

10-31-2024

## Seventeenth-century Malay wordlists and their potential for etymological scholarship

Tom G. Hoogervorst

Royal Netherlands Institute of Southeast Asian and Caribbean Studies (KITLV), Leiden, The Netherlands  
and Universitas Negeri Malang, Indonesia, [hoogervorst@kitlv.nl](mailto:hoogervorst@kitlv.nl)

Follow this and additional works at: <https://scholarhub.ui.ac.id/wacana>



Part of the [Other Arts and Humanities Commons](#), and the [Other Languages, Societies, and Cultures Commons](#)

---

### Recommended Citation

Hoogervorst, Tom G. (2024) "Seventeenth-century Malay wordlists and their potential for etymological scholarship," *Wacana, Journal of the Humanities of Indonesia*: Vol. 25: No. 3, Article 7.

DOI: [10.17510/wacana.v25i3.1782](https://doi.org/10.17510/wacana.v25i3.1782)

Available at: <https://scholarhub.ui.ac.id/wacana/vol25/iss3/7>

This Article is brought to you for free and open access by the Faculty of Humanities at UI Scholars Hub. It has been accepted for inclusion in *Wacana, Journal of the Humanities of Indonesia* by an authorized editor of UI Scholars Hub.

# Seventeenth-century Malay wordlists and their potential for etymological scholarship

TOM G. HOOGERVORST

## ABSTRACT

Early-modern wordlists and dictionaries provide an underexplored area for etymological scholarship. By critically comparing different sources written under the aegis of the Dutch East India Company, often compiled by autodidacts who were unable to gain fluency, this article makes some generalizations about the etymology and contact history of early-seventeenth-century Malay. I demonstrate that the Dutch materials provide concrete instances to study lexical change, both phonologically and semantically. When used advisedly, the material also casts light on the nature of language contact in an era in which the Indo-Malayan Archipelago was at the nexus of trade networks connecting speakers of Arabic, Persian, Chinese, Portuguese, and various Indian and Indonesian languages. Finally, early-modern lexicography offers valuable data to reconstruct elements of the society being studied, including in the realms of religion, social hierarchies, and material culture.

## KEYWORDS

Malay, seventeenth century, VOC, lexicography, dictionaries, loanwords, archaisms, regional varieties.

---

TOM G. HOOGERVORST is a historical linguist employed as a Senior Researcher at the Royal Netherlands Institute of Southeast Asian and Caribbean Studies (KITLV) in Leiden and as an Adjunct Professor at Universitas Negeri Malang. He is interested in language contact in Southeast Asia and the Indian Ocean, including in the historical information which can be deduced from lexical borrowing in the realms of culture, literature, and cuisine. Tom Hoogervorst can be reached at: [hoogervorst@kitlv.nl](mailto:hoogervorst@kitlv.nl).

© 2024 Faculty of Humanities, Universitas Indonesia

TOM HOOGERVORST | DOI: 10.17510/wacana.v25i3.1782.

## 1. INTRODUCTION

Various Malay wordlists and dictionaries appeared during the eastward expansion of the Dutch East India Company (VOC), which was founded in 1602 and went bankrupt in 1799.<sup>1</sup> This article examines the historical development of Malay using these rich but arcane sources. I focus on loanwords and regionalisms, foregrounding words no longer in use today, to reconstruct the broader linguistic landscape and contact situation of seventeenth-century Malay. This temporal variety (chronolect) obviously consisted of numerous regional dialects, courtly registers, distinct spoken and written forms, and conventionalized foreigner talk. As the sociolinguistic history of early-modern Malay has already attracted prodigious scholarship (James T. Collins 1998: 22-54; Henk M.J. Maier and Jan van der Putten 2002; Waruno Mahdi 2007), I limit myself here to its lexicon and the associated etymological insights in keeping with the scope of this issue.

The earliest reliable Dutch-authored lexical source on Malay, the work of naval officer Jacob Cornelisz van Neck, appeared in 1598, a few years before the establishment of the VOC. His *Vocabulaer vande Javaensche ende Malaysche woorden* (Vocabulary of Javanese and Malay words; henceforth “vN”) is a Dutch-Malay-Javanese wordlist compiled on Ternate, an important political and economic hub in North Maluku. It has been analysed by James T. Collins and Hans Schmidt (1992), while Mahdi (2007: 120-122, 323-334) has paid attention to its early German reworkings. A much larger wordlist was compiled by the navigator Frederik de Houtman during his captivity in Aceh, North Sumatra. His 1603 *Spraeck ende woord-boeck* (Phrasebook and dictionary; henceforth “Ho”) includes twelve dialogues in Dutch and Malay, three dialogues in Dutch and Malagasy, a Dutch-Malay-Malagasy wordlist (see also Alexander Adelaar, in this issue), a Dutch-Arabic-Turkish wordlist, and some astronomical tables. A lexical analysis of the Malay section in this material has been published by Denys Lombard (1970: 168-263), and Nicoline van der Sijs (2000) provides a useful history of the publication itself. In addition, De Houtman’s Malay dialogues have been used to reconstruct elements of North Sumatran society (A.J.S. Reid 1993: 92, 144, 237-238).

A relatively large Malay-Dutch and Dutch-Malay dictionary was published in 1623 by the clergymen Caspar Wiltens and Sebastiaan Danckaerts. Compiled in Ambon, the Malay recorded in their *Vocabularium ofte Woort-boeck* (Vocabulary or Dictionary; henceforth “WD”) displays significantly different features from De Houtman’s data. The dictionary contains separate lists of Portuguese and Ternatan words reportedly in common use in the local Malay variety. Although some of the examples given have disappeared from Ambon Malay, others have persisted. A largely expanded version – now reorganized into a Dutch-Malay dictionary – was published in 1650 by Justus

<sup>1</sup> This article has benefited from the useful feedback of Alexander Adelaar and Waruno Mahdi. For overviews of Malay lexicographical material from the VOC period, see Caspar Adam Laurens van Troostenburg de Bruyn (1884: 425-431), W. Linehan (1949), and John Landwehr (1991).

Heurnius (henceforth “He”), a missionary, linguist, doctor, and botanist based in Batavia.<sup>2</sup> In 1677, the French clergyman Frederik Gueynier compiled yet another enlarged and revised edition (henceforth “G”). Published in Batavia, where he was stationed, this appears to be the first book in Malay printed in the Malay Archipelago (Annabel Teh Gallop 1990: 87-92).

While these dictionaries form the focus of the present study, seventeenth-century Dutch-Malay lexicography did not exist in a vacuum. The institutionalized study of the Malay lexicon certainly predates the European presence in the region, as seen from a fifteenth-century Chinese-Malay wordlist compiled in Malacca by Yang Lin 楊林, an interpreter working for the imperial *Siyiguan* 四譯館 (College of All Foreigners). This valuable source was extensively studied by Evangeline Dora Edwards and Charles Otto Blagden (1931). Four Malay-Cham dictionaries, undated but of an archaic linguistic character, have likewise been preserved and studied (Pho Dharma 1999). The Italian scholar-explorer Antonio Pigafetta compiled the earliest European wordlist of Malay, which contains some Bruneian words clearly marked as such, along with some presumably misplaced entries from Philippine languages (see Alessandro Bausani 1960). Few lexicographic studies have been found in the Portuguese archives, but a Portuguese-Malay wordlist dated around 1700 (see Denys Lombard and Luis Filipe F.R. Thomaz 1981) offers interesting points of comparison with Dutch-authored studies. The first substantial British-authored contribution is a Malay-English and English-Malay dictionary, also containing several dialogues, published in 1701 by the mariner Thomas Bowrey. The extent to which this dictionary was based on the work of others is not entirely clear.<sup>3</sup> A liberal number of his dictionary entries – but not apparently his dialogues – seem to have been lifted from Dutch sources.

In addition to dictionaries, important lexical insights can be gained from scholarship on a variety of early primary sources. In this regard, we can mention the Tanjung Tanah manuscript, a fourteenth-century legal code from the Kerinci region in Sumatra, whose vocabulary has been examined by a research team under the supervision of Uli Kozok (2015: 110-129). In a mid-sixteenth-century context, the writings about the Malay used in Maluku by the Portuguese soldier and administrator António Galvão are informative (Hubert Jacobs 1971: 367-376). Along similar lines, the publications of Dutch missionaries and Bible translators have been used to study Malay language development in the seventeenth century (Collins 1992, 2004; Lourens de Vries 2018). The 1692 *Orientalisch-Indianische Kunst- und Lustgärtner* (Oriental-Indian Art and Pleasure Gardener) by George Meister has likewise been subjected to lexicographic and historical linguistic studies (Collins 1995; Mahdi 2007). We should also call attention to a recently described compilation of poems, the

<sup>2</sup> See Koos Kuiper (2005: 109-111) for a short biography. As the author discusses, Heurnius was also the first Dutch person to study and publish about Chinese linguistics.

<sup>3</sup> See Anna Winterbottom (2016: 54-81) on the history of Bowrey’s dictionary and Collins (1998: 50-53) on its value to the study of Malay language history.

*Livro de Pantuns*, written in Batavia around 1700. Presumably originating from the Mardijker milieu, a Batavian community descended from freed slaves, this manuscript has cast new light on the language history of both Malay and Creole Portuguese (Ivo Castro et al. 2022).

Lexicographic studies based on written rather than spoken Malay are also beyond the scope of this article. In this category, a 1604 Dutch-Malay wordlist compiled by Pieter Willemsz van Elbinck should be mentioned because it contains entries written in the Arabic-derived Jawi script (Philippus Samuel van Ronkel 1896). The sizeable, hitherto unpublished dictionary of Melchior Leydekker (1642-1721) likewise falls beyond its purview, as it is largely based on written texts.<sup>4</sup> The 1780 *Nieuwe Woordenschat uyt het Nederduitsch in het gemeene Maleidsch en Portugeesch* (New Vocabulary from Dutch into common Malay and Portuguese), a Dutch-Malay-Creole Portuguese dictionary, is mentioned only briefly, as it also deserves a separate study.

The use of VOC materials presents considerable challenges. Given the archaic character of both the Malay and the Dutch, innumerable glosses are difficult for present-day speakers of these languages to understand. Even though the original spelling affords some phonological inferences, orthographic inconsistency was the norm rather than the exception. In both Dutch and Malay, <oe>, <u>, and <ou> were used interchangeably for /u/, whereas <i>, <j>, <ij>, and <y> could denote both /i/ and /j/. Another difficulty is the unfamiliarity of early European lexicographers with certain Malay phonemes. They had the greatest trouble with the palatals /tʃ, c/, /dʒ, ʃ/, and /ɲ/, which were transcribed in a bewildering variety of ways. On top of this, it is not always clear whether idiosyncratic transcriptions of the historical schwa and/or word-final nasals reflect regional influence – for example, from eastern Indonesian Malay – or mishearing on the part of the lexicographer. The material also abounds in misspellings introduced by typesetters who clearly lacked any knowledge of Malay. As a result, one often encounters <n> instead of <u> or <c>, <o> instead of <a> or <e>, <f> instead of <t> or <s>, <r> instead of <c> or <t>, and vice-versa. Another point of concern was the habit of lexicographers to lift their entries partly from previous works, thereby also perpetuating earlier mistakes. Jacob van Neck, Frederik de Houtman, and the duo Casper Wiltens and Sebastiaan Danckaerts appear to have worked in isolation from one another, so that their material can be considered original. All subsequent lexicographers have copiously included entries from previous scholarship. As a result, the work of Heurnius, Gueynier, and especially Bowrey constitutes a hodgepodge of regionalisms from across the Indo-Malay Archipelago, without any indication of their provenance or range of usage.

The structure of this article is as follows. Section 2 deals with loanwords in seventeenth-century lexicographic sources, focusing on influence from different parts of the Indo-Malayan Archipelago (2.1), the wider Indian Ocean (2.2), and Portuguese and Dutch (2.3). Unless indicated otherwise, the

<sup>4</sup> See P.H. van der Kemp (1914) on the history of this dictionary.

identifications and etymologies of loanwords are mine. Section 3 examines vocabulary no longer encountered in the same form or meaning in modern Malay, paying attention to phonology (3.1), morphology (3.2), and semantics (3.3). Section 4 highlights the value of seventeenth-century wordlists and dictionaries for reconstructing elements of the Indonesian past, foregrounding religion (4.1), social stratification (4.2), and other areas (4.3).

## 2. LOANWORDS

Although etymology can provide insight into how certain concepts entered a specific society, there is not always a direct correlation between borrowed vocabulary and the objects or ideas they represent. In Malay in particular, we can observe high levels of lexical replacement in which loanwords from Indic, Arabic, or European languages, all denoting similar concepts, are known to have gradually replaced each other over time (Hoogervorst 2024). A closer look at seventeenth-century Malay provides a better understanding of these overlapping layers of borrowing.

### 2.1 REGIONAL

Unsurprisingly, De Houtman recorded a variety of Malay which was heavily influenced by Acehese, both lexically and grammatically. Several examples of Acehisms are highlighted in G.W.J. Drewes (1972), and the list can be expanded (Table 1).

Acehese	Attestations
<i>glang</i> 'worm'	<i>glang</i> 'wormen'
<i>guda</i> 'horse' (Ml. <i>kuda</i> )	<i>gouda</i> 'peert'
<i>jumurang</i> 'opposite' (Ml. <i>səbərang</i> )	<i>jybrant</i> 'over, tegenover'
<i>paluəng</i> 'bald'	<i>palang</i> 'cael'
<i>puurulan</i> 'to swallow'	<i>pirlan, perlan</i> 'slicken, swelghen'
<i>singkeə</i> 'elbow'	<i>sinkoeu</i> 'ellenboghe'
<i>sua</i> 'to incite (fighting cocks)'	<i>suwa</i> 'terghen'

Table 1. Acehese loans in De Houtman not listed in Drewes (1972).

In addition to this Acehese influence, De Houtman also includes entries reflecting local Malay varieties. Table 2 contains some entries which can be identified as North Sumatran Malay.<sup>5</sup> The modern dialects of Langkat and Deli most closely resemble these words, although very little is known about the region's dialectal landscape roughly four centuries earlier.

<sup>5</sup> In the dictionary of Chairani Nasution et al. (2018), North Sumatran Malay encompasses the sub-dialects of Asahan, Bilah Panai, Deli, Langkat, and Serdang.

North Sumatran Malay	Attestations
<i>bubol</i> 'net'	<i>mouboel</i> 'net'
<i>əmpus</i> 'garden' (Langkat, Deli)	<i>empus</i> 'acker, hof'
<i>gədubang</i> 'machete'	<i>gedoubang</i> 'houmes'
<i>gurin</i> 'to lie down' <sup>6</sup>	<i>goering</i> 'legghe, rusten'
<i>kəleh, kəlih</i> 'to see' (Langkat, Deli)	<i>kelih, kelich</i> 'sien'
<i>kəmun</i> 'nephew' (Langkat, Deli)	<i>kemon, comon</i> 'neve'
<i>ləkap</i> 'closed up' (Langkat)	<i>lekap</i> 'toe'
<i>lutar</i> 'to throw' (Ml. <i>lontar</i> )	<i>loutar</i> 'werpen'
<i>mandah</i> 'barn' (Langkat)	<i>mandah</i> 'packhuys [storehouse] <sup>7</sup>
<i>maya</i> 'what' (Langkat, Deli)	<i>maya</i> 'wat, welcke'
<i>pangsan</i> 'to faint' (Ml. <i>pingsan</i> )	<i>pangsawan</i> 'beswymen'
<i>soh</i> 'empty' (Langkat, Deli)	<i>soch</i> 'ledighe'
<i>tain</i> 'earlier'	<i>tain, táin</i> 'hadde [had]'
<i>təmbun</i> 'fat'	<i>tombon</i> 'vet'

Table 2. North Sumatran Malay loans in De Houtman.

In some instances, the exact provenance of De Houtman's words is not entirely clear. North Sumatran Malay *kəmun* 'nephew' listed above presumably goes back to Acehnese *kumuən*, which is itself a Mon-Khmer loan (H.L. Shorto, Paul Sidwell, and Doug Cooper 2006: 328-329). The attestation of *agam* in the meaning of 'male' is recorded in Acehnese and North Sumatran Malay but not beyond this region, leaving the directionality of borrowing elusive. The pronunciation of <*bandersa, bendersah*> for the Arabic loanword *madrasa* 'school' resembles Peninsular Malay *bandarsah*.

Similarly, the usage of <*bobo*> mirrors Peninsular Malay *bubuh* 'to put' (although this form is also attested in colonial-era Indonesian Malay). The subjunctive usage of <*takot*> is reminiscent of the etymologically related Peninsular Malay particle *kot*,<sup>8</sup> both forms go back to Ml. *takut* 'afraid'. It is unclear whether these features also existed historically in North Sumatran Malay. The word <*tandas* 'stil, cackhuys' [lavatory]>, also included by Heurnius, could be either from North Sumatran Malay or from Peninsular Malay. I have not found any North Sumatran Malay counterparts of <*tjymelagy, tjimmelagy* 'tamarinde' [tamarind]>, but compare Minangkabau *cumalagi* and Peninsular Malay *cəlagi* with the same meaning. De Houtman's <*mintuwa* 'behoutvader'> reflects *məntua* 'father-in-law', which is attested in North Sumatran Malay,

<sup>6</sup> But spelled *guring* in Hayati Chalil et al. (1985: 23). Also attested in this meaning in some languages of Borneo (Alexander D. Smith 2017: 546, 614) and as Banjar *guring* 'to sleep' and Salako *ma-gurikng* 'to lie down'.

<sup>7</sup> Here and elsewhere, I have translated the seventeenth-century Dutch glosses into English only if they are considerably different from their counterparts in modern Malay.

<sup>8</sup> See Lombard (1970: 239) on the use of <*takot*> in De Houtman's wordlist and Hoogervorst (2018: 309-310) on the particle *kot* in Klang Valley Malay.

Peninsular Malay, and several other dialects, as well as <*minthua*> (Pigafetta), whereas modern Indonesian Malay has *mærtua* in this meaning. The kinship terms <*sepoejang* ‘susterlings kint’ [nephew, niece]> and <*senini, ceniny* ‘neve, susterling’ [cousin]> (Ho) evidently reflect *sə-poyang* ‘having the same ancestor’ and *sə-nenek* ‘having the same grandparent’, but the precise dialectal origins of these terms are unclear.

More than 4,000 kilometres to the east, we find a significantly different Malay landscape. Most seventeenth-century Ambonese spoke one of the island’s closely related indigenous languages, while Malay – including a localized version known as Ambon Malay – was understood to varying degrees.<sup>9</sup> Loanwords from Portuguese and especially Ternatan were relatively common in this variety, as the two appendices in the *Vocabularium ofte Woort-boeck* by Wiltens and Danckaerts (1623) remind us. The Ternatan language originates from North Maluku but has influenced a number of the Malay varieties of eastern Indonesia. In the Bible translations produced under the VOC, the Ternatan word *jou* ‘lord, king’ was used as an honorific second-person pronoun (Collins 2004: 105-112). Various other Ternatan loanwords are found in the main part (rather than the appendix) of the *Vocabularium ofte Woort-boeck* (Table 3). Note that some of these borrowings are still used in Ambon Malay.

Ternatan	Attestations
<i>baikole</i> ‘a black singing bird with white on its breast’	<i>baicôle</i> ‘een vogheltgien...’ [a little bird ...]
<i>barua</i> ‘wooden chest’ (de Clercq 1890)	<i>barouwa</i> ‘grootte kisten ofte cofferen’
<i>belo</i> ‘wooden pole to moor a boat’	<i>bélo</i> ‘staeck’ [stake]
<i>bia</i> ‘snail’	<i>bia</i> ‘mosselen, slecken’
<i>cafarune</i> ‘dirty, filthy’	<i>saffarouini</i> ‘een die vuyl over sijn lichaem is’ [someone with a filthy body]
<i>dodeso</i> ‘snare’	<i>dodesso</i> ‘strick daermen vogelen mede vanght’ [a snare to catch birds]
<i>dofoma</i> ‘provisions in the form of food’	<i>defomo</i> ‘lijf-tocht’ <sup>10</sup>
<i>farfuji</i> ‘a stake to add pressure to the barrel of a bamboo toy gun’	<i>farafodjou</i> ‘schroeve’ [a screw]
<i>gogele</i> ‘to challenge someone to a fight’	<i>gogelo</i> ‘yemant de guyck naesteecken’
<i>gonofu</i> ‘coconut fibre’	<i>gonoffo</i> ‘bolster van calappas; bast van oude calappes’
<i>kadoto</i> ‘palace’ <sup>11</sup>	<i>cadato</i> ‘paleys’

<sup>9</sup> See, for example, a 1692 letter sent by the clergyman Petrus van der Vorm detailing Ambon’s linguistic landscape (Hendrik E. Niemeijer and Th. van den End 2018a: 14).

<sup>10</sup> Also found as <*de fomo* [booty, plunder]> in Brouwerius’ translation of Genesis (Collins 2004: 95). The author proposes a Portuguese provenance.

<sup>11</sup> Compare the earlier form *kadatu, kadatung* (Fredrik Sigismund Alexander de Clercq 1890), ultimately from Javanese *kaḍaton*, ‘palace’.



<i>kora-kora</i> 'a large boat with double outriggers'	<i>corcorre</i> 'corcorre' <sup>12</sup>
<i>kotale</i> 'to try little by little according to one's capabilities'	<i>cottáli, cotáli</i> 'aenhouden om yet te verkrijgē; eyschen met gheduyrich aenhouden' [to persist in obtaining something; to demand persistently]
<i>loleo</i> 'harbour'	<i>loléo</i> 'ancker plaatse' [anchorage]
<i>mahera</i> 'keel of a boat'	<i>mahera</i> 'prau uit een hout' [a dug-out boat]
<i>makurebe</i> 'to compete'	<i>marorebe</i> 'soecken voor malkanderen te gaen'
<i>mia</i> 'monkey'	<i>mea</i> 'aep'
<i>moi</i> 'all gone, extinct'	<i>mof</i> 'uitgestorven of uit-geroeyt van menschen' [extinct or exterminated (of people)]
<i>nganti-nganti</i> 'earrings'	<i>gnate gnate</i> 'oorhangers'
<i>salawaku</i> 'shield'	<i>saloacco, saloáco</i> 'Molucx schilt' [Moluccan shield]
<i>sedu</i> 'to joke'	<i>sedó</i> 'verminckt' [mutilated] <sup>13</sup>
<i>siloloa</i> 'to welcome someone'	<i>silalóa</i> 'bedancken met een solemniteit' [to thank with solemnity]
<i>sinanga</i> 'to fry'	<i>senanga</i> 'braeden inde panne; inde panne snercken'
<i>sowohi</i> 'official who supervises the royal insignia and palace servants'	<i>savoy</i> 'opvoeder, voester' [nourisher, wet nurse]
<i>tepa</i> 'to obstruct something' (De Clercq 1876)	<i>tepa</i> 'borstweeringe' [breastwork (fortification)]
<i>tifa</i> 'a drum'	<i>tifa</i> 'bommeken, trommeltgen'

Table 3. Ternatan loans in the Malay documented by Wiltens and Danckaerts.

The Malay documented by Wiltens and Danckaerts displays phonological characteristics which match the Malay contact varieties of eastern Indonesia more broadly.<sup>14</sup> These include the realizations of all word-final nasals as /ŋ/, the near-absence of the schwa /ə/, and the elision of word-final stops (apocope). Examples in the latter category include <*toula* 'daeghs nae overmorgen'>, <*poló* 'omhelsen'>, <*marata* 'klagen; weenen'>, and possibly <*papodjou* 'een sack daermen yets in steeckt'>, corresponding to modern Ml. *tulat* 'three days from now', *pəluk* 'to embrace', *məratap* 'to lament', and *punjut* 'wrapping cloth' respectively. In addition, some lexical items in *Vocabularium ofte Woort-boeck* reflect Ambon Malay specifically (Table 4).

<sup>12</sup> Note that some of the seventeenth-century Malay entries were so well understood in the Dutch of that period that they were not glossed at all in the early-modern dictionaries.

<sup>13</sup> See also <*sedó* 'gecke, verlemt, verminckt' [crazy, paralysed, mutilated]> (He), which might explain the somewhat morbid semantic connection with 'to joke'.

<sup>14</sup> See Scott H. Paauw (2008) for an overview of these varieties.

Ambon Malay	Attestations
<i>aringang</i> 'light of weight' (Ml. <i>ringan</i> ) <sup>15</sup>	<i>aringan</i> 'licht van gewichte'
<i>babalu</i> 'to forage' (Don van Minde 1997)	<i>babálou</i> 'gaen om den kost te soecken'
<i>balisa</i> 'restless' (Ml. <i>gəlisah</i> , Old Javanese <i>balisah</i> )	<i>balisa</i> 'woelende ende onrustich' (WD)
<i>bodok</i> 'stupid' (not in dictionaries; Ml. <i>bodoh</i> )	<i>bodóc, bodoc</i> 'plomp, niet vernuftich, een plompaert' [clumsy, unrefined, a clumsy person]
<i>caka</i> 'to choke'	<i>tsjacá</i> 'blijft ... steecken'
<i>dati</i> 'tax' (G.W.W.C. van Hoëvell 1876)	<i>dati</i> 'contributie, tol' (WD)
<i>hosa</i> 'out of breath'	<i>esa</i> 'amborsticheyt'
<i>ka</i> 'or'	<i>ca</i> 'ofte'
<i>loko</i> 'to grab'	<i>loco-loco</i> 'grabbelen'
<i>mawe</i> 'to investigate lost items or certain circumstances by using black magic'	<i>mawe</i> 'waer-segghen' [soothsaying]
<i>mongo-mongo</i> 'a boor who does everything wrong' (De Clercq 1876)	<i>mongho-mongho</i> 'slechtelijck, on-ordentlijck, ongeschickt, slordich, soo wat henen' [badly, disorderly, carelessly]
<i>papeda</i> 'a dish made from raw sago flour'	<i>papeda</i> 'dick warden gelijk als stijf sel'
<i>paris</i> 'pair'	<i>sa-páris</i> 'paer'
<i>peka-peka</i> 'fence' (Van Minde 1997)	<i>peca-peca</i> 'staketsels, palisade'
<i>sontong</i> 'squid'	<i>sontong</i> 'zee-katte'
<i>tindis</i> 'to press down' (Van Minde 1997)	<i>tindis</i> 'neder-douwen, neder-perssen'

Table 4. Ambon Malay loans in Wiltens and Danckaerts.

The dictionaries of Heurnius and Gueynier contain a number of entries which appear to reflect Peninsular Malay words (Table 5). Coastal cities like Patani, Sangora (Songkhla), and Ligor (Nakhon Si Thammarat), where the VOC had trading posts (*comptoirs*), presumably played an important role in their transmission. For example, the influential Bible translators Albert Cornelius Ruyl and Jan van Hazel studied Malay in Patani and Sangora respectively.

Peninsular Malay	Attestations
<i>asyik</i> 'a traditional dance in Kelantan and Patani' (← Ar. <i>āšiq</i> 'passionately in love') <sup>16</sup>	<i>assee</i> 'danseresse' [female dancer] (He), <i>aski</i> 'dansseresse' (G)
<i>baleter</i> 'to talk non-stop'	<i>baleter</i> 'blaes-kaecke, blaeter' [braggart] (G), <i>orang bleter</i> 'caeckelaer' [a babbler; <i>orang</i> = 'person'] (He)

<sup>15</sup> Also compare Ternate Malay *haringang* in the same meaning (Betty Litamahuputty 2012: 169). While modern Malay has *ringan*, the Malay dictionary of William Marsden (1812) gives *haringan*, *aringan*.

<sup>16</sup> See further in Table 25.

<i>dəning</i> 'yoke'	<i>danim</i> 't jock der beesten' (G)
<i>dəni</i> 'lintel of door' <sup>17</sup>	<i>darni</i> 'dorpel' (G)
<i>kənalīng</i> 'to tremble from fear or weakness or exertion' (Patani)	<i>kena-kenaling</i> 'vergiftigh gesicht' [poisoned face] (G)
<i>lerek</i> 'to bore through'	<i>larac</i> 'ael' [awl] (G)
<i>mu</i> 'you'	<i>mou</i> 'ghy' (He)
<i>nanggal</i> 'shaft of a plough'	<i>nanggal</i> 'ploegh-steert' (G)
<i>nayam</i> 'blade of a plough-share' (← Pe. <i>niyām</i> 'plough handle')	<i>najam</i> 'ploegh-yser, scheere'
<i>nondi</i> 'limp' (not in dictionaries; ← Tamil <i>nonḍi</i> 'crippled condition')	<i>nondi</i> 'manck gaende' (G)
<i>soh</i> 'body-warmth'	<i>soh</i> 'warm van salfs' [warm (of ointments)] (G)
<i>təpung</i> 'cakes and sweets made with flour' (Patani) <sup>18</sup>	<i>toupong, tepong</i> 'een broodt, coeck, deessem' (He), <i>tepong</i> 'coeck' (G)

Table 5. Peninsular Malay words in Heurnius and Gueynier.

Some words exhibit a phonological shape which can be identified as North-Eastern-Peninsular Malay. Gueynier's dictionary contains <*dakaw* 'plat als een schotel' [flat as a saucer]>, <*kikew* 'de geeren van een mans hemde onder de oxels' [the gussets of a man's shirt under the armpits]>, <*kitsjaw* 'kleyne kinck-hoorn' [small whelk]>, and <*sindjoh* 'verrucken' [to snatch away]>, which reflect MI. *dangkal* 'shallow', *kekeh* 'a short gusset', *kəcar* 'a shell', and *sənjah* 'to grab, to snatch angrily' respectively. All the attested sound changes – the elision of word-final liquids, the loss of nasals in nasal + voiceless stop clusters, and the vowel raising in *-ah* in the word-final position – are characteristic of the Malay varieties of this region (Jiang Wu 2023: 269-341). Another reason to envision that the Malay of Heurnius and Gueynier was influenced by the Kelantan-Patani dialect continuum is the number of Thai loanwords (Table 6). Most of these examples are also attested in Peninsular Malay, although more dialect research is needed to determine their precise distribution.

Thai	Peninsular Malay	Attestations
<i>chop</i> ๑๒๑ 'hoe'	<i>cop</i> 'spade, shovel'	<i>tsjop</i> 'spade, schuppe' (G)
<i>khu</i> คู่ 'pair'	<i>gu</i> 'pair, match' (Patani)	<i>gou</i> 'coppel' (G)
<i>khwam</i> ความ 'lawsuit'	<i>guam</i> 'dispute (in lawsuit)'	<i>baguam</i> (He) 'dingen, pleyten, procederen'
<i>kong</i> กง 'curve'	<i>kong</i> 'rib of a boat'	<i>cong cappel</i> 'ribbe van 't schip' (He)

<sup>17</sup> Also compare Hindustani *deorhī* 'threshold, porch', derived from Sanskrit *dehālī* 'threshold of a door'. If this word has yielded the Malay attestation, the associated phonological innovations require further explanation.

<sup>18</sup> In most Malay varieties, *təpung* means 'flour' rather than the cakes made from it.

<i>ma-hat-lek</i> มหาดเล็ก 'royal servant'		<i>mahat</i> 'lijf-wacht die 't bloote sweerdts draeght' [guard carrying a drawn sword] (G)
<i>na-tang</i> หน้าต่าง 'window'	<i>natang</i> 'French window'	<i>natang</i> 'vensterken' (G)
<i>phra-chao</i> พระเจ้า 'king'	<i>pəracau</i> 'royal'	<i>praetjouw</i> 'majesteyt' (He)
<i>wao</i> วาว 'kite'	<i>wau</i> 'paper kite'	<i>wou</i> 'vlieger van papier' (He), <i>wawou</i> 'vlieger van pampier' (G)

Table 6. Thai loanwords in Heurnius and Gueynier.

Javanese influence is found in all the dictionaries, albeit in relatively small quantities. The second-person polite pronoun *pakənira* features as <*packanira*> (vN), <*pacaniera*> (Ho), <*pakinira*> (WD), and <*pockenere*> (George Meister). Another widespread seventeenth-century Javanese loanword is *tarumpah*, *trumpah* 'a sandal', attested as <*taroupa*, *trompa* 'clickers van hout, clompen, platijnen' [wooden pattens, clogs]> (He), <*tarompah*, *tsjaroupa*, *taroupa* 'clickers, clompen, pantoffel, muyle' [pattens, clogs, slipper]> (G), modern Ml. *tərompah*, and several other borrowings in the Indo-Malayan Archipelago and the Philippines. Other Javanese loanwords are listed in Table 7. However, generally speaking, many common Javanisms in modern Indonesian Malay are absent in the early-modern material. Therefore, we find *esok* 'tomorrow', *itik* 'duck', *kətam* 'crab', *ləmbu* 'cow', *pəlanduk* 'mousedeer', and *pətang* 'evening' rather than *besok*, *bebek*, *kəpiting*, *sapi*, *kancil*, and *sore*. Along similar lines, *jamu* features in the meaning of 'feast', but never as 'herbal medicine'.

Javanese	Attestations
<i>blabur</i> 'an abundance of food'	<i>belábor</i> 'rantsoen van eetwaren ende yet diergelijcken' [food ration] (WD)
<i>cala-culu</i> 'bursting in'	<i>tsjala tsjoulou</i> 'clappen, weder seggen' [to chatter, to oppose in speech] (G)
<i>eling</i> 'to be mindful, to remember'	<i>éling</i> , <i>eling</i> 'gauw toesien; goede wachte houden; wel op sijn hoede zijn; erghens op loeren' [to keep a close watch] (WD)
<i>kunir</i> 'turmeric'	<i>connyr</i> 'safferaen' (vN)
<i>kuwung</i> , <i>kakuwung</i> 'rainbow'	<i>kong</i> 'regen-boge' (G)
<i>kyai</i> 'honorific title for men'	<i>quiha'hy</i> , <i>kyayi</i> , <i>quihay</i> 'heere' [lord] (Ho)
<i>lingsir</i> 'afternoon'	<i>lingsir</i> 'nae de middagh' (He)
<i>mənyan</i> 'incense'	<i>menján</i> 'wieroock, benjuwijn' (WD)
<i>pralente</i> 'lascivious' <sup>19</sup>	<i>parlent</i> 'schalck' [trickster] (vN), <i>parlente</i> , <i>paralente</i> 'schelm, rabaut, fielt' [rogue, rascal, villain] (WD)

<sup>19</sup> This is often claimed to be a Portuguese loanword, including in WD, but I have not yet encountered a detailed etymology. Hugo Schuchardt (1891: 117, fn. 153) proposes a derivation from Pt. *parlante* 'geschwätzig'. It is unclear to me whether such a form was ever used in a manner

<i>rubā, rāruba</i> 'a gift to gain a favour or bribe someone'	<i>rouba-rouba</i> 'gerechtigheydt die de schippers moeten betalen eerse yet moghen vercoopen' [fee paid by shippers before they are allowed to sell their goods] (He)
<i>susu</i> 'haste'	<i>sosou</i> 'heet van gaen' [hasty] (G)

Table 7. Javanese loans in seventeenth-century Malay.

## 2.2 INDIAN OCEAN

Beyond languages of the region, early-modern wordlists exhibit significant influence from Arabic. Interestingly, different sources paint a substantially different picture in this regard. Van Neck does not include many Arabic loanwords, but exhibits two examples not attested elsewhere: <*gyma* 'minnen'> from *jimā* 'sexual intercourse' and <*serpi* 'silver'> from *šarīfī* or *ašrafī* 'a coin'.<sup>20</sup> In the same source, we find <*mallim* 'pyloot'> in the meaning of 'navigator'. This word goes back to Ar. *mu'allim* 'master', which is also used in the meaning of 'navigator' in the Gulf and Oman (Dionysius A. Agius 2005: 130). In seventeenth-century Malay, this usage is confirmed by <*malim kapal, malin kapal* 'stierman, pyloot' [pilot; *kapal* = 'ship']> (Ho) and <*elmou malim* 'schip-manschap' [seamanship; *ilmu* = 'knowledge']> (G). In various sources, we find the Arabic borrowings *qirtās* 'paper', *wa-lākin* 'however', and *zabīb* 'raisin'. A number of textile terms are also widespread: *aṭlas* 'satin', *kattān* 'flax, linen', *muḵmal* 'velvet', *siqillāt* 'scarlet', and *šūf* 'wool'. Ml. *jam* 'hour' is also attested in the seventeenth-century dictionaries and reflects Ar. *zām* 'a unit of time'.<sup>21</sup> The Malay word *Jawi*, borrowed from the Arabic toponym *Jāwī* denoting Muslim Southeast Asia in general (Michael Laffan 2005), is attested as <*bahsa djawi* 'landst-spraecke' [the local language; *bahasa* = 'language']> (G).

Many Arabic loans in De Houtman belong to the domains of maritime knowledge and commerce (Table 8), reflecting the historical position of Aceh at the time.<sup>22</sup> In addition to these examples, I would postulate that <*kabitmal* 'factoor' [overseer]> (Ho) reflects the Arabic term *qābiḍ al-māl* 'tax collector', which is absent from the dictionaries I have been able to consult.

Arabic	Attestations
' <i>araq</i> 'a strong colourless liquor'	<i>arack apy</i> 'brandewijn' [ <i>api</i> = 'fire']
<i>a'wān</i> 'helpers'	<i>hawaen, hawan</i> 'toesien, waernemen, bewaren' [to watch, to protect]
<i>basbās</i> 'mace'	<i>bas bas</i> 'folie, macis'

consistent with its purported Malay and Javanese derivations. Its absence in Portuguese-lexified Creole languages raises scepticism about a Portuguese link.

<sup>20</sup> Hence *xarafim* or *xerafim* in Portuguese sources.

<sup>21</sup> This was originally a three-hour watch, and hence also a nautical measure (Pierre-Yves Manguin 1979; Agius 2005: 137). It is already attested as <站 *zhān* 'season, time, hour'> in the *Siyiguan* wordlist.

<sup>22</sup> See Drewes (1972: 453-457) for an earlier study and Drewes and Voorhoeve (1958) for more background on the actual usage of some of these terms.

<i>dawwār</i> 'spinning'	<i>dawar, dauwer, dauwar</i> 'draeyen, winden, spille, windaes' [to spin, to wind, spindle, windlass]
<i>farda</i> 'pair'	<i>fardā, fardah</i> 'bale'
<i>furda</i> 'seaport, river harbour, small port town' (compare Ar. <i>firda</i> 'tax')	<i>ferdach</i> 'tolhuys' [tollhouse]
<i>ḥarāmī</i> 'thief, robber, bandit'	<i>haramy, harammy</i> 'boef, guyt, rabbaut, schelm, vyleyn'
<i>ḥarf</i> 'letter'	<i>araf abc</i> 'letter, abc'
<i>ḥuqqa</i> 'compass'	<i>hocka</i> 'compas'
<i>miškil</i> (= <i>muškil</i> ) 'difficult'	<i>miskil</i> 'perijkel, questie' [danger, quarrel]
<i>ṭūfān</i> 'typhoon'	<i>toffæen</i> 'storm, onweder, stormwint'
' <i>ušr</i> 'tithe'	<i>ousour, ouchoer</i> 'tol'

Table 8. Examples of Arabic loanwords in De Houtman (largely from Drewes 1972).

A completely different situation is seen in Gueynier's dictionary, which contains numerous Arabic loanwords reflecting textual knowledge rather than a spoken usage (Table 9). Some of these entries are also found in an incomplete version of Leydekker's dictionary manuscript (1835), which was likewise based on written Malay and compiled within the same network of Bible translators. Even so, some attestations eventually gained wider currency, including *MI. khamir* 'leavened bread' (← *ḥamīr*), *najam* 'star' (← *najm*), *rubai* 'quatrain' (← *rubāʿī*), and *zabur* 'Psalms' (← *zabūr*). Many of the same Arabic loanwords are also included in the dictionary of Hermann von de Wall (1877), often without indicating their pronunciation. This situation leaves us rather unsure about how many seventeenth-century Malay speakers actually used (or even pronounced) them.

Arabic	Attestations
<i>asad</i> 'lion'	<i>asada</i> 'leeuw'
<i>baṭrīk</i> 'patriarch' (← Gr. <i>patriárkhēs</i> )	<i>bathric</i> 'patriarch'
<i>buṭm</i> 'terebinth'	<i>bouthom</i> 'terpentijn-boom'
<i>faṭīr</i> 'unleavened'	<i>fathir</i> 'ongesuyrde deegh'
<i>fīrṣād</i> 'mulberry'	<i>firdsad</i> 'moerbezie die soet is' [sweet mulberry]
<i>ḡarīb</i> 'unusual'	<i>garib</i> 'een dinck van waerden' [something of value]
<i>jāwars</i> 'sorghum' (← Pe. <i>gāwars</i> )	<i>djawars</i> 'hirsse'
<i>ḥamīr</i> 'leavened'	<i>chamir</i> 'gesuyrde deegh'
<i>karaza</i> 'to preach'	<i>karaza</i> 'preecken, prediken'
<i>kārīz</i> 'preacher'	<i>kariz</i> 'preecker, predicant'
<i>ḥarṭāl</i> 'oats'	<i>charthal</i> 'haver'
<i>kummaṭrā</i> 'pear'	<i>koumatsra</i> 'peer, vrucht-peere'

<i>najm</i> 'star'	<i>nadjam</i> 'sterre-kijker' [telescope]
<i>nusūr</i> 'vultures'	<i>nousour</i> 'ghier'
<i>qaṭāf</i> (= <i>qitāf</i> ) 'the vintage'	<i>cathaf</i> 'wimme'
<i>qītār</i> 'guitar, lyre'	<i>kitar</i> 'orgelen' [organ]
<i>rubā'ī</i> 'quatrain'	<i>rabiji, rabii</i> 'lof-dicht, gedicht' [eulogy, poem]
<i>rummān</i> 'pomegranate'	<i>roman</i> 'granaet appel'
<i>sabāsib</i> 'desert, wasteland'	<i>Hari Raja Elsabasib</i> 'Palm-Sondag-Feest' [ <i>hari raya</i> = 'holiday']
<i>sadāb</i> 'common rue'	<i>sadsab</i> 'wijn-ruyte'
<i>safarjal</i> 'quince'	<i>safardjal</i> 'que-appel, peere'
<i>ša'ir</i> 'barley'	<i>schangir, schoungour</i> 'gort'
<i>salwā</i> 'quail'	<i>salwa</i> 'quackel'
<i>sinjāb</i> 'grey squirrel'	<i>sondjab</i> 'eekhoren'
<i>ta'lab</i> 'fox'	<i>tsaalab</i> 'vosch'
<i>ṭalaja</i> 'snow'	<i>talaja</i> 'sneeuwen' <sup>23</sup>
<i>zabūr</i> '(Book of) Psalms, Psalter'	<i>zabour</i> 'psalm'
<i>za'farān</i> 'saffron'	<i>tsafaran</i> 'safferaen'

Table 9. Examples from Gueynier of Arabic loanwords in literary Malay.

Some loanwords are no longer attested in modern Arabic but can still be seen in older sources. Gueynier's <*saboucat* 'olier-boom'> corresponds to Ar. *sanbūqa* 'elder tree (*Sambucus nigra*)', which is known from medieval texts.<sup>24</sup> His <*loufahhat* 'mandragora'> appears to represent the archaic *luffāḥa* 'mandrake' rather than its modern equivalent *luffāḥ*. Gueynier's <*Elchamsina* 'Pentecostes, Pincxter-feest' [Pentecost]> reflects *kamsīna* 'fifty; Pentecost', which is an alternative form of *kamsūn* in the same meaning. His <*karadjat* 'accijs, schips-tol' [excise duty, ship toll]> goes back to *karājāt* 'poll taxes', which is an uncommon plural form of *karāj*.<sup>25</sup> The attestations <*toulah* 'de strafe Godts'>, <*mitar, moustar* 'regel om te meeten; linie'>, <*caif* 'fisionomus'>, and <*robiat* 'woecker'> are still found in the 1829 dictionary of John Richardson, which has *tawala, tuwala* 'a misfortune, a dreadful calamity', *mistar* 'ruler', *qā'if* 'physiognomist', and *rubīya* 'usury'.<sup>26</sup> Multiple seventeenth-century Malay dictionaries display variations of *gaibana* in the meaning of 'miserable, desolate', which reflects

<sup>23</sup> Corresponding to modern Ml. *salju* (Indonesia) and *salji* (Malaysia).

<sup>24</sup> This word is ultimately derived from Latin *sambucus* and is also found as *šabūq, šabūquh, šabuquh, sanbuq, sābūq, sunbuq*, and *sunbūq* (Albert Dietrich 1988: 680-681, n. 5). The Malay dictionary of Von de Wall (1877) has *sanbuqa* in the same meaning.

<sup>25</sup> The default form would be *akrija*, but see Joshua Blau (2006: 173) on the medieval attestation of *karājāt*. Also note that the latter form is common in classical Malay texts.

<sup>26</sup> Modern Arabic has *mistarā* for 'ruler'. Note also that *rubyah* in the meaning of 'usury' is also found in Von de Wall (1877). I suspect that modern Ml. *tulah* and *dulah* 'misfortune, calamity', which are often used interchangeably, result from contamination between Ar. *tawala, tuwala* 'a misfortune, a dreadful calamity' (R) and *dawla* 'turn of fortune'.

the obsolete Arabic word *ḡabāna* ‘infirmity of mind’ (R).<sup>27</sup>

Phonologically, the Arabic loanwords in Gueynier exhibit a number of noteworthy features. The voiced pharyngeal fricative /ʕ/ – known in Arabic as *‘ayn* – is frequently realized as a velar nasal /ŋ/, which is common in Javanese but not generally in Malay.<sup>28</sup> For example, <*bidangat*, *bidngat* ‘secte; ketterye’> and <*scheringat* ‘regel, wet’> – from Ar. *bid‘a* ‘innovation, heresy’ and *šarī‘a* ‘law’ – resemble Javanese *bidangat* and *saringat* more than modern Ml. *bidaat* and *syariah*. In a number of Arabic borrowings, the vowels show irregularities (Table 10). This strengthens my impression that many of these loanwords were adopted from written rather than spoken language, as Arabic vowel diacritics were often omitted in the Jawi script. Alternatively, some examples could reflect morphologically different source forms, but if so, I am unable to identify their precise etymological origins.

Arabic	Modern Malay	Attestations
<i>iklās</i> ‘sincere devotion’	<i>ikhlas</i>	<i>achalats</i> ‘trouw, getrouw’
<i>mašhūr</i> ‘well-known, renowned, famous’	<i>masyhur</i>	<i>mitchehor</i> ‘benaemt, deurluchtigh, hoogh-beroemt, ruchtbaer’
<i>qudra</i> ‘power, omnipotence (of God)’	<i>kodrat</i>	<i>cadirat Allah</i> ‘Goddelijcke macht’ [Allah = ‘God’]
<i>sikkīn</i> ‘knife’	<i>sikin</i>	<i>sakin</i> ‘een groot hangende mes’ [a large hanging knife]
<i>‘ubūda</i> ‘slavery’		<i>avodaet</i> ‘slavernye’
<i>zuhara</i> ‘Venus’	<i>Zohrah</i>	<i>bintang zahar</i> ‘avondt-sterre’ [evening star; <i>bintang</i> = ‘star’]
<i>zumurrud</i> ‘emerald’	<i>zamrud</i>	<i>zamouroud</i> ‘emeraude’

Table 10. Irregular vowels in Arabic loanwords in Gueynier.

Another relevant phonological feature is the realization of the *tā’ marbūṭa* <ﺕ>. This letter is silent in Arabic, unless it occurs in the construct state (*idāfa*) – that is, when it is followed by a possessor noun in a genitive phrase – in which case it is pronounced as /t/. In modern Malay, this grapheme is realized as /t/ in some words and /h/ in others, with the former pronunciation generally representing a more archaic layer.<sup>29</sup> Gueynier’s dictionary contains the greatest number of pronunciations with /t/ where modern Malay has /h/ (Table 11). I have found only one example of the reverse: De Houtman has <*toeffa* ‘presenten’ [gift]> for Ar. *tuhfa* ‘gift, present’, whereas modern Malay exhibits

<sup>27</sup> This also appears to be the etymon of Spanish *galbana* ‘laziness’.

<sup>28</sup> By contrast, nasalization of the vowels following a historical voiced pharyngeal fricative is common in North-Eastern-Peninsular Malay (Wu 2023: 50, 315) as well as other Malay varieties.

<sup>29</sup> See Stuart Campbell (1996), Nikolaos van Dam (2010), and Hoogervorst (2024) for more discussion, including on the possible role of Persian and/or Indian languages as vectors of borrowing.



*tuhfat*. However, the question arises whether some of the pronunciations listed in Gueynier are hypercorrections based on misspellings in the Jawi script. For example, <*dibadjat* ‘voor-reden’> goes back to Pe. *dībāja*, *dībāča* ‘preface’, which features the letter *hā* <◦> rather than <◊>. Similarly, <*elbabat* ‘paus’> (← Ar. *al-bābā* ‘the Pope’) exhibits a word-final *alif* <◁> in Arabic, whereas <*dangawaton* ‘daginge’> (← Ar. *da’wā* ‘lawsuit, legal proceedings’) is supposed to be written with an *alif maqṣūra* <◡>. In addition to a non-canonical /t/, the latter word contains the superfluous Arabic case ending *-un*, which is discussed below.

Arabic	Modern Malay	Attestations
<i>fā’ida</i> ‘utility, benefit’	<i>faedah</i>	<i>fahidat</i> ‘voor-deel’
<i>jumla</i> ‘totality, sum’	<i>jumlah</i>	<i>djomalat</i> ‘somme’
<i>ka’ima</i> ‘tent’	<i>kemah</i>	<i>chiamat</i> ‘herders-hutte’ [shepherd’s hut]
<i>nuqṭa</i> ‘dot, diacritical point’	<i>noktah</i>	<i>nocthat</i> ‘vocale’ [vowel]
<i>qubba</i> ‘dome’	<i>kubah</i>	<i>coubat</i> ‘gewelf, gewelfsel’
<i>ṣadaqa</i> ‘alms’	<i>sadakah</i>	<i>tsadekat</i> ‘aelmoesse, de tiende’
<i>ṣarī’a</i> ‘law’	<i>syariah</i>	<i>scheringat</i> ‘regel, wet’
<i>sayyāra</i> ‘planet’ (R)	<i>sayarah</i>	<i>sajarat</i> ‘planeet’

Table 11. Arabic loanwords with *tā’ marbūṭa* Gueynier.

A similar peculiarity is Gueynier’s transcription of the nominative indefinite case ending *-un* in several of Arabic loanwords. The associated diacritic, known as *ḍamma tanwīn* in Arabic and *baris dua dapan* in Malay, is rarely pronounced in the latter language. One might therefore suspect hypercorrection by a classically trained scholar eager to add declensions even in languages which do not have them. Most of the words in this category are highly uncommon in Malay. For example, Gueynier gives <*hhaboul mouloukon* ‘criecke’> from Ar. *ḥabb al-mulūk* ‘a sour cherry’, which designates multiple cherry varieties in the dialects of the Maghrib countries (northwest Africa) and Andalusia (Nawal Nasrallah 2018: 520). Also note that the correct case ending in the above construction should have been *-in* (genitive) rather than *-un* (nominative) according to the rules of Arabic grammar.

Another point of interest is semantic shifts. In various dictionaries, Ar. *ḡurāb* ‘crow’ exhibits its secondary meaning of ‘a ship’, while *ṣuḡl* ‘concern’ has chiefly come to denote ‘sadness’ in seventeenth-century Malay. The word *ḥāḍir* ‘present, attending’ has obtained the more specific meaning of ‘prepared’, as seen in <*boat adar* ‘bereyden’ [to prepare; *buat* = ‘to make’]> (vN), <*ader, adir* ‘ghereedt, veerdich’> (Ho), and <*hadir* ‘ghereedt’> (WD). The word *ḥarām* ‘forbidden’ appears to have yielded <*arrhām* ‘dreygen’ [to threaten]> (WD), whereas *iblis* ‘devil’ might have been the etymon of <*bilis* ‘diffameren’

[to defame]> (He).<sup>30</sup> Ar. *ḥadd* 'border, limit' exhibits a grammaticalized usage, as in <*hat 'juyst, pas* [just now]> (Ho), <*adda hadt 'tis tijt* [it's time; *ada* = 'is']> (Ho), and <*hat 'uytgenomen, behalven* [except]> (He). I suspect the entries <*faicat 'chiromantie* [chiromancy]>, <*safahh, safak 'wreedt* [cruel]>, and <*sorbat 'bier* [beer]> in Gueynier to reflect Ar. *fā'iqā* 'exquisite', *safk* 'to shed (blood)', and *šurba, šarba* 'a drink'.

Seventeenth-century Malay is also rich in Persian loanwords. Most of these still exist in the modern language,<sup>31</sup> while several others are no longer attested or do not have the same meaning (Table 12). This implies that the Malay language was more Persianized in early-modern times, which is confirmed by its classical literature and some cultural borrowings.<sup>32</sup> We can see this, for example, in the seventeenth-century Persianized titles <*Chayalam, cheyálam* 'Coning' [King]> (Ho) (← Pe. *šāh- 'ālam*) and <*djowan-pahalawan* 'opperste heldt, 't hooft van de ridders' [champion, head of knights]> (G) (← Pe. *jahān pahlawān*). In other cases, the Persian etyma have themselves become obsolete, as with '*arabāna* 'a drum, a tambourine' (R) borrowed as <*rabana* 'een kleyn trommelken'> (G).<sup>33</sup> Other entries ultimately reflect Arabic etyma, which must have been acquired through Persian on account of their phonology and/or semantics. Phonologically, the words <*moskit 'kercke* [church]> (Ho) or <*mesagit, mesadjid* 'kerck, tempel' [church, temple]> (G), and <*kitab taurit* 'Wet-boeck Mosis' [Torah; *kitab* = 'book']> (He) suggest borrowing from Pe. *mazgit* and *taurīt* rather than directly from Ar. *masjid* and *tawrā*.<sup>34</sup> Semantically, <*nadirat* 'phenixvogel' [phoenix bird]> (G) reflects Pe. *nādīra* 'rarity, phoenix' rather than Ar. *nādīra* 'rarity'. Semantic shifts also occurred after transmission into Malay. For example, Pe. *angūr* 'a grape' features as <*anghor* 'wijn-gaerd' [vineyard]> (WD) but has become *anggur* 'wine' in modern Malay.

<sup>30</sup> Also attested as <*bilis* 'bezwalken' [to besmirch]> (Von de Wall 1877). In colloquial Indonesian and Javanese, *bəlis* can denote 'troublemaker'. See also Adelaar (in this issue) for a semantic shift of the same Arabic etymon to 'soul' and 'hell' in early seventeenth-century Malagasy.

<sup>31</sup> This concerns the words *camca* (← *čamča* 'spoon'), *dəstar* (← *dāstar* 'cloth wrapped round the turban'), *haram zadah* (← *ḥarām-zāda* 'a bastard'), *kələbut* (← *kālbud* 'a model, form, a mould'), *kawin* (← *kāwīn* 'a dower, marriage-portion'), *kəlasī* (← *kalāšī* 'sailor'), *kimkha* (← *kimkā* 'damask silk in different colours'), *kurma* (← *kurmā* 'date fruit'), *mojah* (← *mūza* 'a boot; a stocking'), *naḥīri* (← *naḥīrī* 'a kind of trumpet'), *nakhoda* (← *nākhudā* 'sea captain'), *pəlita* (← *palīta* 'the wick of a candle or lamp'), *pinggan* (← *pingān* 'a bowl, a cup'), *piala* (← *piyāla* 'a cup, goblet, or drinking-glass'), *rəbab* (← *rubāb* 'a four-stringed instrument'), *sakar* (← *šakar* 'sugar'), *samsir* (← *šamšīr* 'a sword, scimitar, sabre'), *syahbandar* (← *šāh-bandar, šah-bandar* 'the chief of customs-house officer'), *səluar* (← *šalwār, šulwār* 'drawers reaching to the feet'), *sərahi* (← *šurāḥī* 'a long-necked flask, goblet'), and *sərunai* (← *surnāy* 'a trumpet, clarion'), all of which are attested in the seventeenth-century material. More Persian loanwords in Malay can be found in Russel Jones (2007).

<sup>32</sup> See Tomáš Petrů (2016) for a recent overview.

<sup>33</sup> See also modern ML *rəbana* and related forms in various other languages of Southeast Asia and the Indian Ocean (Hoogervorst 2013: 34, fn. 76).

<sup>34</sup> The corresponding Javanese forms are *məsigit* and *toret*.

Persian	Attestations
<i>bastā</i> 'bundle'	<i>basta</i> 'bale, pack' (Ho)
<i>darzī</i> 'tailor'	<i>durjy</i> 'cleermaker' (Ho)
<i>Frangī</i> 'a European Christian' (← Ar. <i>Ifranjī</i> 'Frank, European')	<i>sa orang Frangi</i> 'Portugijs' [a Portuguese; <i>səorang</i> = 'a person'] (He), <i>labo frangi</i> 'pompoene' [pumpkin; <i>labu</i> = 'gourd'] (He)
<i>gaz</i> 'a cubit'	<i>gas</i> 'ellen' (vN, Ho), <i>gasa</i> 'elle, mate' (G)
<i>gumbad</i> 'arch, dome'	<i>comoet miskijt</i> 'torren van een kerck' [tower of a church/mosque; <i>masjid</i> = 'mosque'] (Ho)
<i>halāl-zāda</i> 'legitimate child'	<i>halalsjada</i> 'echte kint' (WD)
<i>hindū</i> 'Indian; infidel'	<i>oerang hindoe</i> 'heydens' [heathen; <i>orang</i> = 'person'] (Ho)
<i>ijās</i> 'prune' (R)	<i>idjats</i> 'pruym' (G)
<i>kār-kun</i> 'a director, manager; an operator'	<i>cornon, korkon</i> 'schryver, secretaris' [scribe, clerk] (Ho)
<i>kavš</i> 'shoe'	<i>cous, kous</i> 'pantoffle, schoen' [slipper, shoe] (Ho), <i>cous</i> 'schoen' (WD), <i>kawts</i> 'pantoffel, muyle, schoen' [slipper, shoe] (G)
<i>kumīzdān</i> 'chamber pot'	<i>camistan</i> 'becken' (Ho)
<i>kušī</i> 'pleasure, happiness'	<i>tidor chousi</i> 'gerustelick slaepen' [to sleep peacefully; <i>tidur</i> = 'to sleep'] (He)
<i>līmūn</i> 'lemon, lime'	<i>limon</i> 'cytroen, lymoēn' (WD)
<i>māzaryūn</i> 'spurge olive'	<i>madjeloin</i> 'galnote' (He)
<i>panīr</i> 'cheese'	<i>pānir, panir</i> 'kase' (Ho)
<i>pāy-afrāz</i> 'a shoe'	<i>bai paras</i> 'schoen' (vN)
<i>rumāl</i> 'handkerchief'	<i>romal</i> 'servette' (vN)
<i>sar</i> 'amala' 'head of workers'	<i>sermala</i> 'schrijnwercker' [joiner] (Ho)
<i>tāfta</i> 'a garment woven from flax'	<i>taffa, taffeta</i> 'taffa, taffata' (Ho)
<i>zūr</i> 'power, violence'	<i>zur</i> 'gewelt, kracht' (He)

Table 12. Persian loanwords no longer used in modern Malay.

In the same way that many Arabic loanwords might have entered the Malay languages via Persian, many Persian loanwords reached Southeast Asia through the springboard of India. This is occasionally revealed by phonological clues. Malay <*tendour* 'oven'> (He) reflects Hindustani *tandūr* 'oven', which itself goes back to Pe. *tanūr* (ultimately from Akkadian) in the same meaning. Similarly, Malay <*māaf* 'aflaet'> (Ho) – corresponding to *maaf* 'forgiveness' in modern Malay – ultimately reflects Ar. and Pe. *mu'āf* 'forgiven, exempted', yet its phonological form suggests borrowing through Hindustani *māf* 'forgiven' (Hoogervorst 2024: 312). It is not immediately clear whether <*zyreh, zyreh* 'harnas, rustinghe tot krijch, wapen' [armour]> (Ho) reflects Hindustani *zirah* 'coat of mail' or its Persian precursor *zirih*, but its modern Malay equivalent *zirah* favours the former. The Arabic word *mušajjar* 'figured with designs of plants' acquired the meaning of 'damask' in Hindustani, which subsequently yielded seventeenth-century Malay <*mossayjar, mossayar* 'dammast, damast'>

(Ho). In other cases, the direct etymon of a word is more difficult to establish. Malay <*caffiri* 'heydenen' [heathens]> (He) and <*bandari* 'marckt-ganger' [person who goes to the market]> (He) appear to reflect *kāfirī* 'relating to infidels' and *bandarī* 'person belonging to a port town', but it is unclear whether such a hypothetical form was borrowed through Arabic, Persian, or Hindustani.

Early-modern Malay sources also exhibit many direct South Asian loanwords. De Houtman's <*iere manis, jieramanis* 'anijs zaet, venckelsaet, vincket' [aniseed, fennel; *manis* = 'sweet']> reflects Hindustani *jīrā* 'cumin-seed'.<sup>35</sup> The words <*tanty* 'gheweven, woven garen' [woven, woven yarn]> (Ho) and <*tantī* 'woven passement of diergelijcken' [weaving haberdashery or the like]> (WD) appear to go back to Hindustani *tāntī* 'weaver'. A rather enigmatic entry is <*morey* 'liedt' [song]> (Ho), which might have been mistranslated by De Houtman in view of <*mouri* 'fluytken van dun riet' [small flute made from a thin reed]> (He). This seventeenth-century Malay attestation can be connected to a similar type of instrument from the Indian subcontinent. In Orissa, this reed aerophone is known as *mohorī*.<sup>36</sup> Other widespread South Asian loanwords, from Hindustani or a closely related language, include *roti* (← *rotī*) 'bread' found in all the early-modern dictionaries and *cap* (← *chāp*) 'stamp, seal, mark' attested as <*si'ap, tsiap, tjap* 'merck'> (Ho). Also ubiquitous is the first-person pronoun *beta* from Hindustani *betā* 'son, boy, child', which is discussed in Section 4.2.

Tamil, a language spoken along the south-eastern coast of India, has also left an imprint. Leaving aside the many Tamil borrowings which are still used in modern Malay, early-modern dictionaries contain a number of obsolete words (Table 13). The attestation <*atjoeu* 'patroon nae te wercken' [pattern to imitate]> (Ho) and its Tamil precursor *accu* suggest an earlier \**acu*, whereas modern Malay exhibits only derived forms such as *acuan* 'mould'. In addition to the tabulated forms, modern Indonesian *capek* 'tired' goes back to Tamil *cappai* 'that which is weak, lean, emaciated', but in the seventeenth-century its primary meaning was 'lame' – see <*tjapick, tjapit* 'creupel, hincken, lam, paraltijck'> (Ho) and <*tsjapick* 'lam, paralatijcq'> (WD) – as remains the case in Peninsular Malay (see also Collins 2003: 21-22). *MI. kəməndikai* 'water melon' has been derived from Tamil *kommaṭṭi-kāy* 'unripe water melon' (Van Ronkel 1903; Hoogervorst 2015: 78) and De Houtman still contains a form without a prenasalized intervocalic stop: <*commedekey* 'kawoerde' [gourd]>.<sup>37</sup>

<sup>35</sup> Also compare Acehnese *jura, jara, jira* 'cumin'. Most other Malay dictionaries exhibit the related word *jintan* 'cumin', a formal by-form with the suffix *-ntan* (Hoogervorst 2017b: 301). The latter form is already attested in the *Siyiguan* wordlist as <正丹 *zhèngdān*> in the meaning of 'fennel [茴香]'.

<sup>36</sup> This instrument was known as a *madvarī* in medieval Indian texts. It features as *mahurī* or *muhurī* in a twelfth-century encyclopaedia, *Mānasollāsa*, and as *mauhārī* or *mohārī* in early Bengali (Reis Wenger Flora 1983: 212).

<sup>37</sup> However, prenasalized forms are attested as <悶的格 *mèndegé*> (*Siyiguan*) and <*mandicai*> (Pigafetta).

Tamil	Attestations
<i>accu</i> 'mould'	<i>atjoeu</i> 'patroon nae te wercken' (Ho)
<i>kucavan</i> 'potter' (← Pe. <i>kuzā</i> 'pot')	<i>cousawan</i> 'potbacker' (Ho)
<i>maṇvetṭi</i> 'hoe, spade'	<i>mamouti</i> 'spade om te graven' (G)
<i>vēci</i> 'courtesan, whore'	<i>bisi</i> 'hoere' (Ho)
<i>viricu</i> 'kind of rocket'	<i>berissó</i> 'een basse' (WD)

Table 13. Tamil loanwords in seventeenth-century Malay.

Seventeenth-century Malay also exhibits some Chinese loanwords, the majority of which can be identified as Southern Min (Table 14). One of the earliest examples might be ML. *cəngkih* 'cloves', already attested as <真皆 *zhēnjiē*> in the *Siyiguan* wordlist and <*ghianche*> by Pigafetta.<sup>38</sup> In addition, we encounter a widespread form of address corresponding to modern ML. *əncik*: <*entje* 'heere' [lord]> (Ho) and <*entsje* 'cock, cockinne, Monsieur, singnor, jouffrouwe' [cook, Mr, mistress]> (WD).<sup>39</sup> If this word goes back to the Southern Min term for 'uncle' (see Jones 2009: 196; Hoogervorst 2021: 34, 182, n. 28), the semantic shift to a more generic form of address invites further examination. Another possible borrowing from Southern Min is an unidentified term for 'camphor laurel (*Camphora officinarum*)', <*tjuden* 'campher van Sina of Japon'> (He), the first syllable of which resembles *chiu*<sup>n</sup> 樟 in the same meaning. The word <*lang* 'doodt-kiste' [coffin]> (G) might correspond to modern ML. *lung* in the same meaning, in which case the etymon would be Southern Min *láng* or *lóng* (depending on the dialect) 籠 'large bamboo trunk'. The expression <*rambot separti tangtchi* 'blondt' [blond]> (G) apparently consists of ML. *rambut* 'hair', *səpərti* 'like', and Southern Min *tāng-chi*<sup>n</sup> 銅錢 'copper coin', possibly reflecting an informant's impromptu attempt to describe this uncommon hair colour. In terms of notable omissions, the widespread Malay pronouns *gua* 'I' (← Southern Min *góa* 我) and *lu* 'you' (← Southern Min *lú* 汝) are absent in the early-modern dictionaries, although they did spread to Sri Lankan Malay (Nordhoff 2013: 22).

Southern Min	Attestations
<i>bāng-lô</i> 網羅 'fishing-net'	<i>bangloc</i> 'visch-korf' (G)
<i>chhiō<sup>n</sup>-kī</i> 象棋 'Chinese chess'	<i>tjoky</i> 'schaeckspel' (Ho) <sup>40</sup>
<i>kiáu</i> 傲 'to gamble'	<i>kiaw</i> 'caerte-spel' [card game] (G)
<i>ko-iōh</i> 膏藥 'herbal plaster, poultice'	<i>coyo</i> 'salve' [ointment] (He), <i>kojoc</i> 'salve' (G)
<i>pe-si</i> 飛絲 'gossamer'	<i>peessy</i> 'satijn' (He)

<sup>38</sup> See Mahdi (1994/1: 188) and Hoogervorst (2013: 68-69) on possible Southern Min etymologies.

<sup>39</sup> This word is also found as *ince* in the *Livro de Pantuns* and – in combination with the pronoun *ia* followed by velarization – as the third-person pronoun *incayang* in Sri Lankan Malay (S. Nordhoff 2013: 22).

<sup>40</sup> This word already features in the *Siyiguan* wordlist in which it was misidentified as a card game, *cəki*, by Edwards and Blagden (1931: 734-735).

<i>sam-pán</i> 舢板 'flat-bottomed boat'	<i>sampan</i> 'sloepen, jacht, sloup' (Ho)
<i>tháng</i> 桶 'barrel'	<i>tang</i> 'coel-vat, cuype, tobbe' (G)

Table 14. Southern Min loanwords in seventeenth-century Malay.

A significantly smaller number of Chinese loanwords appear to reflect Cantonese influence (Table 15). It is also worth pointing out that Southern Min *tê* 茶 'tea', a perennial example of a widespread loanword from this language, is absent from the seventeenth-century Malay dictionaries. We do find <亞衣茶 *yàiyīchá*> – with the first two characters representing *air* 'water' – in the *Siyiguan* wordlist and <*cha* 'tea'> in Thomas Bowrey's dictionary, which suggests that Cantonese *chàh* 茶 was the more common form in early-modern Malay. Interestingly, <羅巴 *luóbā*> (*Siyiguan*) in the meaning of 'turnip [蘿蔔]' reflects Cantonese *lòh baahk* 蘿蔔 'radish' (and modern Ml. *lobak*) rather than its Southern Min equivalent *chhài-thâu* 菜頭. Finally, the word for 'weighing scale, balance' – <*d'ajing, d'atjing* 'ghewicht'> (Ho), <*datsjin, datsjng, datsjim* 'gewichte, weeck-schale, schaele, balance'> (WD), and modern Ml. *dacin* (Indonesia) or *dacing* (Malaysia) – appears to be from a Sinitic language, but the precise etymology remains to be demonstrated (the second character is probably 秤 'scales').

Cantonese	Attestations
<i>dāng-lùhng</i> 燈籠 'lantern'	<i>tangelong</i> 'lanteerne' (He), <i>tanglong</i> (G) 'lanteerne'
<i>túng</i> 桶 'barrel'	<i>thong</i> 'tonne' (He), <i>tong</i> 'vat, tonne' (G)

Table 15. Cantonese loanwords in seventeenth-century Malay.

### 2.3 EUROPEAN

The seventeenth-centuries dictionaries contain various Portuguese loanwords (Table 16). Some of these early-modern examples constitute missing links between the Portuguese etyma and their modern Malay derivations. For example, Pt. *espingarda* 'rifle' yielded Ml. *istinggar, satinggar* 'matchlock gun' through such intermediate forms as <*espingard, bedil espingarde* 'roer te schieten' [rifle; *bədil* = 'firearm']> (Ho) and <*istingard, astingarda, estingarda, stingarda* 'busse, roer, *snap-haen*' [rifle, flintlock rifle]> (G). As the Appendix in Wiltens and Danckaerts indicates, the Portuguese influence was even greater in Ambon Malay.

Portuguese	Attestations
<i>bouba</i> 'yaws'	<i>boba</i> 'pocken' [smallpox] (WD) <sup>41</sup>
<i>calafate</i> 'caulker' (← Ar. <i>qilāfa</i> 'caulking')	<i>calfat</i> 'calfaten' (Ho)
<i>capitão</i> 'captain'	<i>capitaen, capitain, capitayn</i> 'capiteyn' (Ho)

<sup>41</sup> Also borrowed into Swahili as *mbuba* 'yaws'.

Castela 'Spanish'	<i>kastyla</i> 'Spaingaerden' (Ho) <sup>42</sup>
<i>capa</i> 'cloak, cape'	<i>capa</i> 'deecken' (WD)
<i>catana</i> 'Japanese sword' (← Japanese <i>katana</i> 'sword')	<i>catána</i> 'een Javaensche sabel' [Javanese sword] (He)
<i>dado</i> 'dice'	<i>bedadoe</i> 'dobbelen, spelen' [to play dice] (Ho)
<i>dedal</i> 'thimble'	<i>gidal</i> 'vingher-hoet' (WD)
<i>Deus</i> 'God'	<i>dios</i> 'Godt' (vN)
<i>feitor</i> 'overseer'	<i>fetor</i> 'factoor' (He)
<i>Francês</i> 'French'	<i>Francis, Francies</i> 'Franscoys' (Ho)
<i>fusta</i> 'a galley'	<i>fusta</i> 'fusten' (Ho)
<i>jogar</i> 'to play, to gamble'	<i>juggar</i> 'teerlinck-spel' [game of dice] (He)
<i>leilão</i> 'auction' (← Ar. <i>al-i lān</i> 'the announcement')	<i>lelong</i> 'vercopen by erf-huys ofte den meest biedenden' [to sell through a public auction room or to the highest bidder] (WD)
<i>padre</i> 'priest'	<i>padry</i> 'canonick' (Ho)
<i>Páscoa</i> 'Easter'	<i>Hari raja Pascha</i> 'Paesschen, Paesch-Feest' [ <i>hari raya</i> = 'holiday'] (G)
<i>tabaco</i> 'tobacco' (← Taíno)	<i>tabáco</i> 'tabac' (He)
<i>tartaruga</i> 'turtle' (← Gr. <i>tartaroukhos</i> )	<i>tetrougo</i> 'schild-padde' (WD)
<i>varanda</i> 'veranda, porch'	<i>baranda</i> 'afdack van een huys, gelderije' (WD)

Table 16. Portuguese loanwords in seventeenth-century Malay.

In addition, we find some more obscure Portuguese etymologies. The word <*estemy* 'brief' [letter]> (Ho) has been derived from Pt. *estima* 'esteem', although the semantic shift to 'letter' involved requires further qualification and alternative etymologies have been postulated.<sup>43</sup> Collins (1996: 196-197) argues that <*balio* 'gelubde' [eunuch]> (WD) goes back to early-modern Pt. *bailío*, *balío* 'bailiff, magistrate' through an intermediate meaning of 'court eunuch'. This word is often confused with Pt. *baileu* 'ship's forecastle', which the author points out yielded Ambon Malay *baileo* 'town hall'. In another publication, he derives <*secouta* 'borghe' [surety]> (WD) from Pt. *escolta* 'escort' (Collins 2004: 95). The Portuguese word *bastião* 'bastion' yielded <*bastion* 'loop-schanse'> (WD) in the same meaning, but also <*bastijong* 'busschieter' [cannoneer]> (Ho). In the latter case, the semantic shift from a location to the person working at said location is also seen in such Malay words as *kəbun* 'garden; gardener' (Lombard 1970: 255). The otherwise obscure term <*cartas sapzy* 'obligatie' [bond]> (vN)

<sup>42</sup> Also seen as <*oubi castella* 'patattes' [sweet potato; *ubi* = 'tuber']> (WD) and <*labo-castella* 'pepoen' [pumpkin; *labu* = 'gourd']> (He). Modern Malay has *kətela* in the meaning of 'sweet potato', which arrived in Asia through the Spanish galleon trade across the Pacific Ocean.

<sup>43</sup> The term is linked to Acehnese *esutumi* or *sutumi* 'letter with royal seal, royal edict' and classical Malay *istimi* or *sitmi*, and appears as *estemie* in Dutch sources. Frederick de Haan (1912: 776) suggested a Portuguese origin, while Gallop (Forthcoming) has recently proposed a connection to inscriptional Telugu, where the term *istimi* 'we have given' appears in the context of grants or bestowals from a ruler.

strikes me as a loan from Pt. *cartas de serviços* ‘service letters’ which were used to request a government grant (Sanjay Subrahmanyam 2005: 16). The word <*mainato* ‘wasscher’ [laundry person]> (He), corresponding to *bənatu* or *pənatu* in later Malay, reflects Creole Portuguese *mainato*, which ultimately goes back to a South Asian language. Finally, some obsolete Portuguese loanwords attested in seventeenth-century Malay can still be found in older Portuguese dictionaries, especially A. Alewyn and Joannes Collé (1718) (Table 17).

Early-modern Portuguese	Attestations
<i>belleguím</i> ‘servant of justice’	<i>belegyn ganton</i> ‘hangman’ [executioner; <i>gantung</i> = ‘to hang’] (vN)
<i>brága</i> ‘fetters’	<i>barágga</i> ‘boeye’ (WD)
<i>esfëra</i> ‘a round bullet’	<i>spera</i> ‘grof-gheschut’ (WD) <sup>44</sup>
<i>fitár</i> ‘to shoot at a target’	<i>mintar</i> ‘een geschut af sien ende stellen’ [to aim artillery] (WD), <i>mitar bedil</i> ‘t geschut passen’ [to calibrate artillery; <i>bədil</i> = ‘firearm’] (G) <sup>45</sup>
<i>Ingrís</i> ‘English’ (Pt. <i>Inglês</i> )	<i>engrees</i> ‘Enghels’ (Ho) (Ml. <i>Inggəris</i> )
<i>pelóuru</i> ‘iron bullet, iron ball’	<i>piloro</i> ‘bol, bal, bus-klooten, kogels’ (G) (Ml. <i>pəluru</i> )

Table 17. Obsolete Portuguese loanwords still found in Alewyn and Collé (1718).

Some of the most widespread Portuguese loanwords in modern Malay had apparently not yet entered the language in the seventeenth century, as inherited vocabulary and/or pre-existing borrowings were perfectly able to convey the same concept. Therefore, we find *gulung* ‘wheel’<sup>46</sup> instead of modern Ml. *roda* (← Pt. *roda*), *jumat* ‘week’ instead of *minggu* (← Pt. *Domingo* ‘Sunday’),<sup>47</sup> *minyak sapi* ‘butter’ instead of *məntega* (← Pt. *manteiga*), *panir* or *dadih* ‘cheese’ (← Sanskrit *dadhi* ‘curds’) instead of *keju* (← Pt. *queijo*), *kalam* ‘pen’ (← Ar. *qalam* ‘reed pen’), *bulu mənulis* or *bulu mənuyurat* instead of *pena* (← Pt. *pena*), and several words for ‘church’ predating *gəreja* (← Pt. *igreja*) (see Section 4.1).

The Dutch, who were relatively late arrivals in the Archipelago, had made even less of a dent in its seventeenth-century linguistic landscape. While De Houtman’s Malay dialogues contain the Dutch coin names *croonen*, *gulden*, *ponden*, *schellinghen*, and *stuyvers*, the conversations in which they occur have been lifted in their entirety from the *Vocabulare* – a 1536 phrasebook by the Belgian author Noël de Berlaimont – and therefore do not reflect the realities of the Indo-Malayan Archipelago.<sup>48</sup> Along similar lines, De Houtman’s <*ialan de Lombaert* ‘Lombaerde veste’ [Lombard Street]> reflects the topography of

<sup>44</sup> Also compare Ternatan *uspera* ‘cannon’.

<sup>45</sup> This would imply an intermediate form \*pitar, which underwent verbal prenasalization.

<sup>46</sup> Originally ‘roll’, but already referring to wheeled vehicles in Old Javanese.

<sup>47</sup> Both instances exhibit a semantic shift to the most important day of the week (Ar. *Jum’a* ‘Friday’, Pt. *Domingo* ‘Sunday’) to ‘week’ in general.

<sup>48</sup> See Van der Sijs (2000) on the influence of De Berlaimont’s *Vocabulare* on De Houtman’s *Spraeck ende woord-boeck*.



Antwerp rather than the Indo-Malayan Archipelago. Even the coin names *myte* and *cassen* – the latter from Tamil *kācu* ‘a coin; money’ – do not necessarily reflect any form of spoken Malay, even though they occur in original dialogues. They coexist with ephemeral Malay-Dutch hybrid constructions like <*orang duyts ‘een Duyts’* [a Dutch person; *orang* = ‘person’]>, <*orang Indiaen ‘een Indiaen’* [an Indies person]>, and <*tiere negry Oost Indien ‘op de Oost-Indische manier’* [in the East Indies manner; *cara* ‘manners’, *nəgəri* ‘country’]>. Even the authenticity of De Houtman’s <*bier ‘bier’* [beer]> might be called into question, as it only occurs in a dialogue lifted from De Berlaimont and not in any of the subsequent seventeenth-century dictionaries (despite the existence of modern Ml. *bir*).

One possible early Dutch borrowing is <*stratagema ‘crijghs-listigheydt’* [strategem]> (He), which is itself a borrowing from French or Latin and does not appear to have been widely used in Malay. Only Gueynier’s dictionary contains a modest number of credible Dutch loans: <*filosouf ‘philosooph’* [philosopher]>, <*myl ‘mijl’* [mile]>, <*oranji ‘oranje’* [orange]>, and <*elcatif ‘tapisserye, tapijt’* [carpet]>. The latter, *alkatief* in Dutch and *alkətip* in modern Malay, is itself a borrowing from Pt. *alcativa* and ultimately from Ar. *al-qatīfa* ‘the carpet’. As with Portuguese, the absence of Dutch loanwords, including those known from later times, is telling. Arabic equivalents were used for the names of the months, while we find *bea* or *biaya* ‘expenses’ instead of *ongkos* (← *onkost*), *bilik* ‘room’ instead of *kamar* (← *kamer*), *damaran* ‘lamp’ instead of *lampu* (← *lamp*), *gərudi* ‘drill’ instead of *bor* (← *boor*), *jamban* ‘latrine’ instead of *kakus* (← *kakhuis*), *juru dapur, pəngghulu dapur*, or *əncik dapur* ‘cook’ instead of *koki* (← *kokkie*), *kaca minum* ‘drinking glass’ instead of *gəlas* (← *glas*), *surat* or *kitab* ‘book’ instead of *buku* (← *boek*), and versions of <*espingarde*> or <*estingarda*> ‘rifle’ instead of *sənapan* (← *snapphaan*). Only later lexicographic sources, such as the 1780 *Nieuwe Woordenschat* and a roughly contemporaneous manuscript from Batavia, the *Syair nasihat belajar menyurat* (Poem of exhortations to learn to write), contain a greater yet still relatively modest number of Dutch loanwords. Additional examples, ignored by seventeenth-century lexicographers, can be found in numerous Malay manuscripts and letters.<sup>49</sup>

### 3. ARCHAISMS

Having discussed the lexical influence of local as well as foreign languages on seventeenth-century Malay, I now foreground a number of dictionary entries which can illuminate its phonological, morphological, and semantic history.

#### 3.1 PHONOLOGY

Early-modern sources contain various Malay words reflecting an earlier and/or regional stage of development (Table 18). We should also note a small number of words which display the segment /aba/, corresponding to /awa/

<sup>49</sup> See E.P. Wieringa (2005: 57-59) on Dutch loanwords in the *Syair nasihat belajar menyurat* and Suryadi (2015) for instances in Malay epistolary.

in generic Malay: <labang 'naghel, spijker'> (Ho), <bertabba, tertabba 'belachen, lachen'> (Ho), and <panabar 'medicijne'> (WD) instead of *lawang* 'nail', *tærtawa* 'to laugh', and *pənaʷar* 'medicine'. In addition, all the dictionaries and other Malay materials from this period document the absence of the /r/ in the word *pərgi* 'to go': <pigy 'gaen'> (vN), <pegy 'gaet'> (Ho), <peghí 'gaen'> (WD). The realization of this word as *pəgi* or *pigi* remains ubiquitous in Malay contact varieties to the present day.

Modern Malay	Attestations
<i>bəlarja</i> 'to learn'	<i>bereyer, bereyjar</i> 'leeren' (Ho)
( <i>bər</i> ) <i>niaga</i> 'to trade'	<i>viniága</i> 'handelen' (Ho), <i>vínjága</i> 'coop-handel doen' (WD) <sup>50</sup>
<i>bərsih</i> 'clean'	<i>bisey</i> 'schoon, fraey, pertinent' [clean, beautiful, accurate] (WD)
<i>cəguk</i> 'to hiccup'	<i>segóu</i> 'de hick' (WD)
<i>dəndam</i> 'spite'	<i>dam-dam</i> 'quaedtwiligheydt' (G)
<i>gasal</i> 'uneven in number'	<i>gassar</i> 'oneven in getal, ongelijck, onpaer' (He)
<i>hantar</i> 'to bring'	<i>antat</i> 'brenghen' (Ho), <i>antat</i> 'aanbieden' (WD) <sup>51</sup>
<i>istana</i> 'palace'	<i>asstana rayja</i> 'palleys des Coninck' [royal palace; <i>raja</i> = 'king'] (Ho), <i>astana-radja</i> 'Conincks-paleys' (WD), <i>astana</i> 'paleys' (G)
<i>laut</i> 'sea'	<i>laouer</i> [sic!] 'zee' (WD)
<i>məndəngar</i> 'to hear'	<i>menengar</i> 'verhooren' (G) <sup>52</sup>
<i>məntimun, kətimun</i> 'cucumber'	<i>antimon</i> 'comkommers' (vN) <sup>53</sup>
<i>mərpəti</i> 'dove'	<i>marapati, merepati</i> 'duyf, dyve' (Ho), <i>merepatti</i> 'duyve' (WD) <sup>54</sup>
<i>pəndek</i> 'short'	<i>panda</i> 'cort' (vN), <i>pendack</i> 'cort' (Ho), <i>pandak</i> 'cort' (WD)
<i>pəriuk</i> 'cooking pot'	<i>pryot, priot</i> 'ketel, pot' (Ho), <i>priot berkintjing</i> 'pis-pot' [chamber pot; <i>bərkəncing</i> = 'to piss'] (He)
<i>pintal</i> 'twisting, twining'	<i>pintar</i> 'tweernen, twijnen' (He)
<i>səmu</i> 'to deceive'	<i>orang hismou</i> 'bespieder' [a spy; <i>orang</i> = person], <i>hismouan seitan</i> 'verleydinge des Satans' [Satan's temptations; <i>Setan</i> = Satan], <i>berhismoukan</i> 'lagen legghen' [to deceive] (H)

<sup>50</sup> See also <*biniaga*> (Pigafetta). Etymologically, this word goes back to Prakrit *vāṇiyaga* 'trade; merchant, trader'. It did not originally have the prefix *bər-* in Malay, which is a back-formation (Hoogervorst 2017a: 413).

<sup>51</sup> Eighteenth-century Malay dictionaries still contain the form *antat*, as does Salako.

<sup>52</sup> The form *məndəngar* is common across Malay dialects and is also attested in Jawi manuscripts as well as <門能牙兒 *ménéngyá'er* 'listen'> (Siyiguan).

<sup>53</sup> Also see <*antimon* 'melon'> (Pigafetta).

<sup>54</sup> Also see *marapati* in the *Livro de Pantuns* and *pərapati* in the Tanjung Tanah manuscript. The word goes back to Sanskrit *pārāpati* 'female pigeon' (Jan Gonda 1973: 165; Hoogervorst 2017b: 302).

<i>səndi</i> ‘joint, sinew’	<i>sidy sidy, sidy</i> ‘leden, lidt’ [joints] (Ho), <i>sidi-sidi</i> ‘leden des lichaems’ [joints of the body] (He), <i>sidi-sidi</i> ‘knockel, lidtmaeten’ [knuckle, joints] (G) <sup>55</sup>
<i>undur</i> ‘to retreat’	<i>ondor</i> ‘achterwaerts’ (WD), <i>ondor</i> ‘achterwaerts gaen’ (G)
<i>untung</i> ‘luck’	<i>ontong</i> ‘avontuer, aventueren’ (Ho), <i>ontong</i> ‘aventuyr’ (WD), <i>ontong</i> ‘toeval, voorbate’ (G)

Table 18. Phonologically different forms in seventeenth-century Malay.

### 3.2 MORPHOLOGY

Early-modern data reveal or confirm the etymological origins of a number of words which were originally compound forms (Table 19). In addition to the listed examples, it has been argued that *ML. tənggala* ‘a plough’ reflects an earlier \**nənggala* (← Prakrit *naṅgala* ‘a plough’), later reinterpreted as a prenasalized verb (Hoogervorst 2017a: 396-397; see also Lombard 1970: 255). This hypothesis is confirmed by seventeenth-century Malay <*nəngala* ‘ploeghen, ploech’ [to plough, a plough]> (Ho) and <*de nəngala* ‘ploeghen’ [to plough]> (Ho); the latter is a passive form yet retains the word-initial /n/. We also encounter the stem <*gala* ‘ploegh’ [a plough]> (He) formed through back-formation, including in combination with the prefix *bər-* in <*bangala tana* ‘ackeren’ [to cultivate land; *tanah* = ‘land’]> (vN) and *məN-* in <*məngala* ‘ploegen’ [to plough]> (He).<sup>56</sup> Only Gueynier’s dictionary contains the modern form: <*tənggala* ‘ackeren, ploegen’>. Another noteworthy example is the word for ‘disease’. Modern *ML. pənyakit* ‘illness’ is reflected in <*pingakit* ‘cranck’ [ill]> (vN), <*penjakit* ‘sieckte’ [disease]> (Ho), and <*panjakiet* ‘de ziekte’ [disease]> (*Nieuwe Woordenschat*).<sup>57</sup> In addition, we find the apparent synonym \**sakitan*: <*sakytan* ‘sieckte’ [disease]> (vN), <*sakittan, sakitten* ‘hart, sieckelijck’ [difficult, sickly]> (Ho), and <*sakietan* ‘de ziekte’ [disease]> (*Nieuwe Woordenschat*), which has disappeared from the modern language.

Modern Malay	Etymology	Attestations
<i>bəgitu</i> ‘like that’	* <i>bagi itu</i>	<i>begy ytoe, bagy ytoe</i> ‘alsulcken, sulcx’ (Ho), <i>begytoe</i> ‘soo’ (Ho)
<i>bətapa</i> ‘in what manner; however much’	* <i>buat apa</i>	<i>boutapa</i> ‘hoe dat, hoe so, hoedanighlijck, waer mede, waer op, op wat wijze’ (G)
<i>dələpan</i> ‘eight’	* <i>dua alapan</i>	<i>doualapan</i> <i>pouloh</i> ‘tachtentigh’ [eighty; <i>puluh</i> = ‘ten’] (G) <sup>58</sup>

<sup>55</sup> Compare Old Javanese *sandhi* ‘joint, limb’. The ultimate etymon is Sanskrit *sam̐dhi* ‘junction, connection’.

<sup>56</sup> Note that the prefix *bər-* is consistently realized as *ba-* in vN, while <*ng*> in both vN and He can be used for /ŋ/ or /ŋg/. The form <*bangala*> would therefore seemingly correspond to \**bər-ŋgala* in standard Malay, which is ungrammatical.

<sup>57</sup> See Adelaar (1992: 178-184) on the development of the Malay prefix *pəN-*, which historically had a broader semantic meaning.

<sup>58</sup> This earlier pronunciation is also attested as <*dualapan*> (Tanjung Tanah), <都刺板 *dōulābǎn*> (*Siyiguan*), and *dualapan* in certain Jawi manuscripts.

<i>kəmarin, kəlmarin</i> 'yesterday'	*kəlam hari(an)	<i>kalmarin</i> 'gisteren' (Ho), <i>callamári</i> 'ghisteren' (WD), <i>kalmarin</i> 'daeghs te vooren, ghisteren' (G)
<i>kənapa</i> 'why'	*karəna apa	<i>karn'appa, carnappa, carna appa</i> 'waerom, hoe comt dat' (Ho)

Table 19. Original compound forms.

## 3.3 SEMANTICS

Various seventeenth-century glosses reveal earlier meanings than those attested in modern Malay (Table 20). Needless to add, the dictionaries are not always consistent in this regard. The word *cucu* means 'grandchild' in both standard Malay and in Proto-Malayic,<sup>59</sup> but we also find <*tzouzou* 'neef' [male cousin]> (vN), <*tjoetjoe lalakie* 'een neef' [male cousin; *ləlaki* = 'man']> (*Nieuwe Woordenschat*), and <*tjoetjoe parampoeang* 'een nicht' [female cousin; *pərəmpuan* = 'woman']> (*Nieuwe Woordenschat*). Modern MI. *halaman* 'lawn' was historically used in the meaning of 'main road' <*alámang* 'heer-baene, heeren-straete, den grooten wech'> (WD) or 'courtyard' <*alaman* 'voor-hof des volcks'> (G). The word *kancing* designates a button in modern usage, but originally referred to a bolt or any object to fasten something. For this reason, we encounter it in the metonymic sense of "prison" in De Houtman (Lombard 1970: 200). Modern MI. has *biru* 'blue' and *hijau* 'green', although these semantic categories tend to overlap in Malayic varieties and cross-linguistically.<sup>60</sup> Indeed, Van Neck exhibits <*birou* 'groen' [green]> and <*idgo* 'blau' [blue]>, whereas De Houtman has <*biero* 'blaeu' [blue]> and <*yjou, yiou, jioou* 'groen, blaeu' [green, blue]>.

Modern Malay	Earlier meaning	Attestations
<i>(bər-)səmayam</i> 'to sit on the throne'	'royal residence'	<i>somajam</i> 't speel-huys des Coninghs, vertreck des Coninghs' [royal playhouse, royal quarters] (G)
<i>cinta</i> 'to love'	'to worry, to think'	<i>tzinta</i> 'meynen' [to think] (vN), <i>tjinta</i> 'bedenckt, gewenschē, suchen' [to think, to wish, to worry] (Ho), <i>tsjinta</i> 'droef zijn' [to be sad] (WD), <i>tsjinta</i> 'achterdocht, sorge' [suspicion, worries] (He)
<i>cuci</i> 'to clean, to wash'	'clean, pure'	<i>tzoutze, tsiusi, tsiuzi</i> 'puyr, suyver, reyn' [pure, clean] (vN), <i>t'ioutjy, tjoutjy</i> 'schoon, cuysch, suyver' [clean, pure] (Ho), <i>tsjoutsji</i> 'schommelen, schoon, reyn' [to clean, clean, pure] (G)
<i>indah</i> 'beautiful, precious'	'to treat with respect'	<i>inda, berinda</i> 'sorge dragen, acht nemen' (WD) <sup>61</sup>

<sup>59</sup> Adelaar (1992: 122, 201) reconstructs \**cucu*? 'grandchild'. Its possible connection with Proto-Mon-Khmer \**cuu*? 'grandchild' has been mentioned by Shorto (2006: 79).

<sup>60</sup> See David Gil (2024) for a recent study on these colour terms in Malayic varieties.

<sup>61</sup> Compare <*ida* 'obey; respect'> (Tanjung Tanah) and modern MI. *məngindahkan* 'to pay attention to'.

<i>ingat</i> 'to remember'	'attentive'	<i>hing</i> 'at 'andachtich' [attentive] (Ho), <i>ing</i> 'at <i>ing</i> 'atken 'sorghen' [to ensure] (Ho) <sup>62</sup>
<i>kaya</i> 'rich'	'noble'	<i>kaya</i> 'edel, ryck' [noble, rich] (Ho), <i>parampouan kaya</i> 'jofvrou' [mistress; <i>pərampuan</i> = 'woman'] (Ho) <sup>63</sup>
<i>nenek</i> 'grandmother'	'grandparent'	<i>nyny laki</i> 'grootvader' [grandfather; <i>laki</i> = 'man'], <i>nyny parampouan</i> 'grootmoeder' [grandmother; <i>pərampuan</i> = 'woman'] (Ho) <sup>64</sup>
<i>orang hutan</i> 'a large ape'	'a jungle dweller'	<i>orang outan</i> 'wildt volck' [savage people] (H), <i>orang houtan</i> 'wildt volck' (G) <sup>65</sup>
<i>tuan</i> 'lord'	'royal personage'	<i>tu'an</i> 'heere' [lord] (Ho), <i>touan</i> 'jofferken' [young lady] (G)
<i>tunggu</i> 'to wait'	'to watch over'	<i>tongou</i> 'regeeren' [to govern] (Ho), <i>jure-tongou</i> 'opsiender' [overseer; <i>juru</i> = 'trained worker'] (He)

Table 20. Earlier meanings in seventeenth-century Malay.

Polysemantic words come with additional difficulties, as lexicographers rarely list all their meanings. Ml. *pəlihara*, *piara* 'to look after, to bring up' surfaces as <*palihara*, *plaira*, *playra*, *plara*, *piarra* 'seghende, opvoeden' [bless, to raise]> (Ho) and <*miāra* 'scheppen' [to create]> (WD), but once it might have encompassed all these meanings and more. Ml. *saja* or *sahaja* is glossed as <*sadja* 'allenlijck' [only]> (WD), <*sadja*, *saidja* 'vergeefs, om niet, louter, onghemengt, maer, alleenlick' [in vain, just, only]> (He), and <*sahdja*, *sadja* 'eygen wille, vrye wille, enckelijck, gewilligh, gewilligheydt' [deliberate, only, willingly]> (G), which undoubtedly only reflect a small fraction of the word's semantic range. The primary meaning of *jadi* is 'to become', but a number of phrases suggest that it also meant 'born': <*beta jádi berjádi* 'ick ben gheschaepen' [I am created]> (WD), <*jaddi derri kiël-kiël sondelnja* 'geboren uyt hoererye' [born of whoredom]> (He), and <*boutaja* [sic!] *deri pada djadjinja* 'hy is blindt geboren' [he was born blind]> (G). The word *kərja* primarily means 'work' in modern Malay, but could also designate a kind of ceremony as seen in <*bekerryja gawin* 'bruyloft' [wedding; *kawin* = 'to marry']> (Ho) and confirmed by <*karəja* 'a ritual ceremony'> in the

<sup>62</sup> Compare <宜牙 *yíyá* 'minutely, carefully'> (*Siyiguan*).

<sup>63</sup> Compare <烏郎加亞 *wúláng jiāyà*> (*Siyiguan*) for *orang kaya* 'master, lord (老爺)' [*orang* = 'person']. Also used as an honorific third-person pronoun in Heurnius <*orang caya* 'u'>.

<sup>64</sup> This is corroborated by <你你刺吉 *nīnī lāji* 'duke, gentleman, sir, Mr, male [公]'> (*Siyiguan*) and <你你不論般 *nīnī bùlūnbān* 'old woman [婆]'> (*Siyiguan*), <*niny* 'grandfather'> (Pigafetta), as well as <*nene lakie lakie* 'groot vader' [grandfather]> (*Nieuwe Woordenschat*) and <*nene parampoeang* 'groot moeder' [grandmother]> (*Nieuwe Woordenschat*). See Adelaar (1992: 118-119) on the word-final glottal stop in Malayic kinship terms and the Proto-Malayic reconstruction \**nini*? 'grandparent'.

<sup>65</sup> The *Nieuwe Woordenschat* has <*orang oetan* 'een bosch mensch' [a forest person]>, which is used synonymously with <*orang douloe calla*>, reflecting Ml. *orang dahulu kala* 'people of earlier times'. Wayan Jarrah Sastrawan (2020) points out that its better-known usage of "ape" was already attested in the first millennium CE, although "forest person" must have been its original meaning.

Tanjung Tanah manuscript.<sup>66</sup> The default meaning of *sambah* is ‘to worship, to pray’, but it also occurs as <*samba* ‘*dancken*’ [to thank]> as well as a number of idiomatic expressions in De Houtman.<sup>67</sup> In modern Malay, *pəranakan* denotes ‘womb’ or ‘person of mixed ancestry’, whereas its earlier meaning of ‘native to a particular place’ is still reflected in <*peranackan* ‘*inboorlingh*’> (G).

Predictably, the dictionaries contain some words no longer used or understood today (Table 21). In addition to these examples, the word <*sentaby*> (Ho) conveys politeness and can be translated as ‘please’ (Lombard 1970: 256). It is etymologically related to <*tabec*> (G), which was used as a greeting. Both words reflect developmental stages of the Sanskrit loan *kṣantavya* ‘to be pardoned’, which has been adopted into several Austronesian languages (Hoogervorst 2023: 43). De Houtman displays *bunga lawang* for ‘cloves’ <*bong’a lawang* ‘*nagelē*’>, which has been replaced by the word *cəngkih*.<sup>68</sup> The form <*gnappa*, *mengappa* ‘*tis de pijn waert*’ [it’s worth the pain]> (WD) is reminiscent of modern Ml. *tidak mənɡapa* ‘it doesn’t matter’, while the apparently related form <*gapa* ‘*waer toe*’ [to what end?]> (G) recalls modern Jakartan Indonesian *ngapain* ‘what are you doing? what’s the point?’. We might also call attention to some meaningful absences. The default word for ‘no’ is *tida*, whereas colloquial forms like *tra*, *trada*, *kagak* or *ənggak* remain unattested. The modern pronoun *məreka* ‘they’ only occurs in its earlier meaning of ‘people’: <*marikaitoe* ‘*lieden*’ [people; *itu* = ‘those’]> (G).<sup>69</sup>

Reconstruction	Attestations	Comparisons
*aluis ‘bird of prey’	<i>alowijs</i> ‘ <i>ghier, kieckendief</i> ’ (WD), <i>alouwis</i> ‘ <i>wauwe, kiecken-dief</i> ’ (He)	<i>aloewis</i> ‘ <i>een kieckendief</i> ’ (Von de Wall 1877)
*antəlih ‘throat’	<i>antelē</i> ‘ <i>strot</i> ’ (WD)	<i>antelai</i> , <i>antelih</i> ‘ <i>de strot</i> ’ (Von de Wall 1877), compare Javanese <i>təlih</i> ‘crop of a bird’
*bali ‘monkey’	<i>baly</i> ‘ <i>ape</i> ’ (Ho), <i>bali</i> ‘ <i>meercatte</i> ’ (WD)	<i>bali</i> ‘ <i>meerkat, aap</i> ’ (Von de Wall 1877) <sup>70</sup>
*bilas ‘disassembled’	<i>bilas</i> ‘ <i>ontschicken</i> ‘ <i>thuysraet</i> ’ [to disassemble furniture] (WD)	<i>bilas</i> ‘deranged, out of order of condition’ (Marsden 1812)

<sup>66</sup> This double meaning is already seen in its Sanskrit etymon *kārya* ‘work or business to be done, duty, affair; a religious action or performance’. See Adelaar (In press) for more background and several analogous examples in Malayo-Polynesian languages.

<sup>67</sup> These are <*minta* *somba* ‘*bede*’ [to pray, to request; *minta* = ‘to ask for’]>, <*boat* *somba* ‘*nyghen*’ [to bow; *buat* = ‘to do’]>, <*somba* *fittenah* ‘*verclicken*’ [to defame; *fitnah* = ‘slander’]>, <*sombahan* ‘*present; offeren*’ [present; to offer]>, and <*tersomba* ‘*ghebeden*’ [prayed]> (Lombard 1970: 233).

<sup>68</sup> In modern Malay, *bunga lawang* came to mean ‘star anise’. See Mahdi (1994/1: 188) and Hoogervorst (2013: 68-69) on the adoption of this word into various Indian Ocean languages.

<sup>69</sup> This word has been derived from the Old Javanese distal demonstrative *marika* ‘that; those (people)’ (Adelaar 1992: 125-126).

<sup>70</sup> There might be an etymological connection with *Bālī*, the simian king in the *Rāmāyaṇa* epic, as suggested by Lombard (1970: 259).

*bolelo 'to scatter, to overturn' <sup>71</sup>	<i>bolelò barang barang</i> 'tgoed verhoijen, 'tonderste boven keren' [to scatter goods, to overturn; <i>barang</i> = 'goods'] (WD)	<i>bolelo</i> 'het onderste boven keeren, verhoorien' [to overturn, to scatter] (Von de Wall 1877)
*sa 'one'	<i>dua poelosa</i> 'eenentwintich' [twenty-one] (Ho)	also encountered in Jawi manuscripts; compare ML. <i>asa</i> 'God's unity' and the ML. prefix <i>sə-</i>
*gānggut 'to doze'	<i>gangout</i> 'suffen, suffelen' (G)	compare ML. <i>gəlānggut</i> 'to doze'
*karəna Allah 'for God's sake'	<i>Carna Alla</i> 'Om Gods wille' (Ho)	compare Cape Malay <i>kanala</i> 'please' (Hoogervorst 2021: 44)
*nijis 'carpenter's plane'	<i>ber ny'is, bernyjis</i> 'afschaven, schaven' [to plane] (Ho), <i>nidijs</i> 'schaef' [carpenter's plane] (WD)	compare Acehnese <i>nyheh</i> 'carpenter's plane'
*pupu 'great-grandparent' <sup>72</sup>	<i>poupou</i> 'trap van gheslachte' [degree of kinship] (WD)	<i>poepoe lakie lakie</i> 'over groot vader' [great-grandfather], <i>poepoe prampoeang</i> 'over groot moeder' [great-grandmother] ( <i>Nieuwe Woordenschat</i> )
*təgal 'because'	<i>taghal</i> 'om' (WD)	common up to the 1950s and still attested as <i>tagal</i> in Ambon Malay.
*tərleleh 'uncovered'	<i>talille</i> 'naeck, ongedeckt' [naked, uncovered] (WD), <i>bertelili</i> 'vertellen, verhalen' [to tell, to reveal] (He)	<i>lēleh</i> 'uncover, strip, expose; explain' (Marsden 1812)
*tupas 'tree shrew'	<i>toupas</i> 'weselken' [a small weasel] (He)	<i>monjet toupas</i> 'spookdier' [tarsier; <i>monyet</i> = 'monkey'] ( <i>Nieuwe Woordenschat</i> ); compare ML. <i>tupai</i> 'tree shrew'

Table 21. Obsolete words in early-modern Malay.

#### 4. SOCIETY

Seventeenth-century dictionaries serve as valuable tools for reconstructing aspects of the societies in which they were produced. Given that most early lexicographers were funded and employed by the VOC to translate the Bible, it is no surprise that a significant part of the terminology documented encompassed the realm of Christianity. Bible translators actively created and promoted neologisms, which included the repurposing of pre-existing words – inherited as well as those borrowed from Sanskrit, Arabic, Portuguese,

<sup>71</sup> This does not resemble a Malay word. The closest match I have been able to find is Ambon Malay *bolengkar* 'to rummage'.

<sup>72</sup> R.J. Wilkinson (1932) glosses *pupu* as 'grade or degree of relationship measured by removes from a common ancestor'. Modern Malay has *sə-pupu* 'cousin', which presumably denotes 'having the same great-grandparent'.

and other languages – to convey biblical concepts.<sup>73</sup> In what follows, several examples from the seventeenth-century materials are discussed.

#### 4.1 RELIGION

The earliest available source, Jacob van Neck's *Vocabulaer*, does not delve deeply into the religious realm. Its compiler evidently had trouble eliciting some key concepts. His wordlist contains <*dilanghyn* 'hemel'> for 'heaven' and <*del auatana* 'helle'> for 'hell', reflecting ML. *di langit* 'in the sky' and *di bawah tanah* 'under the ground'.<sup>74</sup> All other early-modern dictionaries exhibit the default Malay forms *surga* (← Sanskrit *svarga*) and *nāraka* (← Sanskrit *naraka*) for these respective concepts. For 'God', <*dios, Deos*> (← Pt. *Deos*) coexists with <*Alla, Allah*> (← Ar. *Allāh*), while *Tuhan* 'Lord' is seen in the sentence <*ako Allah Touhanmou* 'ick ben de Heeren uwen Godt' [I am the Lord, your God; the Malay translates as 'I am God, your Lord']> (G).

Bible translators did not always agree on the preferred Christian terminology, with some individuals preferring a rather erudite idiom with Arabic and Persian loanwords ("High Malay"), others a more colloquial language with European loanwords ("Low Malay"), and yet others a sort of middle ground. This is also reflected in the words they chose to include in their dictionaries. The work of Heurnius, for example, contains some Portuguese loans (Table 22).

Portuguese	Attestations
<i>apóstolo</i> 'apostle'	<i>apostolo</i> 'apostel'
<i>Cristão</i> 'Christian'	<i>Ecclesia Christaon</i> 'De Christelicke Kercke' [the Christian Church; <i>eklesia</i> = 'church']
<i>grego</i> 'Greek'	<i>orang-Grego</i> 'Griek' [ <i>orang</i> = 'person']
<i>paraíso</i> 'paradise'	<i>paraíso</i> 'paradijs'
<i>sacramento</i> 'sacrament'	<i>sacramento</i> 'sacrament'

Table 22. Portuguese loans for Christian terminology in Heurnius.

Heurnius and Gueynier also included a small number of Greek borrowings (Table 23), whose existence might have been limited to the direct circles of Bible translators. In addition, Heurnius contains the word <*Euangelium* 'Euangelie'>, from Ecclesiastical Latin (and ultimately Greek) *Evangelium* 'the Gospel', along with the derivations <*Euangelista* 'Euangelist' [evangelist]> and <*Euangelician* 'euangelisereen' [to evangelize]>. The Malay of Bible translators also reveals a small number of Hebrew loanwords. The word <*bedola* 'bdellium'> in Brouwerius' translation of Genesis reflects Hebrew *bdólakh* 'bdellium', possibly

<sup>73</sup> For concrete examples of this commonly observed phenomenon, refer to Collins (2004) and De Vries (2018).

<sup>74</sup> De Houtman's wordlist contains similar instances of apparent confusion between him and his Malagasy consultant in the religious realm (see Adelaar, in this issue).



through seventeenth-century Dutch *bedolah* (Collins 2004: 95, fn. 40). Heurnius' dictionary has <*sabith* 'sabbath'>, which might reflect Hebrew *šabbāt* 'Sabbath' if not Ar. *sabt* in the same meaning. The etymology of <*kebith* 'kercke' [church]> (He), which also features in early Bible translations in the meaning of 'temple', is more ambiguous. A derivation from Hebrew *habbayit* 'the Temple' would raise questions about the associated correspondence between the word-initial consonants. Another possible etymon is Ar. *ka'ba* 'object of veneration', which also appears to have yielded <*kahebah*, *kabah* 'kerk' [church]> (*Nieuwe Woordenschat*). Less ambiguous is <*Tuan Zebaoth* 'Heere der Heyrscharen' [God of Hosts]> (He), which is a loan blend consisting of Ml. *Tuhan* 'God' and Hebrew *ts'vaot* 'host (armed forces)'. The original Hebrew expression is *Elohé Ts'vaot* 'God of Hosts' and one imagines that its partial translation into Malay reflects an impromptu coinage rather than a fixed expression.

Etymon	Attestations
<i>apsínthion</i> 'wormwood'	<i>afsintinon</i> 'alssem' (G)
<i>ekklésia</i> 'community'	<i>ekkklesia</i> 'gemeynschap' (He)
<i>kánkamon</i> 'kataf tree ( <i>Commiphora kataf</i> )'	<i>kamkamon</i> 'cinamome, canneel' [cinnamon] (G)
<i>Khristós</i> 'Christ'	<i>Christos</i> 'Christus' (G)
<i>órganon</i> 'organ'	<i>organon</i> 'orgelen' (G)

Table 23. Biblical terms from Greek.

As the language of Islam, Arabic already possessed many terms which could be used in a Christian context (see Collins 2004: 98-102; De Vries 2018: 66). Here, again, different lexicographers had different preferences. For example, 'Bible' was translated as <*Kitap-Agamma*> (He) 'the Book of Religion', <*ferman Allah*> (G) 'the Decree of God', or <*Alkitab*> (G) 'the Book'. The latter term is still commonly used in modern Malay. The terms <*beytoul* 'tempel' [temple]> (He) and <*orang chlaya* 'heyden' [heathen; *orang* = person]> (He) – apparently reflecting Ar. *bayt al-muqaddis* 'sacred house' and *kalā'iq* 'creatures' – were dismissed by Leydekker as unintelligible neologisms (Niemeijer and Van den End 2018b: 139).

Early Bible translators did indeed occasionally make up words. De Vries (2018: 67-68) lists various new terms formed by adding the suffix *-an*. In the work of Heurnius and his associates, the suffix *-ni* or *-ani* served a similar function. Its function was apparently to form adjectives, including nominalized adjectives, from Arabic words (Table 24). We can assume that these derivations were inspired by Arabic pairs such as *rūḥ* 'spirit' and *rūḥānī* 'spiritual', which occurs in Heurnius as <*roahni* 'geestelick'>. However, some non-Arabic words were also derived in this way: <*nang berpitsjayahani* 'de geloovige' [the believer]> (He) from *pərcaya* 'to believe' and <*terpileharani* 'uytverkorene' [the chosen one]'> from *tərpəlihara* 'looked after'.

Etymon	Suffixed form
<i>dunyā</i> 'world'	<i>dunjani</i> 'wereltlick, wereltdts' [worldly]
<i>ḥayrān</i> 'confused'	<i>cardja herani, bacarjahan heirani</i> 'wonder-werck' [wonderwork; <i>kərja</i> = 'work']
<i>nabi</i> 'prophet'	<i>nabbi-nabbini</i> 'propheten' [prophets]
<i>salām</i> 'peace'	<i>salamani</i> 'saligh-maker der werelt' [saviour of the world]
<i>sulwān</i> 'comfort'	<i>saluwani-dunja</i> 'saligh-maker der werelt' [saviour of the world; <i>dunia</i> = 'world']

Table 24. Arabic-derived neologisms in Heurnius formed through the suffix *-(a)ni*.

Gueynier's use of the suffix *-at* represents another way to form Arabic-derived neologisms. Presumably informed by such pairs as *MI. akhir* 'end' (← *Ar. ākir*) and *akhirat* 'Hereafter' (← *Ar. ākira*), the addition of this suffix allowed Bible translators to form abstract nouns which do not exist as such in Arabic, including <*kafirāt* 'ongeloove, valsche Godts-dienst' [unbelief, false religion]> (G) from *Ar. kāfir* 'unbeliever' and <*thabibat* 'medicijne' [medicine]> (G) from *ṭabīb* 'doctor'. Semantically, the word <*chalifat* 'hooge priesterschap' [high priesthood]> (G) resembles Dutch *kalifaat* 'caliphate' rather than *Ar. kalifa* 'caliph', which would imply that Gueynier introduced a hypercorrected Arabic form in his Malay dictionary (the corresponding Arabic word is *kilāfa*).

I have found no satisfactory etymology for some religious terms (Table 25). As these words are not to my knowledge attested outside the biblical realm, it is difficult to reconstruct how they entered seventeenth-century Malay.

Attestations	Comments
<i>askijt, askyt</i> 'abel, bevallichlijck, geestich, vermaeckelijck, wellust, playsier' [pleasant, lust, pleasure] (WD)	Collins (2004: 99-101) derives this word from <i>Ar. āšiq</i> 'passionate love'. Also compare < <i>askai</i> 'aangenaam, vrolijk' [pleasant, cheerful]> (Von de Wall 1877)
<i>berbistouwi</i> 'ergernisse, verergeren, sich yewers aen ergeren' [annoyance, to be annoyed] (He)	also written as < <i>berbistuwi</i> > in early Bible translations
<i>golesonda</i> 'hypocrijt, geveynst' [hypocrite, feigned] (He)	
<i>horgay</i> 'vremt, vreemdelinck' [alien] (WD)	also written as < <i>horgai</i> > in early Bible translations
<i>mohsina, mousina</i> 'algemeyn, gemeynschap' [common, community] (He)	compare <i>Ar. muḥsina</i> 'of unblemished reputation'
<i>plauwan baring</i> 'oogst' [harvest] (He)	<i>MI. baring</i> = 'lie down'
<i>tjangla, batjangla</i> 'weyden, hoeden' [pastures, to let graze] (He)	compare <i>MI. jənggala, cəngkəla</i> 'forest' (← <i>Sanskrit jāṅgala</i> 'rough and arid terrain')
<i>tjel-gagga</i> 'adder-slange' [viper] (He)	

Table 25. Words attested only in the publications of Bible translators.

## 4.2 SOCIAL STRATIFICATION

Seventeenth-century Malay dictionaries contain some clues to social stratification. Slavery, serfdom, and other types of unfree labour, which consisted of inherited forms as well as practices introduced by Asian and European merchants and rulers, were prevalent in early-modern times. It is often claimed that Indonesian slavery was of a fundamentally different nature to its European counterpart and in some aspects bore a closer resemblance to servitude. Semantically, enslaved people, servants, seasonal workers, subjects, followers, and henchmen were indeed overlapping categories. Besides this, many of the associated words could be used as polite first-person pronouns (“your humble servant”). Therefore, Ml. *hamba* features as <*hamba, amba, hamma, amma* ‘slave, ick’ [slave, I]> (Ho), <*amba* ‘slave, ick’> (WD), and <*hamba* ‘ick, knecht, slave, slavinne [I, slave, servant]> (G), whereas *saya* or *sahaya* is attested as <*saya* ‘ick’ [I]> (WD) and <*sahja* ‘ick, knecht, slave, slavinne [I, servant, slave]> (G). Heurnius contains the polite first-person pronoun <*saya oulun* ‘ick’>, consisting of *saya* followed by *ulun* ‘slave, servant, I’.<sup>75</sup> As mentioned previously, the Malay first-person pronoun *beta* goes back to Hindustani *betā* ‘son, boy, child’. The associated semantic shift becomes clear upon recalling that *beta* meant ‘servant’ in seventeenth-century Malay, as seen in <*jure betā besar* ‘hof-meester’ [steward; *juru* = ‘skilled worker’, *basar* = ‘big’]> (He) and <*oran beta* ‘dienaer’ [servant; *orang* = ‘person’]> (vN) (see also Adelaar, in this issue).

Another common word found across the Indian Ocean was *laskar*, ultimately derived from Pe. *laškar* ‘army’. In many South Asian languages, *laškar* referred to both soldiers and sailors. The word was presumably acquired in this meaning from a European language rather than directly from Persian.<sup>76</sup> In early-modern Malay, the word primarily denoted ‘slave, servant’: <*lascar* ‘slave’> (WD), <*laskar, laschkar* ‘knecht, slave, slavinne, trauwant [servant, slave, henchman]> (G), <*orang Laeskar* ‘Sclaven’ [slaves; *orang* = ‘person’]> (George Meister), and <*hazal lascar* ‘slaven geslagt’ [slave-descended; *asal* = ‘origins’]> (*Nieuwe Woordenschat*). Javanese *laškar* is already attested in Van Neck’s Dutch-Malay-Javanese wordlist as <*lascar* ‘dienaer’ [servant]>. The associated semantic shift – from soldier or sailor to slave or servant – probably took place in early-modern times, as the VOC often used South Indian “lascars” as reserve militias and ship’s crews. Many of these people were effectively serfs, shaping the meaning of *laskar* in the Indo-Malayan Archipelago (De Haan 1912: 770-771).

Various other words for subservient people are found in the early-modern material. The Malay word *sakai* ‘subject, dependent’, from an earlier \**sakay*

<sup>75</sup> Malay *saya* is from Sanskrit *sahāya* ‘companion, follower’, while *ulun* is connected to *hulu* ‘upriver (people)’ in Malay and related languages (Alexander Adelaar and John Hajek 2024: 919).

<sup>76</sup> In Persian, ‘soldier’ would have been *laškarī*, which does indeed feature in early-modern Portuguese as *lascarim, lasquarin, or lascari*. In Tamil, *laskar* had the additional meaning of ‘labourer, tent-pitcher, peon’.

'stranger' (Robert Blust 2010: 76-77), features as <*sakey, sakai, sakei* 'knecht, *cnape, dienaer, ondersaten* [servant, errand-boy, subjects]> (Ho). Classical Malay *dangan* 'slave, servant' is found as <*dingán, dengán* 'slave'> (WD). A specific term appears to have been used by two people sharing the same enslaver, <*sombóa* 'seggen twee slaven tegen makanderen die een meester dienen'> (WD), the etymology of which eludes me. Ar. 'abd' 'slave, servant' has yielded modern Malay and Javanese *abdi*, which is absent in the seventeenth-century material. However, we do find <*abdoun* 'slave, *slavinne*'> (G) from the same precursor, probably reflecting written rather than spoken Malay. The word *budak*, which denotes 'slave' in modern Indonesian, is not attested in this meaning. In early-modern Malay, it designated boys – as it still does in Malaysia – or young people in general: <*boudack* 'jonghen' [boy]> (Ho), <*boudac* 'jongen, meysken' [boy, girl]> (WD), and <*boudac* 'trauwant' [henchman]> (G).

The seventeenth-century Malay word for a free person was <*mardeka* 'vryman, geen slaef zijn' [freeman, not to be enslaved]> (Ho), <*mardéca* 'vrij-man die geen slave en is' [a freeman who is not enslaved]> (WD), or <*maredheyka* 'borger' [citizen]> (G), whereas someone born free was <*atsal maredheyka* 'vry-geboren' [freeborn; *asal* = 'origins']> (G). This word predates colonialism. In Old Javanese, *maharddhika* (← Sanskrit *maharddhika* 'very prosperous or powerful') was a sage or holy man. As a result, this person was exempt from forced labour, taxes, and other forms of subservience (Gonda 1973: 505). Freed slaves, known as *mardicas* in Portuguese and *mardijkers* in Dutch, constituted an important part of the early-modern lower-middle classes. In modern Malay, *mārdeka* has come to mean 'independent' in general.

Various terms existed for different classes of skilled people. De Houtman exhibits the greatest variety in this regard. The default term for a skilled person was *utas*: <*outos* 'subtijl' [professional]> (Ho). Its Javanese equivalent, *tukang*, is only seen in Gueynier: <*toekang* 'ambachts-man, handt-wercker' [craftsman, artisan]>. The word *pandai* more specifically referred to artisans and/or craftsmen working with metal or wood, whereas *juru* specialized in areas which did not involve craftsmanship. In Heurnius' dictionary, *juru* were typically in charge of something: <*jure betá besar* 'hof-meester' [steward]>, <*jure dapor* 'cock of cockinne' [cook]>, and <*jure-tongou* 'opsiender' [overseer]>. In De Houtman, this nuance was expressed with the term <*pongolo* 'meester'>, reflecting Ml. *pəngulu* 'superintendent; headman' (from *hulu* 'head, upper part').

#### 4.3 OTHER AREAS

The lexical material contains some clues to early-modern settlements. The default term for city in modern Indonesian Malay is *kota*, which goes back to Sanskrit *koṭṭa* or *koṭa* 'fort, stronghold'. Early-modern dictionaries show that this word denoted a fortified compound as well as the area within its walls: <*kouta, koutta* 'bolwerck, casteel, fortresse, schantse, slot, sterckte, vesten, wallen' [castle, fortress, fortifications of a city]> (Ho) and <*cotta, cota* 'casteel, vesten van een stadt' [castle, fortifications of a city]> (WD). The same idea is conveyed in

<谷達 *gǔdá* ‘city, city wall’ [城]> (*Siyiguan*). A trading town located near the sea could be designated by the Persian loanword *bandar* ‘city, port, trading town’, hence <*bendar*, *bender* ‘stadt, coopstadt’ [city, trading town]> (Ho), <*orang bendar* ‘burgher’ [citizen; *orang* = ‘person’] (Ho), <*pongolo lyma bendar* ‘stadthoudert’ [governor; *pəngghulu lima* = ‘executive officer’]> (Ho), and modern Malaysian Ml. *bandar* ‘city’. The word *nəgəri* (← Sanskrit *nagari* ‘town, city’) was a more generic term for a settlement. It might have been a type of city-state in early-modern times: <*negri* ‘stadt’ [city]> (vN), <*negry* ‘landt’ [land]> (Ho). In addition, we find the derivations <*nigry citgil* ‘dorp’ [village; *kəcil* = ‘small’]> (vN), <*bendar negry* ‘stadt’ [city]> (Ho), and <*cotta negri* ‘wallen der stadt’ [city walls]> (He).

One of the words in use for smaller settlements was *dusun*, which historically referred to a rural hinterland community (see Mahdi, in this issue). Earlier sources have <路孫 *lùsūn* ‘village’> (*Siyiguan*) and <*dusun* ‘village, village common, community’> (Tanjung Tanah). In seventeenth-century Malay, *dusun* apparently denoted a sparsely populated patch of gardens: <*douson* ‘hof, thuyñ’ [garden]> (WD) and <*douson* ‘landt-huys, vleckē, dorp’ [farm, small village]> (G). The word *kampung* meant ‘gathering together’, as seen in <*campon* ‘t ‘samen komen, vergaderinge’ [come together, gathering]> (WD). It could also refer to a cluster of houses, both in an urbanized area or in the countryside, as in <*campong* ‘straet, dorp’ [street, village]> (Ho) and <*kampong* ‘dorp’ [village]> (G). The word also occurs as <*oerang campong* ‘buerman’ [neighbour; *orang* = ‘person’]> (Ho). In the *Nieuwe Woordenschat*, <*campoong* ‘dorp’ [village]> is listed as a synonym for <*doeson* ‘dorp’ [village]>. The word *desa* (← Sanskrit *deśa* ‘province, country’), denoting a rural village in modern Indonesian, is unattested in the seventeenth-century material, but does already occur in the latter meaning in the fourteenth-century Tanjung Tanah manuscript.

A few observations can be made in the realm of trade. The generic words for money in modern Indonesian, *uang* and *duit* (← Dutch *duit* ‘a coin’), are absent from the seventeenth-century material. Instead, we find other words which historically designated specific coins but temporarily came to designate money in general: <*derham* ‘ghelt, handtgift’ [money, handgift]> (Ho) from the Arabic coin *dirham*, <*dinnar* ‘geldt’> (He) from the Arabic coin *dīnār*, <*pittis* ‘geldt’> (He) from the locally produced *pitis*, and <*pera* ‘geldt’> (He) from Ml. *perak* ‘silver; silver coin’.

We also encounter some information about specific trade products and other manufactured items. The term <*ka’in berlylin* ‘ghewast laken’ [waxed fabric]> (Ho) strikes me as an early mention of batik, whereas <*sedelinggam kambajat* ‘vermillioen dat noch een steen is’ [vermilion still in stone form]> (G) probably denoted cinnabar. The latter compound consists of *sədəlinggam* ‘vermilion’ (← Tamil *cādilingam*) and the toponym *Kəmbayat* (← Gujarati *Khambhāt*). For spectacles, Van Neck gives <*sarou matta* ‘bril’> reflecting *sarung mata* ‘eye cover’ (Collins and Schmidt 1992: 324), whereas the other dictionaries display the modern form *cərmin mata* ‘eye-glasses’: <*tjierement matta* ‘bril’> (Ho) and <*tsjeremyn-matta* ‘bril’> (WD). Ml. *təsmak* ‘spectacles’ (← Pe. *čāsmak*) might have also been in use, but is not attested in the seventeenth-century material.

The seventeenth-century lexicographic material holds some relevant culinary insights, although contemporaneous archival documents, travelogues, diaries, and cookbooks contain more information.<sup>77</sup> Early Dutch lexicographers typically translated Indonesian dishes by referring to the European dishes with which they were familiar. The word for salad, for example, was used to gloss two distinct raw-vegetable dishes: <*olam* 'slaet'> (vN) and <*crabo* 'salade'> (He), that is *ulam* 'uncooked leafy vegetables eaten with rice' and *kərabu* 'cucumber and sour fruits mixed with spices'. Gueynier translates the Malay word for sweet potato as parsnip, <*oubi kastela* 'pasternaken'>, seemingly on account of their superficial similarity. The sweet coconut custard pudding known in Malay as *sərikaya* was translated with the early-modern Dutch word *vlaye*, which denoted a type of flat oven-baked cake with milk and sugar: <*sircaya* 'een vlaye'> (WD).<sup>78</sup> For the Arabic loanword *ḥalwā*, referring to various types of sweetmeats, the word for 'confiture' was chosen; <*haluwa* 'confijt, confitueren'> (Ho). A specific flat round cake made of rice flour, known in modern Malay as *apam*, features as <*apam* 'coeck van rijs' [cake from rice]> (Ho). The word <*djawadah* 'coeck' [cake]> (G) corresponds to modern Ml. *juadah*, which denotes a specific type of sweetmeat, sweetmeats in general, or food for a journey. It reflects Ar. *zuwwāda* or Pe. *zawāda*, both in the generic meaning of 'provisions'. The word *gulai*, which in modern Malay refers to a spicy stew with coconut milk,<sup>79</sup> is simply translated as 'sauce': <*gouley* 'sausse'> (He), <*gouley* 'sausse'> (G). The modern term *rəndang* surfaces not as a dish but as a method of cooking: <*rindang* 'brade inde boter ofte inde oÿ' [sic!], *snercken, fricasseren* [to fry in butter or oil, to fry, to fricassee]> (WD) and <*randang* 'braeden in de panne, fricasseren, fruyten' [to fry in a pan, to fricassee, to sauté]> (G).

#### CONCLUDING REMARKS

Various etymological insights emerge from using and comparing Malay wordlists and dictionaries written during the VOC period. The work of the earliest lexicographers clearly reflects the local characteristics of the region in which they operated, such as North Sumatra in the case of De Houtman and Ambon in the case of Wiltens and Danckaerts. The language of later scholars, particularly Gueynier, had already begun to resemble what would later become standard Malay. And yet, both Heurnius and Gueynier exhibit many words which reveal influence from the Malay dialect of Patani, where several Bible translators spent time. Together, the seventeenth-century sources contain a number of loanwords from Persian and Arabic which were replaced by European equivalents in later times. One might also point out that most of the lexical and morphological characteristics associated with Betawi Malay

<sup>77</sup> See Kathleen Burke (2022) for a study on VOC documents on early-modern Indonesian foodways. Some additional insights might be gleaned from De Houtman's dialogues set in North Sumatra.

<sup>78</sup> This dish is no longer common in the Netherlands, but its name lives on in *vla* 'custard, custard-like filling' and *vlaai* 'a sweet round pie with a filling'.

<sup>79</sup> See Hoogervorst (2017b: 297, 326, n. 6) for a postulated Tamil etymology of this word.

remain unattested in the seventeenth-century material,<sup>80</sup> while the influence of Javanese is relatively modest albeit not absent in this period.

Two logical directions for future research would be to focus on the archival material of Melchior Leydekker – which requires a separate study on account of the sheer quantity of data – and on eighteenth-century Malay lexicography. In both cases, a higher degree of what might anachronistically be called plagiarism is to be expected in the available material. Nevertheless, most lexicographers incorporated substantial quantities of new entries, both expanding and improving those lifted from previous scholarship and excluding erroneous or obsolete data. As such, their publications are useful, albeit best used advisedly. For example, various entries in Bowrey betray his habit of relying on the work of Dutch Bible translators, including <*bōniāco* ‘plunder, spoils’>, <*joo* ‘thou, yee’> (← Ternatan *jou* ‘lord, king’), and <*nen* ‘neither’> (← Pt. *nem* ‘neither’) already seen in Brouwerius’ 1697 translation of Genesis. His <*nardus* ‘spikenard’>, <*tīāra* ‘a mitre’>, and <*ōran hēber* ‘a Hebrew’> also seem to have been taken from early Bible translations. Pertinently, eighteenth-century material contains a number of entries omitted by early Dutch lexicographers yet attested in pre-European sources. Section 2.2 has already drawn attention to the Cantonese loanword for ‘tea’. Another example is a certain plant known in Malay as *pucuk*,<sup>81</sup> which is absent from the Dutch sources but attested as <布竹 *bùzhú* ‘root of a species of Cashmere thistle’> in the *Siyiguan* wordlist, <*poocho* ‘*Costus indicus*’> in Bowrey’s 1701 dictionary, and <*poetjoek* ‘*aloës*’> in the *Nieuwe Woordenschat*. I hope these and other limitations of the present study – including the need for a closer look at the grammar of seventeenth-century Malay – will be taken up in future scholarship.

#### DICTIONARIES USED

Unless indicated otherwise, the dictionaries listed below have been used. I have translated the glosses into English when originally in Indonesian or Dutch and have shortened them where pertinent.

Acehnese	H. Djajadiningrat (1933)
Ambon Malay	D. Takaria and C. Pieter (1998)
Arabic	H. Wehr (1976)
Deli Malay	H. Chalil et al. (1985)
Hindustani	J.T. Platts (1884)

<sup>80</sup> George Meister’s 1692 *Orientalisch-Indianische Kunst- und Lustgärtner* is the first source known to me which displays Betawi influence in the phonological and – to a smaller extent – the lexical realm. The 1780 *Nieuwe Woordenschat* is the earliest source I know to feature the suffix *-in* and several characteristic Betawi words alongside various eastern Indonesian traits.

<sup>81</sup> The word *pucuk* means ‘shoot, leaf bud’ in modern Malay and reflects Proto-Austronesian \**qapucuk* ‘peak of a mountain’ (Blust 1989: 124). It presumably referred to a plant known as Indian costus (*Dolomiaea costus* (Falc.) Kasana and A.K. Pandey).

Javanese	J.F.C. Gericke and T. Roorda (1901)
Langkat Malay	Masindan et al. (1985)
North Sumatran Malay	C. Nasution et al. (2018)
Patani Malay	G. Gray (1962)
Peninsular Malay	R.J. Wilkinson (1932)
Persian	F. Steingass (1892)
Salako	A.K. Adelaar and P.V. Kaslem (2005)
Ternatan	R.A. Atjo (2008)
Tamil	S.V. Pillai and J.S. Chandler (1982)

#### ABBREVIATIONS

Ar.	Arabic
G	F. Gueynier (1677)
He	J. Heurnius (1650)
Ho	F. de Houtman (1603)
ML	Malay
Pe.	Persian
Pt.	Portuguese
R	J. Richardson (1829)
vN	<i>Vocabulaer vande Javaensche ende Malaysche woorden</i> (J. Keuning 1942)
WD	C. Wiltens and S. Danckaerts (1623)

#### REFERENCES

- Adelaar, Alexander and John Hajek. 2024. "Personal pronouns", in: Alexander Adelaar and Antoinette Schapper (eds), *The Oxford guide to the Malayo-Polynesian languages of Southeast Asia*, pp. 907-920. Oxford: Oxford University Press.
- Adelaar, K. Alexander. 1992. *Proto Malayic; The reconstruction of its phonology and parts of its lexicon and morphology*. Canberra: Australian National University.
- Adelaar, K. Alexander. In press. "Austronesian etymologies", in: Philip Durkin (ed.), *Handbook of etymologies*. Oxford: Oxford University Press.
- Adelaar, K. Alexander and P.V. Kaslem. 2005. *Salako or Badameà; Sketch grammar, texts and lexicon of a Kanayatn dialect in West Borneo*. Wiesbaden: Harrassowitz.
- Agius, Dionysius A. 2007. *Classic ships of Islam; From Mesopotamia to the Indian Ocean*. Leiden: Brill.
- Alewyn, A. and Joannes Collé. 1714. *Vocabulario das duas linguas; Portugueza e Flamenga*. Amsterdam: Pieter vanden Berge.
- Atjo, Rusli Andi. 2008. *Kamus Ternate-Indonesia*. Jakarta: Cikoro Trirasuandar.
- Bausani, Alessandro. 1960. "The first Italian-Malay vocabulary by Antonio Pigafetta", *East and West* 11(4): 229-248.



- Blau, Joshua. 2006. *A dictionary of Mediaeval Judaeo-Arabic texts*. Jerusalem: The Academy of the Hebrew Language.
- Blust, Robert. 1989. "Austronesian etymologies: IV", *Oceanic Linguistics* 28(2): 111-180.
- Blust, Robert. 2010. "The greater north Borneo hypothesis", *Oceanic Linguistics* 49(1): 44-118.
- Burke, Kathleen. 2022. *Hearth of empire; Global cuisine in the Dutch Indian Ocean Empire*. PhD thesis, University of Toronto.
- Campbell, Stuart. 1996. "The distribution of -at and -ah endings in Malay loanwords from Arabic", *Bijdragen tot de Taal-, Land- en Volkenkunde* 152(1): 23-44.
- Castro, Ivo, Hugo C. Cardoso, Alan Baxter, Alexander Adelaar, and Gijs Koster. 2022. *Livro de pantuns; Um manuscrito Asiático do Museu Nacional de Arqueologia, Lisboa = Book of pantuns; An Asian manuscript of the National Museum of Archeology, Lisbon*. Lisboa: Imprensa Nacional.
- Chalil, Hayati, Masindan, T. Sita Syaritsa, T. Silvana Sinar, and Oliviana Harahap. 1985. *Kamus Melayu Deli-Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Clercq, Fredrik Sigismund Alexander de. 1876. *Het Maleisch der Molukken*. Batavia: Bataviaasch Genootschap der Kunsten en Wetenschappen.
- Clercq, Frederik Sigismund Alexander de. 1890. *Bijdragen tot de kennis der Residentie Ternate*. Leiden: E.J. Brill.
- Collins, James T. 1992. "Studying seventeenth-century Ambonese Malay; Evidence from F. Caron's sermons (1693)", *Cakalele* 3: 99-122.
- Collins, James T. 1995. "Der Orientalisch-Indianische Kunst-und Lust-Gärtner; An early German contribution to Malay dialect studies", in: Irmengard Rauch and Cornelia Moore (eds), *Across the oceans; Studies from East to West in honor of Richard K. Seymour*, pp. 151-166. Honolulu, HI: University of Hawai'i.
- Collins, James T. 1996. "Of castles and councillors; Questions about Baileu", *Bijdragen tot de Taal-, Land- en Volkenkunde* 152(2): 191-203.
- Collins, James T. 1998. *Malay, world language; A short history*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Collins, James T. 2003. *Mukadimah ilmu etimologi*. Ampang: Dawama.
- Collins, James T. 2004. "A book and a chapter in the history of Malay; Brouwerius' Genesis (1697) and Ambonese Malay", *Archipel* 67(1): 77-127.
- Collins, James T. and Hans Schmidt. 1992. "Bahasa Melayu di Pulau Ternate; Maklumat tahun 1599", *Jurnal Dewan Bahasa* 36 (April): 292-327.
- Dam, Nikolaos van. 2010. "Arabic loanwords in Indonesian revisited", *Bijdragen tot de Taal-, Land- en Volkenkunde* 166(2-3): 218-243.
- Dharma, Po. 1999. *Quatre lexiques malais-cam anciens, rédigés au Campā*. Paris: Presses de l'École française d'Extrême-Orient.
- Dietrich, Albert. 1988. *Dioscurides triumphans; Ein anonymer arabischer Kommentar (Ende 12. Jahrh. n. Chr.) zur Materia medica, pt. 2: Übersetzung und Kommentar*. Göttingen: Vandenhoeck & Ruprecht.

- Djajadiningrat, Hoesein. 1934. *Atjèhsch–Nederlandsch woordenboek*. Batavia: Landsdrukkerij.
- Drewes, G.W.J. 1972. “De invloed van de Atjehse omgeving op het Maleise Spraek ende Woordboek van Frederick de Houtman”, *Bijdragen tot de Taal-, Land- en Volkenkunde* 128(4): 447-457.
- Drewes, Gerardus Willebrordus Joannes and Peter Voorhoeve. 1958. *Adat Atjeh*. The Hague: Martinus Nijhoff.
- Edwards, Evangeline Dora and Charles Otto Blagden. 1931. “A Chinese vocabulary of Malacca Malay words and phrases collected between AD 1403 and 1511 (?)”, *Bulletin of the School of Oriental and African Studies* 6(3): 715-749.
- Flora, Reis Wenger. 1983. *Double-reed aerophones in India to AD 1400*. PhD thesis, University of California.
- Gallop, Annabel Teh. 1990. “Early Malay printing; An introduction to the British Library collections”, *Journal of the Malaysian Branch of the Royal Asiatic Society* 63(258): 85-124.
- Gallop, Annabel Teh. Forthcoming. “Documents from the royal chancery of Aceh; istimi/sitmi in the 17th century”, in: Sher Banu Khan and Maznah Mohamed (eds), *Malay world-making from the 13th to the 21st century; Retelling origins, Islam, Adat and power through Malay manuscripts*. Singapore: NUS Press.
- Gericke, Johann Friedrich Carl and Taco Roorda. 1901. *Javaansch–Nederlandsch handwoordenboek*. Amsterdam: Müller.
- Gil, David. 2024. “How language affects thought; The colour of mosques in Ternate and Padang”. [Paper presented at the 16th International Conference on Austronesian Linguistics, Manila, Philippines, 20-24 June.]
- Gonda, Jan. 1973. *Sanskrit in Indonesia*. New Delhi: International Academy of Indian Culture.
- Gray, Gordon. 1962. *Pattani Malay Dictionary*. Saiburi: O.M.F.
- Gueynier, Fredericus. 1677. *Vocabulaer, ofte Woorden-boeck, nae ordre van den alphabet, in 't Duytsch ende Maleys*. Batavia: Abraham van den Eede.
- Haan, Frederik de. 1912. *Priangan: de Preanger-regentschappen onder het Nederlandsch bestuur tot 1811*. Batavia: Bataviaasch Genootschap van Kunsten en Wetenschappen.
- Heurnius, Justus. 1650. *Vocabularium; ofte Woorden-boeck, nae ordre van den Alphabeth, in't Duytsch en Maleys*. Amsterdam: s.n.
- Hoëvell, G.W.W.C. van. 1876. *Vocabularium van vreemde woorden voorkomende in het Ambonsch-Maleisch*. Dordrecht: Blussé en Van Braam.
- Hoogervorst, Tom G. 2013. *Southeast Asia in the ancient Indian Ocean world*. Oxford: Archaeopress.
- Hoogervorst, Tom G. 2015. “Detecting pre-modern lexical influence from South India in Maritime Southeast Asia”, *Archipel* 89: 63-93.
- Hoogervorst, Tom G. 2017a. “The role of ‘Prakrit’ in Maritime Southeast Asia through 101 etymologies”, in: R. Blench Andrea Acri and Alexandra Landmann (eds), *Spirits and ships; Cultural transfers in early Monsoon Asia*, pp. 375-440. Singapore: ISEAS Yusof Ishak Institute.

- Hoogervorst, Tom G. 2017b. "Lexical influence from North India to Maritime Southeast Asia; Some new directions", *Man in India* 97(1): 293-334.
- Hoogervorst, Tom G. 2018. "Utterance-final particles in Klang Valley Malay", *Wacana, Journal of the Humanities of Indonesia* Vol. 19 No. 2: 291-326.
- Hoogervorst, Tom G. 2021. *Language ungoverned; Indonesia's Chinese print entrepreneurs, 1911-1949*. Ithaca, NY: Cornell University Press.
- Hoogervorst, Tom G. 2023. "Lexical influence from South Asia", in: Francesca R. Moro and M.A.F. Klamer (eds), *Traces of contact in the lexicon; Austronesian and Papuan Studies*, pp. 25-56. Leiden: Brill.
- Hoogervorst, Tom G. 2024. "Non-areal contact", in: Alexander Adelaar and Antoinette Schapper (eds), *The Oxford guide to the Malayo-Polynesian languages of Southeast Asia*, pp. 303-321. Oxford: Oxford University Press.
- Houtman, Frederik de. 1603. *Spraeck ende woord-boeck, in de Maleysche ende Madagaskarsche talen, met vele Arabische ende Turcsche woorden*. Amsterdam: Jan Everts.
- Jacobs, Hubert. 1971. *A treatise on the Moluccas (c. 1544); Probably the preliminary version of António Galvão's lost Historia das Molucas*. Rome: Jesuit Historical Institute.
- Jones, Russell. 2007. *Loan-words in Indonesian and Malay*. Leiden: KITLV Press.
- Jones, Russell. 2009. *Chinese loan-words in Malay and Indonesian*. Kuala Lumpur: University of Malaya Press.
- Kemp, P.H. van der. 1914. "Uit den tijd van C.P.J. Elout's toewijding aan de Maleische taal; Een bijdrage tot de geschiedenis van de beoefening der Indologische wetenschappen", *Bijdragen tot de Taal-, Land- en Volkenkunde* 69(2-3): 141-218.
- Keuning, J. 1942. *De tweede schipvaart der Nederlanders naar Oost-Indië onder Jacob Cornelisz. van Neck en Wybrant Warwijck 1598-1600. III: Het tweede boeck, journael oft dagh-register Amsterdam 1601*. The Hague: Martinus Nijhoff.
- Kozok, Uli. 2015. *A 14th century Malay Code of Laws; The Nītisārasamuccaya*. Singapore: ISEAS.
- Kuiper, Koos. 2005. "The earliest monument of Dutch Sinological studies; Justus Heurnius's manuscript Dutch-Chinese dictionary and Chinese-Latin Compendium Doctrinae Christianae (Batavia 1628)", *Quaerendo* 35(1-2): 109-139.
- Laffan, Michael. 2005. "Finding Java; Muslim nomenclature of insular Southeast Asia from Śrīvijaya to Snouck Hurgronje", *Asia Research Institute Working Paper Series* 52: 1-70.
- Landwehr, John. 1991. *VOC; A bibliography of publications relating to the Dutch East India Company, 1602-1800*. Utrecht: HES.
- Leydekker, Melchior. 1835. *Maleisch en Nederduitsch woordenboek*. Batavia: D. Lenting en J.D.G. Schaap.
- Linehan, W. 1949. "The earliest word-lists and dictionaries of the Malay language", *Journal of the Malayan Branch of the Royal Asiatic Society* 22(147): 183-187.
- Litamahuputty, Betty. 2012. *A description of Ternate Malay*. Utrecht: LOT.

- Lombard, Denys. 1970. *Le 'Spraeck ende woord-boek'; Première méthode de malais parlé (fin du XVIe s.)*. Paris: École française d'Extrême-Orient.
- Lombard, Denys and Luis Filipe F.R. Thomaz. 1981. "Remarques préliminaires sur une lexique portugais-malais inédit de la bibliothèque nationale de Lissebon", in: Nigel Phillips and Khaidir Anwar (eds), *Papers on Indonesian languages and literatures*, pp. 84-96. London: School of Oriental and African Studies.
- Mahdi, Waruno. 1994. "Some Austronesian maverick protoforms with culture-historical implications", *Oceanic Linguistics* 33(1): 167-229; 33(2): 431-490.
- Mahdi, Waruno. 2007. *Malay words and Malay things; Lexical souvenirs from an exotic archipelago in German publications before 1700*. Wiesbaden: Harrassowitz.
- Maier, Henk M.J. and Jan van der Putten. 2002. "Van tolken, papegaaien en predikanten; Het Maleis en de VOC", in: Leonard Blussé and Ilonka Ooms (eds), *Kennis en compagnie; De Verenigde Oost-Indische Compagnie en de moderne wetenschap*, pp. 100-113. Amsterdam: Balans.
- Manguin, Pierre-Yves. 1979. "Note sur l'origine nautique du mot jam", *Archipel* 18: 95-103.
- Marsden, William. 1812. *A dictionary of the Malayan language, in two parts, Malayan and English and English and Malayan*. London: Cox and Baylis.
- Masindan, T. Silvana Sinar, Zulkifli, Muhizar Muchtar, and Oliviana Harahap. 1985. *Kamus Melayu Langkat-Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Minde, Don van. 1997. *Malayu Ambong; Phonology, morphology, syntax*. Leiden: Research School CNWS.
- Nasrallah, Nawal. 2018. *Treasure trove of benefits and variety at the table; A fourteenth-century Egyptian cookbook*. Leiden: Brill.
- Nasution, Chairani, Sriasrianti, Anharuddit Hutasuhut, Zufri Hidayat. 2018. *Kamus Melayu Sumatera Utara-Indonesia*. Medan: Balai Bahasa Sumatera Utara.
- Niemeijer, Hendrik E. and Th. van den End (eds). 2018a. *Bronnen betreffende kerk en school in de gouvernementen Ambon, Ternate en Banda ten tijde van de Verenigde Oost-Indische Compagnie (VOC), 1605-1791; Eerste deel, eerste band: Gouvernement Ambon 1605-1690*. Leiden: Sidestone Press.
- Niemeijer, Hendrik E. and Th. van den End (eds). 2018b. *Bronnen betreffende kerk en school in de gouvernementen Ambon, Ternate en Banda ten tijde van de Verenigde Oost-Indische Compagnie (VOC), 1605-1791. Eerste deel, tweede band: Gouvernement Ambon 1690-1791*. Leiden: Sidestone Press.
- Nieuwe Woordenschat*. 1780. *Nieuwe Woordenschat, uyt het Nederduitsch in het gemeene Maleidsch en Portugeesch, zeer gemakkelyk voor die eerst op Batavia komen*. Batavia: Lodewyck Dominicus.
- Nordhoff, S. 2013. *The genesis of Sri Lanka Malay; A case of extreme language contact*. Leiden: Brill.
- Paauw, Scott H. 2008. *The Malay contact varieties of Eastern Indonesia; A typological comparison*. PhD thesis, State University of New York.

- Petrů, Tomáš. 2016. "'Lands below the winds' as part of the Persian cosmopolis; An inquiry into linguistic and cultural borrowings from the Persianate societies in the Malay world", *Moussons* 27: 147-161.
- Piḷḷai, S. Vaiyāpuri and John Scudder Chandler. 1982. *Tamil lexicon*. Madras: University of Madras.
- Platts, John Thompson. 1884. *A dictionary of Urdū, classical Hindī, and English*. London: W.H. Allen.
- Reid, A.J.S. 1993. *Southeast Asia in the age of commerce, 1450-1680. Vol. 2: Expansion and crisis*. New Haven, CT: Yale University Press.
- Ronkel, Philippus Samuel van. 1896. "Account of six Malay manuscripts of the Cambridge University Library", *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië* 46(1): 1-53.
- Ronkel, Philippus Samuel van. 1903. "De oorsprong van het Maleische woord kamandikai", *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 46: 476-477.
- Sastrawan, Wayan Jarrah. 2020. "The word 'orangutan'", *Bijdragen tot de Taal-, Land- en Volkenkunde* 176(4): 532-41.
- Schuchardt, Hugo. 1891. *Über das malaioportugiesische von Batavia und Tugu*. Vienna: F. Tempsky.
- Shorto, H.L., Paul Sidwell, and Doug Cooper. 2006. *A Mon-Khmer comparative dictionary*. Canberra: Pacific Linguistics.
- Sijs, Noline van der. 2000. *Wie komt daar aan op die olifant?; Een zestiende-eeuws taalgidsje voor Nederland en Indië, inclusief het verhaal van de avontuurlijke gevangenschap van Frederik de Houtman in Indië*. Amsterdam: Veen.
- Smith, Alexander D. 2017. *The languages of Borneo; A comprehensive classification*. PhD thesis, University of Hawai'i.
- Steingass, Franz. 1892. *A comprehensive Persian-English dictionary including the Arabic words and phrases to be met with in Persian literature*. London: Routledge & K. Paul.
- Subrahmanyam, Sanjay. 2005. "From bibliography to history; An introduction", in: Kenneth McPherson and Sanjay Subrahmanyam (eds), *From biography to history; Essays in the history of Portuguese Asia (1500- 1800)*, pp. 9-25. New Delhi: TransBooks.
- Suryadi. 2015. "A linguistic kaleidoscope of the Malay letter", *Wacana, Journal of the Humanities of Indonesia* Vol. 16 No. 2: 304-338.
- Takaria, D. and C. Pieter. 1998. *Kamus bahasa Melayu Ambon-Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Troostenburg de Bruyn, Caspar Adam Laurens van. 1884. *De Hervormde Kerk in Nederlandsch Oost-Indie onder de Oost-Indische Compagnie (1602-1795)*. Arnhem: Tjeenk Willink.
- Vries, Lourens de. 2018. "Iang Evangelivm Ul-Kadus menjurat kapada Marcum; The first Malay Gospel of Mark (1629-1630) and the Agama Kumpeni", *Bijdragen tot de Taal-, Land-en Volkenkunde* 174(1): 47-79.
- Wall, Hermann von de. 1877. *Maleisisch-Nederlandsch woordenboek*. Batavia: Landsdrukkerij.

- Wehr, Hans. 1976. *A dictionary of modern written Arabic*. Ithaca, NY: Spoken Language Services.
- Wieringa, E.P. 2005. "A late eighteenth-century *probatio pennae* from Batavia", *Indonesia and the Malay world* 33(95): 53-65.
- Wilkinson, R.J. 1932. *A Malay-English Dictionary (Romanized)*. Mytilene: Salavopoulos and Kinderlis.
- Wiltens, Caspar and Sebastiaan Danckaerts. 1623. *Vocabularium ofte woort-boeck in 't Duytsch-Maleysch ende Maleysch-Duytsch*. The Hague: H.Jz. van Wouw.
- Winterbottom, Anna. 2016. *Hybrid knowledge in the early East India Company world*. Houndmills / Basingstoke / Hampshire: Palgrave Macmillan.
- Wu, Jiang. 2023. *Malayic varieties of Kelantan and Terengganu; Description and linguistic history*. PhD thesis, Leiden University.