Having come to be a copula in Sri Lanka Malay – an unusual grammaticalization path

Abstract

Sri Lanka Malay has innovated a prominent and productive copula, which sets it apart from other descendents of colloquial Malay varieties. This copula has developed from the verb *dhaathang* 'to come', which is a grammaticalization path not yet attested in the literature. This paper describes the forms and functions of this copula and shows that it cannot be traced to any of the main input languages of Sri Lanka Malay (Trade Malay, Tamil, Sinhala). Comparing the Sri Lanka Malay case to attested grammaticalization paths, this paper concludes that the grammaticalization of COME to a copula is less surprising when assuming intermediate stages of 'resultative' and 'stative'. These subpaths are illustrated by a variety of Creole and non-Creole languages.

1 Introduction[2]

Languages with copulas are widely known in Europe and beyond, but languages without a copula are also frequently encountered. If languages without a copula develop one over time, the source is normally either a former locational verb or a deictic (Stassen 1997:91). In this paper, I want to highlight yet another source for a copula, namely the participle of the verb COME, which has grammaticalized into a copula in Sri Lanka Malay.

I will first give some information on Sri Lanka Malay (Section 2) and the types of copula sentences we find in that language (Section 3). I will then show that neither other varieties of Malay nor the contact languages Sinhala and Tamil provide models for the development of this particular copula in Sri Lanka Malay (Section 4). I then discuss

---

common sources for the grammaticalization of copulas in Section 5 and common targets of grammaticalization for the verb COME in Section 6, showing that the particular change COME→copula is not attested. However, three subpaths, namely COME→resultative, resultative→stative and stative→copula are attested, which, when chained, can explain the Sri Lanka Malay facts.

2 Sri Lanka Malay

Sri Lanka Malay is the language of the descendents of soldiers, exiles, convicts and slaves who were brought to the island by the colonial powers of the Dutch and the British between roughly 1650 and 1850 (Hussainmiya 1990). The grammar of the language has changed dramatically ever since (Adelaar 1991, Smith et al. 2004, Ansaldo 2008, Nordhoff 2009) and has shed most of its Austronesian origins to align with the Sri Lankan sprachbund (Bakker 2006), although some traces of Malay grammar still remain (Slomanson 2006). For instance, the language has changed from SVO to SOV and from prepositional to postpositional. The language has developed more than 20 new bound morphemes, mainly for the marking of TAM and case – something which is virtually unheard of in other parts of the Malay world. Most of the changes the language has undergone since its arrival on the island 350 years ago were triggered by language contact. The copula, however, does not seem to have a clear origin in any contact language or any historical variety of Malay.

3 The copula in Sri Lanka Malay

The copula in Sri Lanka Malay is homophonous to the conjunctive participle of the verb *dhaathang* 'come'. The Sri Lanka Malay verb can take a number of prefixes, one of them being the conjunctive participle. This form is used to indicate that the event depicted in the clause thus marked is prior to the event in the main clause. In this sense, it is similar to an English construction like *Having done X, having done Y, finally Z happened*. An illustration of this form is given in (1).

(1) a.   Oorang pada **asà-**pìrrang,
      man   PL   CP-wage.war
      'After having waged war'
  b.   derang=nang **asà-**banthu,
      3PL=DAT   CP-help
      'and after having helped them'

c. siini=jo   su-ciiñğgal.
   here=EMPH PAST-settle
   'the people settled down right here.' (K051222nar03)

Alternatives to the form with *asà-* shown in (1) are the form with *=apa* shown in (2) and the form with both *asà-* and *=apa*, shown in (3). These three forms seem to be functionally equivalent.

(2) a. Oorang pada thiikam=**apa**
        man   PL   stab=CP
        'People having been stabbed,/People were stabbed and'
    b. oorang pada=nang theem̆bak=**apa**
        man   PL=DAT   shoot=CP
        'people having been shot,/people were shot and'
    c. se=dang bannyak creeveth pada su-aada.
        1s=DAT much   trouble PL   PAST-exist
        'I had a lot of problems.' (K051213nar01)

(3) a. Siithu **asà**-blaajar=**apa**,
        there CP-learn=after
        'After having learned there,'
    b. thaaun nnamblas=ka se skuul **asà**-luppas=**apa**,
        year   sixteen=LOC 1s school CP-leave=CP
        'after having left the school at 16,'
    c. pukuran asà-caari anà-pii.
        work   CP-find PAST-go
        'I looked for work and went (away).' (K060108nar01)

The Sri Lanka Malay copula is built around the lexical verb *dhaathang* 'come' in its conjunctive participle form. All three forms mentioned above are possible for the copular use. The following three examples illustrate this.[1]

(4) Se=ppe naama **asà**dhaathang Cintha Sinthani.
     1s=POSS name COPULA   Chintha Sinthani.
     'My name is Chintha Sinthani.' (B060115prs04)

(5) Se=ppe baapa dhaathang**apa** Jinaan Samath.

_____

[1] Note that the word order in copula sentences is NP Cop NP, while in verbal sentences, it is typically NP NP V.

1s=POSS father COPULA   Jinaan Samath.
'My father was Jinaan Samath.' (N060113nar03)

(6)   *Estate*=pe *field officer* **asà**dhaathang**apa** kithang=pe kaake
         estate=POSS field officer COPULA   1PL=POSS grandfather.
         'The estate field officer was our grandfather.' (N060113nar03)

The three examples above illustrate the use of *(asà)dhaathang(apa)* as a copula. Any of the three forms mentioned can also be used as a conjunctive participle proper when combined with *dhaathang* 'come', as shown in the following three examples.

(7)a.   Moonyeth pada=le **asà-dhaathang**
        monkey   PL=ADDIT CP-come
        'The monkeys would go and'
   b.   creeveth athi-kaasi.
        trouble IRR-give
        'cause (some other) trouble.' (K070000wrt01)

(8)a.   oorang mlaayu siithu=dering **dhaathang=apa**
        man   malay there=ABL   come=CP
        'The Malay men came from there and'
   b.   cinggala raaja=nang=le anà-banthu
        Sinhala king=DAT=ADDIT PAST-help
        'helped the Sinhalese king.' (K051206nar04)

(9)a.   incayang islaam=nang **asà-dhaathang=apa**
        3s.POLITE Islam=DAT CP-come=CP
        'He came to Islam'
   b.   inni siigith=nang asà-dhaathang
        PROX mosque=DAT CP-come
        'and came to the mosque'
   c.   giini girja
        like. this make
        'and does like this (=observes Islamic rites).' (K051220nar01)

An interpretation involving an event of coming is mandatory in examples (7)-(9), but impossible in examples (4)-(6). We are, thus, dealing with polysemy of the form *(asà)dhaathang(apa)*.

In the languages of the world which have a copula, the copula is normally used for one or more types of non-verbal predicates (Stassen

1997). The predication types which are often found to require support from a copula are property assignment (*John is tall*), class inclusion (*John is a doctor*), existence (*There is a solution*) and location (*John is in Dubai*). Additionally, identification (*The murderer was the butler; I am John*) and specification (*The president is Obama*) also often require the use of a copula (Hengeveld 1992). This is exactly what we find in English, where the copula *to be* is present in all these cases.

In Sri Lanka Malay, the copula is mainly used for naming people and for indications of group membership. Property assignment or location are not areas which trigger the use of the copula. Among these uses, naming is the most common and has been illustrated in (4) and (5) above. Apart from naming, the copula is most frequently used for indicating membership in ethnic, professional or gender classes or groups. These three uses are given in (10)-(12). The three forms with *asà-*, *=apa* and *asà-...=apa* appear to be used indiscriminately for these functions.

(10)a. Se=ppe *daughter-in-law*=pe *mother* asàdhaathang **bìnggaali**.
        1s=POSS daughter-in-law=POSS mother COPULA   Bengali
        'My daughter-in-law's mother is Bengali.'

   b. Ithukapang *daughter-in-law*=pe *father* asàdhaathang **mlaayu**.
        then   daughter-in-law=POSS father COPULA   Malay
        'Then my daughter-in-law's father is Malay.' (K051206nar08)

(11)   [Seelong=nang duppang duppang anà-dhaathang mlaayu]
asàdhaathang **oorang ikkang**.
        Ceylon=DAT   before   before   PAST-come   Malay]
COPULA   man   fish
        'The Malays who came to Ceylon very early were fishermen.'
(K060108nar02)

(12)   Kàthama aanak dhaathangapa **klaaki**.
        first   child COPULA   male
        'My oldest child is a boy.' (G051222nar01)

Another use frequently found is the indication of kinship relations, another type of class membership. Example (13) illustrates this use.

(13)   Baapa=pe umma asàdhaathang **kaake=pe   aade**.
        father=POSS mother COPULA   grandfather=POSS younger.sibling.
        'My paternal grandmother was my grandfather's younger

sister.' (K051205nar05)

What all the functions mentioned above, with the exception of naming, have in common is that they indicate group membership. Another function is the use of the copula in identificational or specificational sentences. The following two examples show the use of the copula to further specify or identify the nature of a problem.

(14)　[Itthu vakthu kithang=nang nya-aada] **asàdhaathang** ini JVP katha hathu *problem*.
　　　 DIST time　　1PL=DAT　　　　PAST-exist COPULA　　　 PROX JVP
QUOT INDEF problem.
　　　 'What we had at that time was the so-called JVP-problem.' (K051206nar10)

(15)　suda karang [kithang=nang aada *problem*] **dhaathangapa** kithang=pe aanak pada mlaayu thama-oomong
　　　 thus now　　1PL=DAT exist　　　　problem COPULA
1PL=POSS　　child　PL　Malay NEG.NONPAST-speak
　　　 'So, the problem we are having now is that our children do not speak Malay.' (G051222nar01)

To sum up, of the functions commonly associated with the copula, 'class membership' seems to be central for the Sri Lanka Malay case, while identification is another use. Identificational uses of the copula are very common as far as naming is concerned, but far less common for the types of identification and specification illustrated with examples (14)-(15).
　　　 The domains of property assignment and location do not seem to trigger the use of the copula in Sri Lanka Malay, as illustrated in (16) and (17).

(16)　Theem̌bok (*asàdhaathang) puuthi
　　　 wall　　　　COPULA　　　 white
　　　 (intended reading 'The wall is white.')

(17)　Tony (*asàdhaathang) Kluum̌bu=ka
　　　 Tony COPULA　　　　 Colombo=LOC
　　　 (intended reading '*Tony is in Colombo.', possible reading 'Tony has arrived in Colombo.')

An anonymous reviewer asks whether there are explanations as to why the copula cannot be used for property assingment and location. Synchronically speaking, there seems to be no good motivation for this. However, diachrony might provide a solution: As shown in example (17), *asàdhaathang* does not have the reading of 'copula' when combined with locations; instead, the meaning of 'to come' is preserved. As I will argue below, the development of the copula passes through a resultative reading (come → resultative → copula). It is possible to link the meaning of 'to come' to a result. The result of coming to Colombo is being in Colombo. Hower, the step from resultative to copula is not possible for locational predications since the original motion meaning is preserved in the full verb and will always remain, blocking the purported 'new' stative interpretation.
The grammaticalization path also provides the answer as to why property assignments do not combine with the copula. While for locational predicates, it is the second step of the path come → resultative → copula which fails, property assignment already misses the first step. In SLM, adjectival predications are stative. However, they get a dynamic ('inchoative', ' ingressive') reading when combined with verbal morphology (Nordhoff 2009, to appear).

() Thee diining
   tea cold
 'The tea is cold'

() Thee arà-diinging
   tea nonpast-cold
'The tea is getting cold'

() Thee su-diinging
  tea past-cold
'The tea became cold'

Crucially, no derivational morphology is involved. When combined with the right TAM-morphology, a resultative meaning emerges (). This categorial flexibility is a retention of a pattern widespread in Malay varieties (Gil 1994 and passim) and was already present when the first immigrants arrived in Sri Lanka (Paauw 2004, Nordhoff 2009, Nordhoff to appear). As such, there was no need to mark resultative for predications of property assignment overtly, e.g. by *dhaathang*. This means that step one of the proposed grammaticalization path did not

take place, impeding the further progression along this path. This distinguishes property assignment from other predication types, where a dynamic reading must be indicated overtly, today by *jaadi* 'become' and in former times by *dhaathang* 'come'.

Returning to class inclusion as described above, it often does show the copula but, it is by no means obligatory. The copula can be left out in all the sentences above without affecting truth conditions. This means that together with the sentence in (11), (18) is also grammatical.

(18)   [Seelong=nang duppang duppang anà-dhaathang mlaayu] Ø **oorang ikkang**.
          Ceylon=DAT before before PAST-come Malay] Ø man fish
          'The Malays who came to Ceylon very early were fishermen.'

  It appears that the use of the copula is conditioned more by considerations of information structure than by syntax or semantics. Especially topic-comment structures seem to favour the use of the copula (cf. Li & Thompson 1977:420). The stretch preceding the *(asà)dhaathang(apa)* is the topic, while the following portion is the comment. In that sense, (18) without the copula is less felicitous than (11), where the topic-comment-structure is overtly indicated by *(asà)dhaathang(apa)*. This is especially true since the copula prevents misparsing of the constituents as a right dislocation, which is possibly of the copula is omitted as in (). The non-dislocated sentence is given in ().

()   [Seelong=nang duppang duppang anà-dhaathang [mlaayu **oorang ikkang**]].
          Ceylon=DAT before before PAST-come Malay   man fish
          'It was Malay fishermen who came to Ceylon very early.'

()   [[mlaayu **oorang ikkang**] Seelong=nang duppang duppang su[2]-dhaathang**]**.
           Malay   man fish Ceylon=DAT before before PAST-come
          'Malay fishermen came to Ceylon very early.'

    It is not necessary for the topic to be nominal. In (19), we see an adverbial topic *dovulu* 'in former times', and the comment is preceded

2   The change in verbal prefix from *anà-* to *su-* is mandated by neutral information structure (Nordhoff 2009, Bourdin & Jaffar 2010).

by *abbisdhaathang*, an idiolectal variant of *asàdhaathang*. In (20), we have a spatial topic *siini* 'here' instead of a temporal topic.

(19) **dovulu** abbisdhaathang muuka thama-thuuthup
     earlier COPULA   face   NEG.IRR-close
     'Earlier, the women would not cover their faces.'
(K061026prs01)

(20) **siini** dhaathangapa mixed, mixed community
     here COPULA   mixed mixed community
     'What we get here is a mixed, a mixed community.'
(G051222nar04)

Occasionally, the comment can be quite elaborate as in the following two examples.

(21) Itthu abbisdhaathang *custard powder*=dering=jo arà-kirja.
     DIST COPULA   custard powder=ABL=EMPH NONPAST-make
     'As for this one, it is such that it is made with custard powder.'
(K061026rcp02,K081105eli02)

(22)a. Sepaktakraw=pe *rules* dhaathangapa
     sepaktakraw=POSS rules COPULA
     'The Sepaktakraw rules are as follows:'
  b. inni hathu *badminton court*=ka arà-maayeng
     PROX INDEF badminton court=LOC NONPAST-play
     'You play it on a badminton court;'
  c. *game* hatthu itthe *same measurement* ambel=apa; *height*=le *same*=jo
     game INDEF DIST same measurement take=after height=ADDIT same=EMPH
     '(For) a game, you take the same measurements, and then the height (of the net) is also the same.' (N060113nar05)

Furthermore, use of the copula seems to be strengthened by the presence of loanwords in the sentence, as is already evident from the preceding example. For the Sri Lankan context, most of the Malays have an above average command of English, and it cannot be excluded that the use of lexical material from English (loanwords) activates grammatical structures found in English as well, in this case the requirement for a copula in predications of class membership.

4 Where does the copula come from?

The copula is both synchronically and diachronically analyzable as the conjunctive participle of the lexical verb *dhaathang* 'come'. The first Malays who arrived in Ceylon in the 17th century communicated in Vehicular Malay, a trade language used in and around the Indonesian archipelago (Smith 2003, Smith et al. 2004). The language has changed considerably in the last three centuries and has typologically converged towards the local languages Tamil and Sinhala (Smith et al. 2004, Smith & Paauw 2006, Ansaldo 2005, 2008, Nordhoff 2009). Given these historical facts, three possible origins for the copula can be suggested: Indonesian varieties of Malay, Sinhala or Tamil. In this section, I will discuss the copula-like structures we find in these languages and the possibility of their influence on the development of Sri Lanka Malay.

4.1 Copulas in Indonesian varieties

In Indonesian there are two copulas, *adalah* and *ialah*, which are restricted to very formal varieties. Sneddon (2006:98f) informs us that

> [t]he copulas *adalah* and *ialah* optionally occur in F[ormal] I[ndonesian] to link a subject and nominal predicate, and sometimes an adjectival predicate [...] with *adalah* being considerably more common than *ialah*.
>
> *Adalah* also occurs in C[olloquial]J[akarta]I[ndonesian], although it is very rare; in about half the texts it did not occur. In most of the others it occurred with very limited frequency. It was only in the meetings and in the two most formal interviews [...] that it occurred frequently. [...] It can be said that *adalah* is marked for formality while *ialah,* being entirely absent form CJI, is very highly marked for formality.

The immigrants to Sri Lanka were mostly soldiers temporarily garrisoned in Batavia/Jakarta and spoke a very colloquial variety of Malay, remote from any formal standard (Adelaar 1991, Smith et al. 2004, Nordhoff 2009). It can be excluded that Formal Indonesian as developed in the 20th century had a significant impact on their speech. The sociolinguistic profile of Colloquial Jakarta Indonesian is closer to what would have been appropriate for the immigrants, although the relations between the colloquial varieties of the 17th century and the contemporary varieties are not exactly straightforward (Grijns 1991, Adelaar & Prentice 1996). The copula *adalah* can be found in CJI,

although it is extremely rare. The following is an example of the copula in informal speech.

(23)   Yang pasti   bakal ngebikin gua terpesona banget **adalah** cewek dengan rambut panjang yang tergerai, terus anggun CJI
        which certain will   make   me enchanted very   COPULA
girl   with   hair   long   which flowing then elegant
        'What would certainly very much enchant me is an elegant girl with long flowing hair.' (Sneddon 2006:79)

This specificational use of the copula in (23) is similar to example (14), repeated as (24) for convenience.

(24)   [Itthu vakthu kithang=nang nya-aada] **asàdhaathang** ini JVP katha hathu *problem*. SLM
        DIST time 1PL=DAT   PAST-exist COPULA   PROX JVP
QUOT INDEF problem.
        'What we had at that time was the so-called JVP-problem.' (K051206nar10)

  While the functional properties of the copulas in CJI and SLM are similar, they are morphologically very different. The Indonesian form is composed of the existential plus the imperative/emphatic clitic, but the Sri Lankan form is composed of the verb 'to come' and the conjunctive participle prefix *asà-*. Furthermore, the Indonesian copula is a development of the 20th century, as an anonymous reviewer informs me. The possibility that a construction involving a copula crossed the Bay of Bengal with the immigrants can therefore be excluded An additional argument against a historical relatedness between the SLM and CJI constructions is the use of *dhaathang* 'come' in SLM. *Dhaathang* would have had to replace the existential *ada* without any particular motivation. One reason for the replacement of the morphological composition of the copula could be influence from the adstrates. This will be discussed in the following section.

4.2 Copulas in Sinhala and Tamil

 Colloquial Sinhala has no copula as such, but uses a predicative suffix *-yi* on adjectives ending with a vowel (25). Adjectives which end in a consonant (26), and other word classes ending in any sound do not take this suffix (27)-(28).

(25)   mee amu miris hari   særa-yi   COLLOQUIAL SINHALA
         PROX raw chillis really strong-YI
         'These green chillies are really strong.' (Gair 2003:794)

(26)   mee dawaswələ haal bohomə ganang???(*-yi) COLLOQUIAL
SINHALA
         PROX days   rice very   expensive(-YI)
         'These days (uncooked) rice is very expensive.' (Gair
2003:795)

(27)   mamə govi-yek(*-yi)   COLLOQUIAL SINHALA
         1s   farmer-INDEF.ANIM(-YI)
         'I am a farmer.' (Gair 1998:241)

(28)   maʈə - niwaaɖu(*-yi)   COLLOQUIAL
SINHALA
         1s-DAT vacation(-YI)
         'I'm on vacation.' (Gair 2003:794)

  In Literary Sinhala, a similar suffix *-mi* is also used for nominal
predications, as in (29).

(29)   mama goviyekmi   LITERARY SINHALA
         1s-NOM farmer.NOM.1SG
         'I am a farmer.' (Gair 1998:242)

  There is some discussion of whether *mi* in (29) should indeed be
analyzed as a copula, or rather as an agreement suffix (Gair 1998:242).
Whatever the right analysis of *-mi* and similar forms, it is clear that the
variety of Sinhala the Malays were exposed to did not feature this form.
The most copula-like element they are likely to have encountered is *-yi*,
which does not suggest itself as a possible model for the Sri Lanka
Malay copula. Its form as a suffix is very different from the verbal form
employed by Sri Lanka Malay, and its use for adjectival predications is
precisely one where the use of the Sri Lanka Malay copula is not found.
     In Tamil, the existential *iru* is sometimes analyzed as a copula.
This verb is used for predications of existence as in (30) but also for
locational predications as in (31).

(30)   kumaar-ukku oru paiya<u>n</u> iru-kki<u>r</u>-aa<u>n</u>   TAMIL

Kumar-DAT   one boy   IRU-PRES-3SM
'Kumar has one boy/There exists one boy to Kumar.' (Lehmann
1989:188)

(31)   inta·k kiraama-tt-il   muunru koovil (iru-kkir-atu)   TAMIL
       this   village-OBL-LOC three   temple IRU-PRES-3SN
       'There are three temples in this village.' (Lehmann 1989:173)[3]

   Uses for property word predications (32) and nominal predications
(33) are ungrammatical.

(32)   paippu   mukkiyam (*irukkiratu)   TAMIL
       education importance (IRU.3SN)
       'Education is important.' (Lehmann 1989:172)

(33)   kumaar vakkiil (*irukkiratu/*irukkiraan)   TAMIL
       Kumar lawyer   (IRU.3SN/IRU.3SM)
       'Kumar is a lawyer.' (Lehmann 1989:171)

  While in (32) and (33) the English translational equivalent has a
copula, it does not seem appropriate to grant the same status to the
Tamil form. An analysis as an existential can cover both the meanings
given in (30) and (31), namely 'A boy exists (for Kumar)' and
'Temples exist in this village', while excluding the readings 'Education
exists importance' and 'Kumar exists a lawyer'. This is precisely what
we find in (32) and (33). It is true that *iru* is sometimes found in
nominal predications as in (34), but in that case, the noun is
adverbialized by *aakalaay*, and the construction can actually be shown
to have a locational reading as well (cf. Lehmann 1989:174).

(34)   kumaar vakkiil-aakalaay iru-nt-aan   TAMIL
       Kumar lawyer-ADV   IRU-PAST-3SM
       'Kumar was a lawyer.' (Lehmann 1989:174)

  Whatever analysis one prefers for *iru*, this Tamil form seems to be
more closely related to the Sri Lanka Malay existential *aada*, which

---

3The dot · separates the grammatical word from the first part of a long consonant which is the result of
certain morphosyntactic environments. The words in isolation would be *inta* and *kiraama*. One could also
choose to write *inta kkiraama* to correctly represent pronunciation and orthography, but Lehmann chooses
to leave the first part of the long consonant as the coda of the preceding syllable, although he separates it
with a dot to indicate the special status.

would be used for translations of (30) and (31), given in (35) and (36).[4]

(35)   Ini   nigiri=ka   thiiga koovil aada   SLM
        PROX village=LOC three   temple EXIST
        'There are three (Hindu) temples in this village.'

(36)   Kumaar=nang hatthu aanak klaaki aada   SLM
        Kumar=DAT   one   child male   EXIST
        'Kumar has one boy.'

While *iru* and *aada* are quite similar, the parallels between *iru* and *(asà)dhaathang(apa)* are less striking, and it seems unlikely that Sri Lanka Malay emulated the Tamil form *iru* with *(asà)dhaathang(apa)*. Sinhala and Tamil both have conjunctive participles as well. The conjunctive participle forms of the verb meaning 'to come' in Sinhala (*ævillaa*) and Tamil (*vantu*) are not used in a copular sense in these languages; they can only be used in the original reading of temporal anteriority.

(37)a. lamayaa pantiya=ʈa **ævillaa**,   SINHALA
       child   class=DAT come.CP
       'The child came to the class and/The child having come to
class,'
   b.   liyuma livvaa
      letter   write
      '(He) wrote the letter.' (Karunatillake 2004:161)

(38)a. naaʟai   kumaar **va-ntu**,   TAMIL
       tomorrow Kumar come-CP
       'Tomorrow Kumar will come and/Kumar having come
tomorrow,'
   b.   *eṅkaʟ-ai·c canti-pp-aan*
      1PL-ACC   meet-FUT-3SM
      '(he) will meet us.' (Lehmann 1989:267)

To sum up, while the Malay ancestor language, and Sinhala and Tamil all have structures which come at least close to a copula, none of these structures is similar enough to the SLM structure as to suggest itself as

---

[4]Note that *aada* SLM is only an existential, whereas the historically related form in Colloquial Jakarta Indonesian *adalah* (Section 4.1) is a copula which can also be used for non-existential predications, e.g. as in (23).

the driving force behind the development of the copula in Sri Lanka Malay. It thus seems likely that the copula in Sri Lanka Malay is an independent development.

5 The emergence of copulas

The most common etymological source for copulas is an existential (Munro 1977) or a positional verb (Stassen 1997:94). Stassen illustrates this with data from Kiowa, gathered by Watkins; I shall repeat this data here.[5] In (39) we see the last reflexes of the original positional/locational use of *dɔ́·*. This locational use has given way to the general use of dɔ́· as a copula, which can be seen in (40). In return, most of the original uses of dɔ́· for locational purposes are now expressed by other morphemes, e.g. *-cél* in (41). Dɔ́·: has thus nearly fully completed the grammaticalization from locational verb to copula.

(39)   p'ɔ́· hę́· gyà-dɔ́·   KIOWA
       moon gone PL-be
       'There is (temporarily) no moon.' (Watkins 1984:216)

(40)   té· kɔ́ygú bà-dɔ́·   KIOWA
       all Kiowa 2PL-be
       'You are all Kiowas.' (Watkins 1984:227)

(41)   ę́·gɔ́ yí· ɔ́l ę̀-cél   kí·còy-kyà   KIOWA
       here two hair 3DUAL-be.in soup-in
       'There are two pieces of hair in the soup.' (Watkins 1984:211)

The other common source for copulas are pronouns. This can be seen in the famous example from Chinese (Li & Thompson 1977).

(42)   jì yù qí shēng yòu yù qi sý, **shì** huò yě   ARCHAIC CHINESE
       already wish 3SG live also wish 3SG die this indecision DCL
       'Wishing him to live and wishing him to die, this is indecision.'
(Li & Thompson 1977:424)

The demonstrative *shì* in this construction type was reanalyzed as a copula *shì* (Li & Thompson 1977:424).

---

[5]Stassen uses the 1980 PhD-thesis; in this article, I use the newer version (Watkins 1984) and put stripped diacritics back in place.

(43)   nèi-ge rén **shì** xuéshēng   MODERN MANDARIN
             that-CLF man COPULA student
             'That man is a student' (Li & Thompson 1977:422)

In addition to Chinese, this pattern is for instance found in Hebrew and Arabic as well as in a number of North American and African languages (Stassen 1997:77–91).

    With regard to Sri Lanka Malay, we can say that the existential *aada* is different from *asàdhaathangapa* and does not seem to be very much grammaticalized towards a copula. The deictics *ini* and *itthu* are not recruited for copular purposes either.

    Changes from lexical meanings like 'coming' to more grammatical meanings like 'COPULA' have been investigated under the umbrella of grammaticalization. The verb 'come' is a frequent source of a number of grammatical functions so that an investigation of the directions of development of this verb might shed light on what we find in Sri Lanka Malay. Stassen (1997:92) notes that a minor class of sources for the grammaticalization of copulas includes motion verbs like 'go' and 'come', but unfortunately he does not cite languages illustrating the 'come'-case so that it is not possible to compare the Sri Lanka Malay case to other purported instances of this path.

6 Grammaticalization paths of COME

 In the languages of the world, verbs meaning 'to come' often undergo grammaticalization and can be used for a variety of grammatical functions. In French, *venir de* is used to express immediate past, but the venitive meaning is bleached: it is possible to say *Jean vient de partir*, literally 'John comes from leaving', meaning that John has just left (and not that he has just come back). Heine & Kuteva (2002:68–78) list a number of grammaticalization targets for COME: consecutive, continuous, hortative, venitive, ablative, near past, benefactive, change-of-state, future, proximative and purpose. The copula is not found among them. The closest example to a copular use of COME I am aware of is the use as an auxiliary in passive constructions, as found in Italian. In Italian, the verb *venire* can be used to form passives as in (44). Those passives always have a dynamic interpretation.[6]

(44)   Era chiaro che i ragazzi veni-vano   maltratta-t-i ITALIAN
            was clear   that the children COME-3PL.IMPFT illtreat-PTCPL-PL

---

[6]For some diachronic information on this development, see Michaelis (1997) and Giacalone Ramat (2000).

'It was obvious [that] the children were being ill-treated.'
(Maiden & Robustelli 2000:282)

 As in the Sri Lanka Malay case, the verb meaning 'to come' has lost its original venitive meaning. There is no motion towards the deictic center involved in (44). In Italian as in Sri Lanka Malay, the verb meaning 'to come' is completely bleached and serves a purely grammatical function, which would in both cases be fulfilled in English by the copula 'to be'. However, the similarities end here. In Italian, *venire* is used as an auxiliary for diathesis, which is part of a larger construction involving a past participle. Both the auxiliary and the participle agree in number with the subject of the sentence. In the Sri Lanka Malay case, the absence of other verbs in the sentence suggests that *asàdhaathang* is not an auxiliary. There is no particple either, let alone agreement. Furthermore, the Sri Lanka Malay copula typically surfaces in intransitive, non-verbal predications, while the Italian passive construction with *venire* requires at least a bivalent verb. The Italian construction has a dynamic reading, while in the Sri Lanka Malay cases, a stative reading is the norm. For naming uses for instance, a dynamic reading is completely excluded. This all suggests that the grammaticalization path of Sri Lanka Malay *dhaathang* 'come' is not comparable to Italian *venire*, nor to any other of the cases discussed in Heine & Kuteva (2002). With the knowledge we have today, it seems that this grammaticalization from COME to a copula is unique to Sri Lanka Malay.

7 Possible grammaticalization paths

In the preceding sections, I have ruled out inheritance of the copula construction or emergence through language contact. I have also shown that the existing literature on grammaticalization paths does not cover the phenomenon encountered in Sri Lanka Malay. Given this paucity of literature to hook on to, some speculation as to the possible development may be warranted. One possible path from COME to the copula would lead over an intermediate resultative stage. This resultative reading of come is found for instance in English constructions like (45).

(45) A dream come true

 In this case, the original motion meaning of *to come* has been lost. There is no spatial motion involved. What remains is the change of

state from A to A', which is also what we observe in motion events like *John came home from work*. In (45), the state of the dream changes from unrealized to realized/true.[7] While in English, this resultative reading is restricted to a small set of words (*A man come tall, *A T-shirt come grey*), in other languages, it has a wider application. In To'aba'ita, the verb *mai* 'come' has an ingressive/resultative meaning when used with a property concept as in (46).

(46) fanua-'e   rodo   na-mai   To'aba'ita
     place-it:PFV be.dark PERF-come
     'It has become dark.' (Lichtenberk 1991:487)

Heine & Kuteva (2002:74) note that "[t]his grammaticalization appears to be particularly common in pidgin and creole languages." While it is not entirely clear whether Sri Lanka Malay should be seen as a Creole (Smith & Paauw 2006) or not (Ansaldo 2008, Nordhoff 2009), it is clear that the immigrants to Sri Lanka spoke Vehicular Malay, a trade language for inter-ethnic communication which has been likened to a pidgin (Adelaar & Prentice 1996). As such, it is interesting to note that Fa d'Ambu Creole Portuguese (47), Guyanese Creole French (48), and Seychelles Creole French (49) show the resultative reading of the verb meaning 'to come' as well.[8]

(47) tyipa   bi   sxa dual eli kumu pasa   Fa d'Ambu
     stomach come PART hurt 3s east surpass
     'His stomach hurt; he had eaten to much.' (Post 1992:159)

(48) i vini malad   Guyanese Creole French
     3s come sick
     'He has become sick.' (Corne 1971:90)

(49) i n vin larpâter   Seychelles Creole French
     he COPULA come surveyor
     'He became a surveyor.' (Corne 1977:80)

The grammaticalization from 'to come' to a resultative marker is the first step in a grammaticalization chain I want to propose. The second

---

[7]Note also that the string *come* is present in English *become*, and *venir* 'come' is found in French *devenir* 'become' and several Romance cognates.

[8]Furthermore, this is also found in Ghanaian Pidgin English (Huber 1996) and Chinook Jargon (Grant 1996:236).

case involves the change from resultative to stative. For this case, especially the Seychelles Creole French example in (49) is interesting.[9] It shows a great resemblance to one of the main uses of the SLM copula, namely membership in a (professional class) as in (11) or also in (50) below.

(50)   Umma=pe baapa dhaathangapa hathu *inspector of police*. SLM
        mother=POSS father COPULA   INDEF inspector of police.
        'My mother's father was an inspector of police.'
(N060113nar03)

The difference between the Seychelles Creole French example in (49) and the Sri Lanka Malay example in (50) is the change from resultative to stative. The Seychellois sentence is necessarily resultative, while the SLM sentence is necessarily stative. In the Sri Lanka Malay sentence, the information is not about the grandfather having passed an exam, thereby becoming an inspector of police (unlike the Seychellois surveyor). Rather, it is the first mentioning of the grandfather in this text, indicating his profession as a stative predicate without forcing a resultative interpretation.

However, it is logically necessary that the grandfather must have made the step from prospective inspector of police to certified inspector of police at one point in time since it is unlikely that he was born inspector. With many predicates there is a reciprocal entailment between the stative reading and a prior ingressive/resultative one. This is especially true for predicates of profession: If you become a

9   An anonymous reviewer points out that *vin* in (49) might be a reflex of French *devenir* 'become' rather than *venir* 'come'. Today, there is *devenir* in Seychellois, but this is clearly acrolectal and a recent borrowing from Standard French (Susanne Michaelis, p.c.). In more basilectal speech, which reflects earlier stages of the language, it is not possible to distinguish between reflexes of Standard French *venir* and *devenir*. The phonological weakness of the first syllable of *devenir* makes it likely that both forms would have ended up as *vin* in Seychellois. But the indistinction could also be older: Middle French dialects often use *venir* for *devenir* \citep{Chaudenson, EFW, Bolle}. The lexifier the creators of Seychellois were exposed to thus possibly lacked the *devenir*-form altogether. In that case, the first step in the grammaticalization chain would indeed not have taken place in a Creole, but in a precolonial stage of the lexifier language, French in this case. Ultimately, the very form *de-venir* stems from Latin *venire*, so that the fact whether or not there was a *de-* when the language arrived in the Indian Ocean is not substantial to the nature of the claim that COME developed a resultative reading. It only has a bearing on the timing: in the Creole (Seychellois), in the lexifier (Middle French dialects), or in an earlier stage of the lexifier (Vulgar Latin).

professor, you are then a professor; if you are a priest, you must have become a priest before.[10]

The second step of the grammaticalization chain is then the change from a resultative reading entailing stative to a pure stative reading. This is likely to have been pioneered by predicates of profession, which are still one of the most frequent predicates found with the copula in Sri Lanka Malay. Only later would other predicates where a resultative meaning is impossible (e.g. sex) have followed suit. The SLM copula has not completed the full journey yet, though: predications of location and property assignment are still not possible with the copula. SLM has come a long way in developing a copula, but still falls short of a full-fledged copula as we find it for example in English. The proposed development can then be summarized as follows:

(51) come
        > resultative
                > resultative entailing stative
                        > copula (professions)
                                > copula (class inclusion, naming)
                                        > (copula (location, prop. assgn.))

8 Conclusion

Sri Lanka Malay has a copula, which is mainly used for naming and predications of class membership, although some information structure use, especially with regard to topic and comment, is also found. This copula is derived from the conjunctive participle form of the verb *dhaathang* 'come'. Neither the formal properties of the copula, nor its functional uses show parallels to any of the possible languages of origin, i.e. Indonesian varieties, Sinhala, or Tamil. This suggests an independent development. This independent development, however, is not of a type attested in the literature on grammaticalization and seems to be a new grammaticalization path which has not yet been found. While further research is needed, a promising hypothesis is the

---

10 This entailment is also found in some Indonesian varieties of Malay, as an anonymous reviewer informs me:
() *Dia jadi guru di SMA*.
He become teacher of high.school
'He has become a high school teacher.'
'He is a high school teacher.'

development through an intermediate stage of a resultative marker, as found in other contact languages like Seychelles Creole French. This marker would then lose its resultative reading for predicates of profession and become a semantically void copula, which then spreads to other types of intransitive predication.

Abbreviations

ABL	ablative
ACC	accusative
ADV	adverbializer
ANIM	animate
CLF	classifier
CP	Conjunctive participle
DIST	distal
EMPH	emphatic
FUT	future
IMPFT	imperfective
INDEF	indefinite
IRR	irrealis
LOC	locative
M	masculine
N	neuter
NEG	negative
NOM	nominative
PL	plural
PART	
PERF	
PFV	
POSS	possessive
PRES	present
PROX	proximal
PTCPL	participle
QUOT	quotative
S	singular

References

ADELAAR, K. A. (1991). "Some notes on the origin of Sri Lankan Malay". In H. Steinhauer (ed.), Papers in Austronesian Linguistics No. 1, Pacific Linguistics Series A-81. Canberra: The Australian National University. pp. 23–37.

ADELAAR, K. A. & D. J. PRENTICE (1996). "Malay: its history, role and spread". In Wurm, Stephen A., Mühlhäusler, Peter, and Tryon, Darrell T. (eds.) *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas (Volume II.1: Texts)*. Walter de Gruyter. pp. 673–693.

ANSALDO, U. (2005). "Typological admixture in Sri Lanka Malay. The case of Kirinda Java." Unpublished manuscript, available at http://home.medewerker.uva.nl/u.ansaldo/index.html.

ANSALDO, U. (2008). "Sri Lanka Malay revisited: Genesis and classification". In DWYER, A., D. HARRISON & D. ROOD (eds.) "A world of many voices: Lessons from documented endangered languages", Studies in Language, Amsterdam, Philadelphia: John Benjamins. pp. 13–42.

BAKER, P. & A. SYEA (eds.) (1996). *Changing meaning , changing functions: Papers relating to grammaticalization in contact languages*. London: University of Westminster Press.

BAKKER, P. (2006). "The Sri Lanka *Sprachbund*: The newcomers Portuguese and Malay". In MATRAS, Y., A. MCMAHON & N. VINCENT (eds.) "Linguistic Areas – Convergence in Historical and Typological Perspective", Houndmills, New York: Palgrave Macmillan. pp. 135–159.

CORNE, C. (1971). "Le patois créole français de la Guyane (St. Laurent de Maroni): Esquisse de grammaire". *Te Reo*, 14:81–103.

CORNE, C. (1977). *Seychelles Creole grammar: Elements for Indian Ocean Proto-Creole reconstruction*. Tübingen: Narr.

DEUMERT, A. & S. DURLEMAN (eds.) (2006). *Structure and Variation in Language Contact*. Amsterdam: John Benjamins.

GAIR, J. W. (1998). *Sinhala and Other South East Asian Languages*. New York, Oxford: OUP.

GAIR, J. W. (2003). "Sinhala". In CARDONA, G. & D. JAIN (eds.) "The Indo-Aryan Languages", London: Routledge. pp. 766–818.

GIACALONE RAMAT, A. (2000). "On some grammaticalization patterns for auxiliaries". In J.C.SMITH & D. BENTLEY (eds.) "Historical Linguistics 1995. Vol.I: General issues and non-Germanic languages", Amsterdam: John Benjamins. pp. 125–154.

GRANT, A. P. (1996). "The evolution of functional categories in Grand

Ronde Chinook Jargon: Ethnolinguistic and grammatical considerations". In Baker & Syea (1996), pp. 225–242.

GRIJNS, C. D. (1991). *Jakarta Malay*. Ph.D. thesis, Universiteit Leiden.

HEINE, B. & T. KUTEVA (2002). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

HENGEVELD, K. (1992). *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter.

HUBER, M. (1996). "The Grammaticalization of aspect markers in Ghanaian Pidgin English". In Baker & Syea (1996), pp. 53–70.

HUSSAINMIYA, B. A. (1990). *Orang Rejimen: The Malays of the Ceylon Rifle Regiment*. Universiti Kebangsaan.

KARUNATILLAKE, W. S. (2004). *An Introduction to Spoken Sinhala*. Colombo: Gunasena, 3rd edn.

LEHMANN, T. (1989). *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.

LI, C. N. (1977). *Mechanisms of syntactic change*. Austin: University of Texas Press.

LI, C. N. & S. THOMPSON (1977). "A mechanism for the development of copular morphemes". In Li (1977), pp. 419–444.

LICHTENBERK, F. (1991). "Semantic change and heterosemy in grammaticalization". *Language*, 67(3):475–509.

MAIDEN, M. & C. ROBUSTELLI (2000). *A Reference Grammar of Modern Italian*. Lincolnwood: NTC.

MICHAELIS, S. (1997). "Antikausativ als Brücke zum Passiv: fieri, venire und se im Vulgärlateinischen und Altitalienischen". In DAHMEN, W., G. HOLTUS, J. KRAMER, M. METZELTIN, W. SCHWEICKARD & O. WINKELMANN (eds.) "Neuere Beschreibungsmethoden der romanischen Syntax", Tübingen: Narr.

MUNRO, P. (1977). "From existential to copula: The history of Yuman *be* ". In Li (1977).

NORDHOFF, S. (2009). *A Grammar of Upcountry Sri Lanka Malay*. Ph.D. thesis, University of Amsterdam.

POST, M. (1992). "The serial verb construction in Fa d'Ambu". In "Actas do Colóquio Sobre Crioulos de Base Lexical Portuguesa", Lisbon: Edicoes Colibri. pp. 153–169.

SLOMANSON, P. (2006). "Sri Lankan Malay morphosyntax: Lankan or Malay? " In Deumert & Durleman (2006), pp. 135–158.

SMITH, I. R. (2003). "The provenance and timing of substrate influence in Sri Lankan Malay; Definiteness, animacy and accusative case marking." Paper presented at the South Asian Language Analysis Roundtable XXIII, Austin, Texas.

SMITH, I. R. & S. PAAUW (2006). "Sri Lanka Malay: Creole or convert". In Deumert & Durleman (2006), pp. 159–182.

SMITH, I. R., S. PAAUW & B. A. HUSSAINMIYA (2004). "Sri Lanka Malay: the state of the art." *Yearbook of South Asian Languages*, 2004:197–215.

SNEDDON, J. N. (2006). *Colloquial Jakartan Indonesian*, *Pacific Linguistics*, vol. 581. Canberra: Australian National University.

STASSEN, L. (1997). *Intransitive Predication*. Oxford: Clarendon Press.

WATKINS, L. J. (1984). *A grammar of Kiowa*. Lincoln, London: University of Nebraska Press.