) or dikhabarkan ‘to be predicated’ (ibid.: 32.13).20

20 The number of derivates is relatively low compared to the general lexicon and the specialized lexicon of Qurʾānic exegesis (Riddell 1990: 245-250), possibly because Raja Ali Haji seems to prefer loan translations over loanwords.
Raja Ali Haji’s work remained a relatively isolated attempt at describing Malay within an Arabic framework. Even though he is still revered as a hero of the Malay language,21 modern grammars follow a different framework, that of Western linguistics, as in Moeliono and Dardjowidjojo’s (1988) standard grammar, which has traces of Arabic grammatical terminology, but almost always opts for the Western terminology.

BIBLIOGRAPHIE

Primary sources


---

21 A memorial sign in Tanjung Pinang on the island of Riau calls him “The father of Malay/Indonesian, an intellectual of the beginning of the twentieth century” (Bapak bahasa Mel-ayu-Indonesia budayawan di gerbang abad XX); see e.g. http://liputankepri.com/catatan-sejarah-pulau-penyengat-tanjungpinang (downloaded 30.07.2018).
Secondary sources


