Abstract. This paper offers a synchronic analysis of nineteenth-century Baba Malay ini and itu. On the basis of narrative stories published in the Baba Malay newspaper Bintang Timor, the paper argues that Baba Malay used these forms as both definite articles and as demonstratives; when ini and itu precede a noun, they function as demonstratives, but when they follow a noun, they function as definite articles. The evidence for the distinction between demonstratives and definite articles is found in textual frequencies and discourse functions of ini and itu. Nineteenth-century Baba Malay developed both a proximal and a distal definite article, with the choice between the two definite articles ini and itu reflecting complex discourse considerations.

1. Introduction. The term "Baba Malays" is used to refer to the Straits-born Chinese, who have resided either in Melaka or in Singapore for generations. Chinese settlements in Melaka can be traced back to at least the fourteenth century; Chinese settlements in Singapore began with the founding of the city in 1819. The early Chinese immigrants in Melaka intermarried with the local people, adopting the local Malay culture and the Malay language, but at the same time they preserved some of their Chinese heritage. On the basis of the historical facts documented in the work of historians (Clammer 1983; Freidman 1962, among others) as well as nineteenth-century commentators (Vaughan 1974), it is clear that the first contacts between the Chinese and Malays gave the former full access to Malay culture and language.

The linguistic assimilation of Chinese immigrants in Melaka resulted in the formation of a new variety of Malay, Baba Malay. The Chinese immigrants who settled in Melaka and later in Singapore originally spoke Min dialects—primarily Hokkien, but closely related Teochiu as well. From these dialects they shifted to local varieties of Malay into which they introduced some Hokkien elements (Thurgood 1998).

In the nineteenth century, there must have been a rich variety of speakers of Baba Malay. Baba Malay was the business language used not only by the already established Babas, but also by Malays and the newly arrived Chinese immigrants. In 1887, Lim referred to Baba Malay as the Malay "spoken by all nationalities in the colonies of the Straits Settlements, with the exception of the Europeans" (1887:i). Given such a variety of speakers, it would be unexpected for there to be a single, undifferentiated, unified variety of Baba Malay. The variant of Baba Malay analyzed in this study represents the Baba Malay...
language before it was inundated and dramatically influenced by the arrival of
the great influx of Hokkien and other Chinese speakers at the end of the nine-
teenth century and the beginning of the twentieth century. It is the Baba Malay
that, at that time, had been spoken for generations by the Babas as their first
language.

The paper makes two claims about the systematic variation between the
prehead and posthead positions of the determiners ini and itu in nineteenth-
century Baba Malay. First, the nineteenth-century Baba Malay determiners
function as demonstratives when in prehead position and as definite articles
when in posthead position, that is, there are the two functions distinguished
only by syntactic position. Second, Baba Malay has, not one, but two definite
articles, one developed from the proximal demonstrative and the other from the
distal demonstrative. The evidence for each claim is both qualitative and quan-
titative. The qualitative proof is that Baba Malay articles are found in contexts
that do not allow the use of demonstratives. The quantitative proof is found in
positional frequencies. The frequencies of the demonstrative use correspond to
the frequencies of demonstratives in other languages; the frequencies of the
definite article use correspond to the frequencies of articles. Further, it is shown
that the choice between the proximal article and the distal article is discourse
based, involving the interaction of the animacy hierarchy with a discourse hier-
archy (cf. Chafe 1994).

The emergence of definite articles from demonstratives has also been pos-
ited for a number of other languages, for example, for English (Christophersen
1939); German (Lockwood 1968); Finnish (Laury 1995); and the Romance lan-
guages Vulgar Latin, Portuguese, Spanish, and Rumanian (Williams 1962;
Lathrop 1980; Faingold 1996).1 The descriptions of the development of definite
articles in Swahili (Ashton 1944) and in Montagnais (Cyr 1993) offer particu-
larly useful comparisons for the study of Baba Malay definite articles. Swahili
and Montagnais share with Baba Malay two crosslinguistically rather
rare phenomena. First, in Swahili and Montagnais, the articles and demonstra-
tives are distinguished only by a difference in position. The same claim will
be made for Baba Malay. Second, Montagnais, like Baba Malay, uses the same
form for both, except that the positions are reversed—the postposed determiners
function as definite articles and the preposed determiners function as demon-
stratives.

1.1. Baba Malay data. The analysis of Baba Malay demonstratives is based
on the data that come from the first Baba Malay newspaper Bintang Timor
(published from 2 July 1894 to 2 July 1895). The paper was used to report
local news and news from abroad; it was used to advertise, and to tell stories. It
also included Baba Malay pantun and syair, that is, Malay-style poems, and
dondang sayang, that is, Malay-style pantuns set to music.2

The database for this study consists of four narrative stories published in

Bintang Timor

orfang adek berac
deri hal mak tir
and her stepchild
(FYC) ‘About the
Cherita dua sah:
long). Each of th
variety of Malay
variety of Malay

2. Demonstrative
to language shift:
originated when
of Malay. The st
the two demons
nouns (and class

(1) cit ui siai
this CL gen
‘this gentlemen

(2) hit ui thal
that CL lad:
‘that lady’ (Be

In contrast, te
itu, follow the h
and itu was note
‘they commonly i
before it, as päré
‘that man’” (185
not simply obser
had the possibili
head position. C
on Malay only me
said in 1927 that
words, if any, de
‘any descriptive
or adjectival phr

In Malay, ini
icles. These dual
noted that “the
often, where the

1.1. Baba Malay data. The analysis of Baba Malay demonstratives is based
on the data that come from the first Baba Malay newspaper Bintang Timor
(published from 2 July 1894 to 2 July 1895). The paper was used to report
local news and news from abroad; it was used to advertise, and to tell stories. It
also included Baba Malay pantun and syair, that is, Malay-style poems, and
dondang sayang, that is, Malay-style pantuns set to music.2

The database for this study consists of four narrative stories published in
Bintang Timor between July and October 1894. These stories are Cherita dua orang adek beradek (CDO) 'The story of two siblings' (3,094 words long), Cherita deri hal mak tiri dengan anak tirinya (CDHMT) 'A story about a stepmother and her stepchild' (1,766 words long), Putri yang chinta bapanya sperti garam (PYC) 'About the princess who loved her father like salt' (3,120 words long), and Cherita dua sahabat yang baik (CDS) 'A story of two good friends' (1,674 words long). Each of the stories was published in serial form every other day over a period of two to three weeks. The Baba Malay used in the stories is an informal variety of Malay whose grammatical patterns are not found in a more standard variety of Malay.3

2. Demonstratives and definite articles in the Baba Malay context of language shift. Baba Malay system of demonstratives and definite articles originated when the Chinese immigrants shifted from Hokkien to local varieties of Malay. The shift to Malay required learning a new word order. In Hokkien, the two demonstratives, the proximal cit and the distal hit always precede nouns (and classifiers), as shown in (1) and (2).

(1) cit ui sian-si
    this CL gentleman
    ‘this gentleman’

(2) hit ui thai-thai
    that CL lady
    ‘that lady’ (Bodman 1987:15)

In contrast, the two Malay demonstratives, the proximal ini and the distal itu, follow the head noun. However, a variation in the syntactic position of ini and itu was noted in nineteenth-century Malay by Crawfurd, who reports that “they commonly follow the noun or pronoun, but may also be occasionally placed before it, as pärampüan ini or ini pärampüan ‘this woman’; laki itu or itu laki ‘that man’” (1852:28).4 This observation suggests that, assuming Crawfurd was not simply observing the variation in Baba Malay, nineteenth-century Malay had the possibility of prehead position, in addition to the more common posthead position. Contrary to Crawfurd’s early observations, later commentators on Malay only mention posthead position for ini and itu. Winstedt, for example, said in 1927 that ini and itu follow “the word that they qualify and follow all the words, if any, denoting its attributes” (1957:116). Similarly, Lewis said that “any descriptive words which belong to the noun (i.e., any attributive adjective or adjectival phrase) must come before the ini or itu” (1968:55).

In Malay, ini and itu may function as both demonstratives and definite articles. These dual functions have been commented on in the literature. Crawfurd noted that “the demonstrative pronouns, especially itu, are much used, and often, where they would appear to us, redundant” (1852:28). He suggested that
they seem to be equivalent to our definite article, or to the Latin *ille*, turned into an article . . ." (1852:28). Almost a hundred years later, Winstedt gave a more detailed description of *ini* and *itu* saying that the demonstrative pronoun *itu* means ‘the, that, those’ and refers to “the well-known, the distant in time and space,” and that the demonstrative pronoun *ini* means ‘this, these’ and refers to “the particular, the near in time and space” (1957:116). In Lewis’s description of Malay, both *ini* and *itu* were marked as being “at times the equivalent of the English definite article ‘the,’” when “the thing to which the noun refers has been under discussion, or is familiar to the hearer, or . . . is well known to everybody” (1968:55–56). Here, Lewis observed that, for Malay, it is particularly *itu* that may function as a definite article.

3. The qualitative proof. For a particular language, there is usually agreement as to what a demonstrative is and what a definite article is; occasionally, however, there are problems of definition. These problems arise when demonstratives do not express what is often at least assumed to be their primary meaning—denoting spatial distance (cf. Lyons 1977:646; Fillmore 1982:48; Anderson and Keenan 1985:259; among others). Kirsner (1979:358), for example, has shown that, in Dutch, demonstratives do not have to signal spatial distance at all, but instead they may signal a relative degree of focus placed on a referent by the speaker. As a result, in some contexts, the Dutch proximate demonstrative signals a relatively high degree of attention placed on the referent by the speaker (regardless of the spatial distance), whereas the distal demonstrative signals a relatively low degree of attention placed on the referent. In other contexts, however, these demonstratives do signal spatial distance.

Establishing that a particular demonstrative has actually become a definite article has to be based on criteria that explicitly differentiate between those usages shared by both definite articles and demonstratives and those usages restricted just to definite articles. The criteria used in this study follow those of Hawkins (1978, 1984, 1991), who discusses various different uses of demonstratives and definite articles, and then collapses them into the following four: the anaphoric use, the visible situation use, the larger situation use, and the associative anaphoric use. The first use, the anaphoric use, does not discriminate between demonstratives and articles. The second use, the visible situation use, discriminates between demonstratives and definite articles under certain circumstances. The remaining two, the larger situation use and the associative anaphoric use, are restricted to definite articles; only the posthead occurrences of *ini* and *itu* display these uses. This correlation between the syntactic position and the functions performed by *ini* and *itu* constitute the basis for the claim that, besides demonstratives, Baba Malay also has two definite articles.

3.1. The anaphoric use. Both demonstratives and definite articles display the anaphoric use. In (3) below, the anaphoric use of *ini* is illustrated.

(3) Dahulu ada satu perampuan.6 i woman datang kapad come to-3SG 'In an earlier time the king called

In (3), a new participle the numeral *satu* is referred to as an anaphoric definite Like *ini, itu* and (5).

(4) dan atorkan s) and arrange a.

'and everything

(5) kemudian diai afterwards 3PL 'then they took s

In (4), kemah ‘tent’ precedes the head. rumah ‘house’. In position. A more c

(6) dia tengok api 3SG look fire

ayer sudah water already

tetapi dia ta' but 3SG not

'he saw that the already brough
done all the wor

The anaphoric de events: the fire be the tables and cha events are the refute

In the anaphor position, but also i
In (3), a new participant, raja 'king', is first introduced as satu raja 'a king' with the numeral satu 'one' functioning very much like an indefinite article. Then, he is referred to as raja ini 'the king'. Ini in raja ini marks the phrase as an anaphoric definite NP.

Like ini, itu also marks a particular referent as definite, as shown in (4) and (5).

(4) dan atorkan smoa barang itu didalam satu rumah kemah
and arrange all thing that inside one CL tent
and 'everything was arranged inside a tent' (PYC:265)

(5) kumdian diaorang ambil satu krosi letakkan diluar kemah itu
afterwards 3PL take one chair put outside tent that
'then they took a chair and put it outside the tent' (PYC:272–73)

In (4), kemah 'tent' is marked as a new referent by the classifier phrase that precedes the head noun and consists of the numeral satu 'one' and the classifier rumah 'house'. In (5), kemah 'tent' is marked as definite by itu in the posthead position. A more complicated example of anaphora is shown in (6).

(6) dia tengok api dapor-nya sudah menyala bilek-nya sudah sapu,
3SG look fire kitchen-3SG already light room-3SG already sweep
ayer sudah sedia angkat, meja dan bangku smoa sudah bersh,
water already ready lift table and stool all already clean
tetapi dia ta'tahu siapa yang sudah bikin smoa kerja ini,
but 3SG not.know who that already make all work this
'he saw that the fire in his kitchen was lit, his room was swept, the water was already brought, the tables and chairs were all clean, but he did not know who had done all the work,' (CDHMT:178–81)

The anaphoric definite NP smoa kerja ini 'all the work' refers back to four events: the fire being lit; the room being swept; the water being brought; and, the tables and chairs being cleaned. The anaphoric use of ini signals that these events are the referents of the NP smoa kerja ini 'all the work'.

In the anaphoric use, Baba Malay ini and itu occur not only in the posthead position, but also in the prehead position. Compare (4)–(6) above with (7) below.
(7) dan dia dapat satu mimpi dia tengok ads satu hutan, dalam itu
and 3SG get one dream 3SG look be one jungle inside that
hutan ads satu orang bertapa, ini orang bertapa jikalau dia tidor
jungle be one person hermit this person hermit if 3SG sleep
dua bias tahun
two teen year
′and he had a dream in which he seemed to be in a jungle, in the jungle there was a
hermit, the hermit would sleep twelve years′ (PYC:162–64)

Examples (4)–(7) illustrate the anaphoric use of ini and itu characteristic of
definite articles and demonstratives.

3.2. The immediate situation use. Hawkins (1978:110–15) says that, with a
demonstrative, it is essential for the hearer to be able to see the referent; but,
with a definite article, the hearer does not have to be able to see the referent. He
observes that it is not the function of definite articles to signal visibility "even
when the referent is to be located in the immediate situation of utterance in
which it is potentially visible" (1978:114). Thus, in (8), the hearer is warned
about a dog in the immediate vicinity, but it is not necessarily visible to him.

(8) Don′t go there, chum. The dog will bite you. (Hawkins 1978:112)
In contrast, a demonstrative in the same sentence would instruct the hearer “to
identify the object itself, and thus it actually has a visibility requirement built
into it as part of its meaning” (Hawkins 1978:115).

The four Baba Malay texts examined in this study have been analyzed in
light of Hawkins′s observation. Although ini and itu do not often occur in the
prehead position, when they do and when this is a first mention, the referent is
always visible. For example, it is clear from the story PYC that, in (9), the
interlocutors can see the jungle that is referred to with the NP ini hutan ‘this
jungle′. Similarly, it is clear from the story CDO that, in (10), the interlocutors
can see the water referred to with the NP itu ayer ‘that water′.

(9) mengapa angkau datang didalam ini hutan,
why 2SG come inside this jungle
′Why did you come into this jungle?′ (PYC:198)

(10) jikalau angkau mandi badan bulih dapat kuat dan segar,
if 2SG bathe body can get strong and refreshed
lekas-lah itu ayer nanti sejok
fast-EMPH that water later cold
′′if you wash, you will feel strong and be refreshed, quickly before that water gets
cold,′′′ (CDO:263–64)

However, when their referents are visible, as, for example, red to as raja ini

(11) angkau duda
2SG sit
′please sit an

But the crucial is the hearer is in a

(12) Maka bertrus
then cry.out
selak klami
open mosqi
′Then the ol
(CDO:276–77)

Thus, in (12), itu such cases, ini an
demonstratives i

3.3. Larger situ
noted that there
amely, what be
use." The larger
arious kinds, e.
objects" (Hawkins
lish town has a to
mentioned for th

(13) Halifax is a s
last year. (H

Hawkins (1978:1
ear to the defini
However, when ini and itu occur in posthead position, there are times when their referents are clearly not visible. Sometimes, of course, the referents are visible, as, for example, in (11), where both interlocutors can see the king referred to as raja ini ‘the king’.

(11) angkau dudok tunggu raja ini,
    2SG sit watch.over king this
    ‘please sit and watch over the king,’ (PYC:93)

But the crucial examples are when the referents are not visible, as in (12), where the hearer is in another room and thus cannot see the mosquito net referred to with the posthead determiner itu.

(12) Maka bertriak-lah baboo tua itu, dengan berkata, “Hai, jangan-lah
    then cry.out-EMPH domestic old that with say hey don’t-EMPH
    selak klambu itu.”
    open mosquito.net that
    ‘Then the old nurse called out, saying, “Hey, don’t open the mosquito net,”’
    (CDO:276–77)

Thus, in (12), itu must be an article since the referent is clearly not visible. In such cases, ini and itu pattern like definite articles in posthead position, but like demonstratives in prehead position.

3.3. Larger situation and associative anaphoric uses. Hawkins (1978) noted that there are uses of articles that are not shared by demonstratives, namely, what he terms “the larger situation use” and “the associative anaphoric use.” The larger situation use requires “a general knowledge that situations of various kinds, e.g., weddings, villages, countries, etc., generally contain certain objects” (Hawkins 1978:119). For example, the general knowledge that an English town has a town clerk makes it possible to refer to the town clerk even when mentioned for the first time, as in (13).

(13) Halifax is a sleepy little Yorkshire town. The town clerk was involved in a scandal last year. (Hawkins 1978:120)

Hawkins (1978) shows that referring to general knowledge is also the basis for the associative anaphoric use of the definite article. This use of definite articles requires a linguistic referent which triggers relevant associations, as in (14).

(14) The man drove past our house in a car. The exhaust fumes were terrible. (Hawkins 1978:123)

Hawkins (1978:123) observes that the first NP a car triggers associations that lead to the definiteness of the exhaust fumes. Hawkins notes that, although it is
not clear what the parameters defining the set of possible associates are, "the notion 'part-of' seems to play and important role in defining ... possible associates" (1978:123). He observes that "the trigger must conjure up a set of objects which are generally known to be part of some larger object or situation" (1978:123–24).

This associative use of Baba Malay itu is illustrated in (15) and (16) below.

(15) Tetakala sudah sampai ka-istana-nya, dia suroh orang orang-nya
when already arrive to-palace-3SG 3SG order person person-3SG
riaskan dalam istana itu dengan segala perhiasan
decorate inside palace that with all decoration
'When they arrived at his palace, he ordered his/the people to decorate the inside of the palace' (CDO:226–27)

(16) Maka sahari hari dia bermain main-lah di dalam taman itu,
after that one.day day 3SG play play-EMPH at inside garden that
'Then every day he/she/it played in the garden,' (CDO:234)

In (16), this associative, first-mention use of taman 'garden' occurs with the determiner itu already in posthead position, suggesting that it is an article. The larger context provided by the story makes it clear that the definiteness of taman 'garden' has been established by associating it with istana 'palace' referred to in (15).

In a similar way, but in a different story, the earlier mention of tempat tidur 'the bed' in (17) allows the narrator to use a posthead itu when referring to a nearby corner penjuru itu 'the corner' in (18).

(17) angkau mesti masok tempat tidor,
you must enter place sleep
'you must go to bed,' (CDHMT:19)

(18) baik juga dia ada berdiri ditepi penjuru itu,
good again 3SG be stand at.edge corner that
'it was good that she was standing in the corner,' (CDHMT:24)

Reference to general knowledge is also illustrated in (19), where the determiner ini is used to refer to kebun 'garden' on the first mention.

(19) bagaimana elok nampaknya rumah ini, dan brapa chantek
how handsome see house this and how.much beautiful
kebun ini dengan kolam-nya,
garden this with pool-3SG
'how handsome the house looked, and how beautiful the garden with its pool was,' (PYC:61–62)
Examples (16)–(19) illustrate a very frequent use of Baba Malay ini and itu, one in which the two determiners are used to mark a particular NP as being part of some larger object. In such uses, ini and itu always occur in the posthead position, suggesting that posthead position is associated with definite articles.

Baba Malay determiners in posthead position are also used to refer to other first-mention objects that are neither visible to interlocutors nor anaphorically related to a previously introduced referent. Nonetheless, it can be shown that in such uses ini and itu function as definite articles, as, for example, in (20).

(20) tengok mata-hari itu chahyanya merah lagi, tetapi dia sudah turun
     see eye-day that glow-3SG red more but 3SG already come
diseblah kanan,
     down.at.side right
‘did you see the sun glowing red, but it has already set on the right side,’
(CDS:9–11)

In (20), mata-hari ‘sun’ is mentioned for the first time in the story. Although not visible to the interlocutors, and not anaphorically related to any previously introduced referent, mata-hari ‘sun’ is referred to with itu in posthead position. Obviously, this use results from shared general knowledge about the sun that makes it possible to refer to it with a definite article even on the first mention.

The general knowledge that a Malay state has a governor called a dato (in standard Malay datuk) makes it possible for the narrator of the story Di cheritakan dengan rengkasnya waktu kawin-nya Tuan putri anak-anda Beginda Sultan Johore (DDR) ‘A short report of the wedding of the daughter of the Sultan of Johore’ to refer to the dato with the definite article on the first mention, as in (21).

(21) Dato ini membri satu mata-matanya menyuroh sahya pergi
dato this give one eye-eye-3SG order 1SG go
di rumahnya satu towkay China
to house-3SG one towkay China
‘The Dato gave one of his constables an order that I was to go to the house of a Chinese.’ (DDR:18)

Similarly, the raja ‘king’ in (22) is referred to as raja . . . itu ‘the king’ when mentioned for the first time.

(22) Pada suatu hari raja didalam negri itu datang memburu,
on one day king at.inside country that come hunt
‘One day the king of that country came to hunt,’ (CDO:38)

In (22), however, there is the prepositional phrase didalam negri ‘in the country’ intervening between raja and the determiner itu. This leads to scope ambiguity as in Baba Malay it is not possible to say *raja didalam negri itu itu
3.4. Concluding remarks. Qualitative proof for the existence of the Baba Malay posthead definite articles ini and itu is found in their four basic uses. In each case, when a use differentiates between demonstratives and articles, the demonstrative use is associated with prehead position and the article use is associated with posthead position. Thus, demonstratives and articles are readily differentiated on the basis of these tests and on the basis of their syntactic position.

4. The quantitative proof. The quantitative proof for the existence of Baba Malay definite articles is found in positional frequencies. The numbers given in table 1 show the frequencies of ini and itu per full NP in the four Baba Malay stories analyzed for this study. The frequency of ini and itu is far, far higher in posthead position than in prehead position, as might be expected. In the four of the Baba Malay stories analyzed here, posthead ini and itu are used with between 32.9 percent and 40.2 percent of the NPs. In contrast, in none of the stories does the frequency of prehead ini and itu exceed 6.8 percent. These textual frequencies by themselves suggest that posthead ini and itu are articles and prehead ini and itu are demonstratives.

Table 1. Textual Frequencies of ini and itu in Two Syntactic Positions

<table>
<thead>
<tr>
<th>TEXT</th>
<th>NP + ini/itu</th>
<th>ini/itu + NP</th>
<th>TOTAL NPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDO</td>
<td>38.4% (144)</td>
<td>3.5% (13)</td>
<td>100% (375)</td>
</tr>
<tr>
<td>CDHMT</td>
<td>39.5% (75)</td>
<td>3.7% (7)</td>
<td>100% (190)</td>
</tr>
<tr>
<td>PYC</td>
<td>40.2% (159)</td>
<td>6.8% (27)</td>
<td>100% (395)</td>
</tr>
<tr>
<td>CDS</td>
<td>32.9% (84)</td>
<td>4.7% (12)</td>
<td>100% (255)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>38.0% (462)</td>
<td>4.9% (59)</td>
<td>100% (1,215)</td>
</tr>
</tbody>
</table>

The Baba Malay textual frequencies correlate with the frequencies found in other languages. In her study, Cyr (1993:221) compares the textual frequencies of definite articles with the frequencies of demonstratives in such languages as French, Italian, Cree, Swedish, Montagnais, and German. Her statistics are based on narrative texts, short stories, and condensed versions of larger tales (Cyr 1993:222). Cyr’s frequencies for definite articles and demonstratives are presented in table 2 and, for comparison, the textual frequencies of Baba Malay ini and itu have been added.

The posthead frequencies of Baba Malay ini and itu correspond to the frequencies in other languages for definite articles; and, the prehead frequencies of ini and itu correspond to the frequencies of demonstratives in other languages. These numbers provide additional support for the claim that, in posthead position, Baba Malay ini and itu function as definite articles, but, in prehead posi-
In Baba Malay, the NPs with ini and itu represent less than 50 percent of all NPs in the data, leaving the majority of Baba Malay NPs marked in another way or unmarked. Some Baba Malay NPs are marked as definite, not by ini or itu, but by free or bound pronominal forms. As shown in table 3, the most common bound form is the enclitic -nya, which marks primarily the third person as the possessor, and, of course, as definite.

Table 3. Frequencies of NPs with -nya and with Pronouns in Possessive Constructions

<table>
<thead>
<tr>
<th></th>
<th>NP-nya</th>
<th>NP + PRONOUN</th>
<th>TOTAL NPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEXT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PYC6</td>
<td>11.6%</td>
<td>3.3%</td>
<td>100% (375)</td>
</tr>
<tr>
<td>CDO</td>
<td>31.5%</td>
<td>4.0%</td>
<td>100% (395)</td>
</tr>
<tr>
<td>CDHMT</td>
<td>33.3%</td>
<td>1.6%</td>
<td>100% (190)</td>
</tr>
</tbody>
</table>

If they refer to characters of primary or secondary importance (Chafe 1994), first-mention referential indefinite NPs occur with classifiers. The textual frequencies are given in table 4.

Table 4. Textual Frequencies of Indefinite NPs with Classifier Constructions

<table>
<thead>
<tr>
<th>TEXT</th>
<th>CLASSIFIER + NP</th>
<th>TOTAL NPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDO</td>
<td>6.1% (23)</td>
<td>100% (375)</td>
</tr>
<tr>
<td>PYC</td>
<td>9.9% (39)</td>
<td>100% (395)</td>
</tr>
<tr>
<td>CDHMT</td>
<td>10.5% (20)</td>
<td>100% (190)</td>
</tr>
</tbody>
</table>

Nouns that refer to props and characters of trivial importance, that is, characters that are “sources of information, but trivial in the discourse” (Chafe 1994:90) occur unmarked, not just on their first mention, but also on the
subsequent mentions, a pattern that Baba Malay shares with Malay (Cumming 1991:96). The textual frequencies of unmarked nouns are given in table 5.

Table 5. Textual Frequencies of Unmarked Nouns

<table>
<thead>
<tr>
<th>Text</th>
<th>Unmarked NPs</th>
<th>Total NPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDO</td>
<td>16.5% (62)</td>
<td>100% (375)</td>
</tr>
<tr>
<td>PYC</td>
<td>15.1% (60)</td>
<td>100% (395)</td>
</tr>
<tr>
<td>CDHMT</td>
<td>10.5% (20)</td>
<td>100% (190)</td>
</tr>
<tr>
<td>CDS</td>
<td>8.6% (22)</td>
<td>100% (255)</td>
</tr>
<tr>
<td>Total</td>
<td>13.5% (164)</td>
<td>100% (1,215)</td>
</tr>
</tbody>
</table>

Thus, in their article use, ini and itu refer to primary and secondary characters; references to props and trivial characters remain unmarked. The question that remains to be answered is what Baba Malay does with the two definite articles, ini and itu.

5. The two definite articles in Baba Malay. Baba Malay has two definite articles—one developed from the proximal demonstrative and the other from the distal demonstrative. As shown in table 6, both of the articles are used with roughly equal frequency. The question is what determines which is used where.

Table 6. The Distribution of the Two Definite Articles

<table>
<thead>
<tr>
<th>Text</th>
<th>NP + ini</th>
<th>NP + itu</th>
<th>Total NPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDS</td>
<td>57% (48)</td>
<td>43% (36)</td>
<td>100% (84)</td>
</tr>
<tr>
<td>CDO</td>
<td>52% (75)</td>
<td>48% (69)</td>
<td>100% (144)</td>
</tr>
<tr>
<td>CDHMT</td>
<td>44% (33)</td>
<td>56% (42)</td>
<td>100% (75)</td>
</tr>
<tr>
<td>PYC</td>
<td>39% (62)</td>
<td>61% (37)</td>
<td>100% (159)</td>
</tr>
</tbody>
</table>

In the following sections, it is shown that the choice between the two definite articles reflects the complex interaction of two different tendencies. The crucial interplay is between Kuno's (1976) notion of "empathy," that is, in part, the tendency to empathize with the subjects, and Chafe's (1994) notion of referential importance.

5.1. Narrator's empathy. Kuno suggests that grammatical choices can be influenced by "the speaker's attitude towards the participants of the event that he is describing" (1976:431). He uses the term "empathy" to "characterize the speaker's identification, in varying degrees, with a participant in an event" (Kuno 1976:431). Kuno describes the speaker's empathy hierarchy as follows:

Surface Structure Emptaphy Hierarchy (revised): It is easiest for the speaker to empathize with the referent of the subject than with the referents of other NPs in the sentence. [Kuno 1987:211]
This hierarchy of "empathy" can be seen in the Baba Malay distribution of the proximal article and the distal article. Throughout the Baba Malay stories, the narrators use the proximal definite article ini to mark high empathy with the subject. The referents of the other NPs, in contrast, are usually marked as distal by itu. That is, typically, the speaker is "closer" to the subject than to the other NPs. Compare (23) and (24) below.

(23) maka raja ini mulai berchakap kepada anak perawan itu
then king this begin speak to child woman that
'then the king began speaking to the girl' (CDO:217)

(24) dan itek ini bernang-lah kolam itu.
and duck this swim-EMPH pool that
'and the duck swam on the pond.' (CDHMT:103)

In (23), the subject raja 'king' occurs with the proximal definite article ini, but anak perawan 'female child' occurs with the distal definite article itu. Likewise, in (24) from a different story, the two NPs are grammatically marked as definite in different ways depending on whether the particular NP is the subject or not. Examples (23)–(24) illustrate a very common pattern found throughout the Baba Malay narratives: referents of subjects are marked with the proximal definite article ini; referents of other NPs are marked with the distal definite article itu.

Table 7 presents the statistical frequencies in the use of ini with subjects versus nonsubjects in three Baba Malay stories. In all three of the Baba Malay stories, the use of ini with subjects exceeds 70.6 percent, while the use of ini with nonsubjects is under 30 percent. However, as shown in table 8, the situation is reversed with itu, which occurs predominately with nonsubjects.

<table>
<thead>
<tr>
<th>Text</th>
<th>Subjects</th>
<th>Nonsubjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDHMT</td>
<td>87.5% (28)</td>
<td>12.5% (4)</td>
</tr>
<tr>
<td>PYC</td>
<td>78.5% (51)</td>
<td>21.5% (14)</td>
</tr>
<tr>
<td>CDO</td>
<td>70.7% (53)</td>
<td>29.3% (12)</td>
</tr>
</tbody>
</table>

Table 8. The Use of itu with Subjects vs. Nonsubjects

<table>
<thead>
<tr>
<th>Text</th>
<th>Subjects</th>
<th>Nonsubjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDHMT</td>
<td>22.0% (9)</td>
<td>78.0% (32)</td>
</tr>
<tr>
<td>PYC</td>
<td>24.7% (23)</td>
<td>75.3% (70)</td>
</tr>
<tr>
<td>CDO</td>
<td>32.4% (22)</td>
<td>67.6% (46)</td>
</tr>
</tbody>
</table>

The use of itu with subjects never passes 32.4 percent, while with nonsubjects it occurs over 67 percent of the time. Clearly, the occurrence of ini and itu also correlates with the distinction between subjects and nonsubjects, and
reflects what Kuno described as a tendency to empathize with the subject. Although this tendency usually completes the animacy hierarchy, at times the two conflict.

5.2. The animacy hierarchy. The animacy hierarchy can be represented as the distinction between human, nonhuman animate, and inanimate categories in language. In the animacy hierarchy, humans are ranked above animals, and animals are ranked above objects (Comrie 1989):

human > nonhuman animate > inanimate

The animacy hierarchy reflects the human point of view of the world: of primary importance to humans are other humans, of secondary importance to humans are animals, and of tertiary importance are objects. This hierarchy is sometimes reflected in grammatical choices and sometimes in discourse choices.

In Baba Malay, the animacy hierarchy is reflected in the use of the two definite articles ini and itu. Example (25) illustrates a case in which the grammatical categories do not coincide with the animacy hierarchy.

(25) dan pintu itu pun terbuka lalu rusa ini masok dan pintu itu tertutup kembali.

'and the door opened, the deer entered, and the door closed again.' (CDO: 152–53)

Here, the first and the third clauses have the inanimate noun pintu 'door' as their subjects; the second clause has the animate noun rusa 'deer' as its subject. Notice that it is only the animate subject rusa 'deer' that is marked by ini; the inanimate pintu, even though in the subject position, is marked by itu.

In (23)–(24), the grammatical category predicts the choice between the proximal article ini and the distal article itu. Example (25) shows that the grammatical category may not determine the choice if it conflicts with the animacy hierarchy. Note that the animacy hierarchy, even more directly than the grammatical category, reflects Kuno’s (1976) empathy. In (25), where the animacy considerations override the grammatical category, it is the most empathized referent that is marked with ini.

The statistical frequencies presented in table 6 show that the definite article ini is used primarily to refer to human referents. In all three Baba Malay stories, ini is used with human referents more than 84.5 percent of the time, while the use of ini never exceeds 15.4 percent.
with the subject. At times, however, it can be represented as inanimate categories above animals, and

Table 9. The Use of ini with Human Referents

<table>
<thead>
<tr>
<th>TEXT</th>
<th>HUMAN</th>
<th>NONHUMAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDO</td>
<td>100% (75)</td>
<td>0% (0)</td>
</tr>
<tr>
<td>CDHMT</td>
<td>87.5% (28)</td>
<td>12.5% (4)</td>
</tr>
<tr>
<td>PYC</td>
<td>84.6% (55)</td>
<td>15.4% (10)</td>
</tr>
</tbody>
</table>

Unlike ini, the use of itu correlates neither with human referents nor with nonhuman referents, as shown by the statistical frequencies presented in Table 10.

Table 10. The Use of itu with Human and Nonhuman Referents

<table>
<thead>
<tr>
<th>TEXT</th>
<th>HUMAN</th>
<th>NONHUMAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDO</td>
<td>50.0% (34)</td>
<td>50.0% (34)</td>
</tr>
<tr>
<td>CDHMT</td>
<td>41.5% (17)</td>
<td>59.0% (24)</td>
</tr>
<tr>
<td>PYC</td>
<td>50.5% (47)</td>
<td>49.5% (46)</td>
</tr>
</tbody>
</table>

5.3. The influence of the discourse hierarchy. The choice between the two definite articles also interacts with a discourse hierarchy reflected in Chafe's (1994) notion of referential importance. Chafe (1994) distinguishes between characters of primary importance, secondary importance, and trivial importance, with certain props, of course, falling outside of the classification. For example, consider (26) below.

(26) dia-orang berdua itu angkat-lah permesuri ini
3SG-person both that lift-EMPH queen this
'they both lifted the queen' (CDO:265)

The expression dia-orang berdua 'both of them' refers to the stepmother and her child who are characters of secondary importance. As a consequence, the expression dia-orang berdua 'both of them' co-occurs with the distal itu, even though it is in a subject position. In contrast, the noun permesuri 'queen' co-occurs with the proximal article ini, because the queen is one of the two major characters in the story and, thus, of primary importance. Similarly, in (27), taken from another story, the raja 'king' is followed by ini, even though it is in a nonsubject position, because the king is more important than the orang bertapa 'hermit'.

(27) orang bertapa itu banyaklah suka dengan raja ini,
person hermit that many-EMPH like with king this
'the hermit liked the king very much,' (PYC:196)

In (26)–(27), the influence of the grammatical category was overridden, whereas in (28) below, the effects of both the grammatical category and the animacy hierarchy are overridden.
"Then the king saw the deer come,' (CDO:191)

Here, the raja is marked by the distal article itu, while rusa 'deer' is marked by the proximal article ini. The ignoring of the grammatical category and of the animacy hierarchy is imposed by the discourse importance of the rusa 'the deer', which is of more importance in the story than the raja.

6. Conclusion. This study has examined ini and itu in the literary Baba Malay of the nineteenth century. Using narrative stories published in the Baba Malay newspaper Bintang Timor as a database, the paper has argued that ini and itu function both as definite articles and as demonstratives. When ini and itu precede nouns, they function as demonstratives; when they follow nouns, they function as definite articles. Further, the Baba Malay system of definite articles is particularly interesting because it distinguishes between a proximal definite article ini and a distal definite article itu, a choice determined by the interaction of grammatical and discourse considerations.

Baba Malay developed its system of articles in the course of language shift, when Hokkien speakers shifted to Malay. The findings of this study suggest that the system of nineteenth-century Baba Malay determiners reflects a compromise between the distributional patterns found in Malay and in Hokkien. The statistically more common Baba Malay pattern head + article comes from Malay, while the statistically less common pattern demonstrative + head comes from Hokkien, which itself has demonstratives in that position.

The Malay-like Baba language analyzed in this study was later influenced by two events: a large influx of new Chinese immigrants, and the growth of Singapore and other British controlled towns. From a linguistic point of view, the first event brought an enormous number of new speakers of not just Hokkien, but of other Chinese languages as well. The second event resulted both in Baba Malay, not the Malay of the Malays, becoming the dominant language of the Straits Settlements, and in English, not Malay, being the most important second language of the region. Thus, it would be interesting to examine modern Baba Malay texts to see in what ways its determiner system resembles the system of the nineteenth century and in what ways it has changed.

Notes

Acknowledgements. I would like to thank Mohamed Risham Fauzi and Solakhia Januari for assisting with text translation. I am indebted to Graham Thurgood, George Grace, and Roderick Jacobs for helpful comments and discussions. I am grateful to Barbara Andaya for generous assistance with the historical aspects of this paper, and to Benjamin Te’ou for his information about the Min dialects. Thanks are also due to Hein Steinhauser and to an anonymous referee for their thoughtful comments. All errors, of course, remain mine.

Abbreviations

person; 3 = third singular.

1. Frajzyngier demonstrates that definite markers d example, verbs of s
tells how when he Baba from Malac
ing in the comp 489).

2. Although be widespread the syn
calculative and tomonts' (19
telling how when he Baba from Malac
ing in the comp 489).

3. Although be widespread the syn
calculative and tomonts' (19
telling how when he Baba from Malac
ing in the comp 489).

4. The Malay s
telling how when he Baba from Malac
ing in the comp 489).

5. The Baba M
telling how when he Baba from Malac
ing in the comp 489).

6. Not included
telling how when he Baba from Malac
ing in the comp 489).

7. See, however, modern Baba Mal

Anderson, Stephen 1985 Deixi
daic Camb

Ashton, E. O. 1944 Swah

Bodman, Nicholas 1987 Spoke

Chafe, Wallace 1994 Disco

Christophersen, P 1939 The A

Clammer, John 1983 The S Capit

Collins, J. T. 1980 Ambo
dan P
Abbreviations. The abbreviations used in the glosses are as follows: 2 = second person; 3 = third person; CL = classifier; EMPH = emphatic particle; FL = plural; SG = singular.

1. Frajzyngier (1996) challenges the assumption that definite articles develop from demonstratives (cf. Greenberg 1978). On the basis of Chadic languages, he argues that definite markers do not have to derive from deictic markers, but may originate from, for example, verbs of saying (see Frajzyngier [1996] for discussion).

2. Vaughan says that “the Malacca Babas are exceedingly musical and are very clever in extemporizing words to their tunes, and will for several hours at a time amuse themselves and their guests by singing their pantuns, and lagus, accompanied by fiddles and tom-toms” (1974:39). Vaughan’s reports are confirmed by Munshi Abdullah, “who tells how when he was accompanying his master from Johor on a voyage in 1871, four Baba from Malacca were brought in to entertain the company in the evening by competing in the competition of pantun and songs with an old blind Malay” (Salmon 1987: 489).

3. Although beyond the scope of this work, it might be interesting to investigate how widespread the syntactic patterns found in Baba Malay have been among other regional varieties of Malay, which include Kelantan Peranakan Hokkien (Teo 1993), Chindo (Rafferty 1982), and Ambones Malay (Collins 1980; Dix Grimes 1980, 1994).

4. The Malay spelling is the original orthography used by Crawfurd (1852).

5. The Baba Malay spelling is the original orthography used in Bintang Timor.

6. Not included here are the fifty occurrences of the term si Putri used to refer to the main character of the story, the princess. The term si Putri in this Baba Malay story is used very much like in Malay (cf. Marsden 1973:156), with “a titular prefix” si (Wilkinson 1969:1101) marking a title-like use of the noun putri ‘princess’.

7. See, however, Pakir (1986) and Lim (1988) for their analyses of different aspects of modern Baba Malay.

References


Comrie, Bernard

Crawfurd, John

Cumming, Susanna

Cyr, Danielle

Dix Grimes, Barbara


Faingold, Eduardo D.

Fillmore, Charles J.

Frajzyngier, Zygmunt

Freedman, Maurice

Greenberg, Joseph H.

Hawkins, John A.


Kirsner, Robert S.

Kuno, Susumu
1976 Subject, Theme, and the Speaker’s Empathy—a Reexamination of
Williams, E. B.  

Winstedt, Richard O.  