Nominalizers (and Copulas) in Malay

Foong Ha Yap
Chinese University of Hong Kong

1. Introduction

In the course of daily—even mundane—communication, we often need to refer not only to concrete entities such as animate beings, inanimate objects, locations and times (i.e. first order ontological entities according to Lyons, or very simply the who, what, where and when elements in news reports and the someone, something, someplace and sometime elements in narratives); we also need to refer to abstract entities such as events, perceptions, and perspectives (i.e. second and third order ontological entities). In linguistic terms, we need to deal with more complex constructions involving gerundives, infinitives, and other complementation structures. While first order entities are often readily expressed as noun phrases, second and third order entities are often realized as nominalized clauses. Previous studies from various languages have shown that the same morpheme that serves as nominalizing particles frequently serves other functions as well, including genitive and relative clause marker (Matisoff 1972; Delancey 1986; Herring 1991; Noonan 1997), complementizer (Horie 1998), and speaker mood or stance marker (Noonan 197; Yap, Matthews & Horie 2004; Watters, in press; Yap & Matthews, in press). Detailed analyses of versatile nominalizers have been undertaken for a number of Asian languages, among them Tibeto-Burman languages, including Atong (van Breugel 2006), Burmese (Simpson 2006), Bodic and Tamangic languages (Noonan 1997, in press), Kiranti (Bickel 1999), Mongsen Ao (Coupe 2006), Rawang (LaPolla 2006), Nimpuk Singhpo (Morey 2006), and Thulung Rai (Lahaussois 2003), and among the East Asian languages, Japanese and Korean (Horie, in press; Rhee, in press), and the range of functions identified with nominalization phenomena in these studies are proving useful not only for syntactic theory but also for scholars working on erstwhile nominalizers in classical languages such as Archaic Chinese, Old Japanese, and Old Korean.

The present paper examines nominalization phenomena from another language phylum, namely Western Austronesian. More specifically, we examine three Malay nominalizers—namely, yang, -nya and punya. Each of these nominalizers serves a number of other functions, some of which overlap to some extent. The focus of the present paper is to identify the functions and grammaticalization pathways of each of these morphemes, and how they interact with each other, and in due course carve out distinct semantic niches for themselves. Particular attention is given to three factors that have contributed to the functional differentiation of these nominalizers, namely, their etymological differences and their
morphosyntactic environments—i.e. semantic and structural constraints respectively—and the sociological factors of language contact.

This paper is organized as follows: §2 focuses on the various functions and possible diachronic development of *yang*, based on a review of extant literature, while §3 similarly discusses the range of functions and grammaticalization pathways of *-nya*, drawing on insights from Englebretson's (2003) conversational analysis of colloquial Indonesian. §4 examines *punya* in similar fashion, using data from classical texts made electronically available via the Malay Concordance Project housed at the Australian National University (courtesy of Ian Proudfoot); in this section, we further add to the discussion the impact of language contact on the rise of *punya* as a nominalizer and stance marker. §5 compares similarities and differences in the grammaticalization trajectories of these three nominalizers, with implications for robust crosslinguistic tendencies.

2. Nominalizer *yang*

Previous scholars (e.g. van Wijk 1909; van Ophuijzen 1910; Mees 1969; Verhaar 1983; Kaswanti Purwo 1983; Simin 1988; Steinhauer 1992; van Minde 2003) have identified a wide range of grammatical functions for *yang*, including relative clause marker, complementizer, theme/topic marker, and ‘defining article’ (particularly in contrastive contexts), as illustrated in (1) to (4) below.¹

(1) *Buang saja bunga yang sudah layu.*  
throw  just  flower  REL  PERF  wither  
‘Just throw away the flowers that have withered.’

(2) *Dia tahu (yang) orang ini tak boleh dipercayai.*  
3SG  know  COMP  person  this  NEG  can  PASS:trust:TRANS  
‘S/he knows (that) this person cannot be trusted.’

(3) *Yang dia, sedikit pun dia tak kesah.*  
TOP 3SG little.bit also 3SG NEG care  
‘As for him, he didn’t even care the slightest bit.’

(4) *Yang bapa ketawa; yang ibu menangis.*  
‘ART’  father  laugh  ‘ART’  mother  cry

¹ Except where specified, all the Malay examples used in this paper are of the Straits variety spoken in Perak and Selangor, West Malaysia.
'The father laughed; the mother cried.'

An interesting observation about yang is its link (sometimes subtle, sometimes strong) to referential definiteness. Not much is known about the origin of yang, but as highlighted in van Minde (2003), it is worth noting that Mees (1969) has argued for a link between Classical Malay yang and pre-Classical Malay yam, the latter identified as a definiteness marker in seventh-century Old Malay inscriptions from the Palembang and Bangka region in Indonesia. Adelaar (1985) has also further suggested the possibility that yang is derived from ia + -η (where ia is the third person pronoun while -η is a ligature or ‘linker’). This implicates a pronominal source for yang, and as such an inherent definiteness feature associated with (at least some of) its grammatical functions. We see evidence of this in (1) to (4) above, where the presence of yang contributes to referential definiteness via implicit contrastive comparison.

In relative clauses, as seen in (1), we obtain weak definiteness effects from an implicit contrast between ‘flowers that have withered’ vs. ‘flowers that have not withered’, the former expressed explicitly while the latter is left to inference. In factive complementation clauses, as seen in (2), the presence of yang suggests that the speaker is referring to information that is known to the hearer, with emphasis being given to the matrix verb—that is, the focus is on ‘the referent knowing (vs. not knowing) about some particular information’. In other words, the information being referred to in the complement clause is ‘definite’ by virtue of being shared information between the interlocutors. Absence of yang, on the other hand, renders a different interpretation: the person being referred to is said to know something (where ‘something’ is new information to the hearer), as opposed to that person being said to know something (where ‘something’ is already given information to the hearer).

A stronger definiteness interpretation emerges when yang is used as a topic marker, as seen in (3), where we obtain the reading ‘as for him (in contrast to others such as me/you/them/someone else), he couldn’t care less’. The definiteness reading comes not only from the definiteness feature presumably inherent in yang; it often also comes from definiteness features from other sources, such as when pronominal referents like dia

---

2 Other scholars have identified nan, nang, nyang, lang and jang as either dialectal cognates or orthographic variants of yang (e.g. Teeuw 1990; Wilkinson 1959; see also van Minde 2003).

3 Naturally, there is also a contrast with ‘flowers of other kinds’, or ‘flowers with other distinguishing characteristics’. Such a contrast, presumably underlies much (if not all) of cognitive processing, and need not be linguistically marked.
'him/her' are used. In addition, a definiteness interpretation also often comes for free by virtue of the topic position, which is typically reserved for definite referents.\footnote{This position is also available for generic referents. In the case of new referents in the discourse, these need to be immediately identified, often via cataphoric reference using relative clauses, e.g. Bunga yang sudah layu, buangkan saja ‘The flowers that have just withered, just throw (them) away’.}

Here it is worth noting that \textit{yang} encodes more than definiteness alone; it also encodes contrast as well. We see a clear example of this contrastive feature in the 'definite article' usage of \textit{yang} in (4), where the construction \textit{yang bapa ketawa, yang ibu menangis} can be uttered with subtle contrast to yield ‘the father laughed, (while) the mother cried’, or it can be uttered in topic-like fashion to yield the strong contrastive reading ‘as for the father, \textit{he} laughed; as for the mother, \textit{she} cried’.\footnote{A strong topic-like contrast is produced by stressing the topic noun phrase introduced by \textit{yang} (in this case, \textit{yang bapa} ‘(as for) the father’ and \textit{yang ibu} ‘(as for) the mother’, and separating it from the comment structure (e.g. \textit{ketawa ‘laugh’ and menangis ‘cry’}) with a distinctive ‘comma length’ pause. The definiteness effect of \textit{yang} emerges contextually from the discourse and via structural parallelism. Its association with topic marking emerges from the topic construction. In other words, while the use of topic constructions is discourse-motivated, the association of \textit{yang} with the topic marking function is structurally facilitated.}

Another usage of \textit{yang} that involves a strong contrastive and definiteness reading of the intended referent is found in cleft environments, as seen in (5) below. The intended reading here is ‘\textit{She} (not me/you/her/someone else) is the one who is getting married tomorrow’. In colloquial speech, the use of focus marker \textit{lah} is often optional, and a contrastive focus interpretation is still available without \textit{lah}. It is possible that \textit{yang} acquires some contrastive value from frequent association with focus marker \textit{lah}, much like the development of French negator \textit{pas} from \textit{ne...pas}-type constructions; at the same time, it should be noted that \textit{yang} in other contexts also yield an implicit contrastive reading, suggesting that contrastiveness (along with definiteness) may be an inherent feature of \textit{yang}. One could say then that certain syntactic environments such as cleft constructions, and also topic constructions as discussed earlier, strengthen the implicit contrastive feature of \textit{yang}.

\begin{verbatim}
(5) Dia (lah) yang nak nikah esok.
    3SG FOC NOMZ FUT marry tomorrow

‘She’s the one who’s getting married tomorrow.’ / ‘It’s she that’s getting married tomorrow.’
\end{verbatim}

We now come to another major function of \textit{yang}, namely, nominalization.\footnote{As noted in van Minde (2003), Mees (1969) identifies definiteness marking and nominalization as the two major functions for \textit{yang}.} As seen in (5) above, \textit{yang} introduces a nominal predicate, which is referential and bears the meaning ‘the
one who VP*, where VP must be specified. In (5) above, the use of yang yields the interpretation ‘the one who’s getting married’. In other words, yang is here used as a nominalizer that takes a clause [(NP) VP] and makes it available for use as a referential argument, whether as agent or patient nominal, in either subject or object position. An example of yang used in agent nominalization is highlighted in (6a), and in patient nominalization in (6b). Both examples involve yang-nominalized clauses in subject position.

(6) a. *Yang* nak nikah bukan aku.
    NOMZ want marry NEG 1SG
    ‘The one who wants to get married is not me.’

    b. Sebaliknya, yang dipaksakan nikah tu lah aku.
    on.the.contrary NOMZ PASS:force marry DEF FOC 1SG
    ‘On the contrary, the one that is being forced to marry, *that*’s me.’

Examples such as (5) above show that clauses nominalized by yang readily appear in cleft constructions. Examples such as (6) further highlight the ease with which clauses nominalized by yang occur in topic position. This is more clearly seen in (7) below, where a distinct pause separates the nominalized clause in topic position (yang nak nikah esok tu ‘the one who is getting married’) from the comment structure (bukannya aku ‘that person isn’t me’).\(^7\)

(7) *Yang* nak nikah esok tu bukannyaaku.
    NOMZ FUT marry tomorrow DEF(<DEM) NEG.3SG 1SG
    ‘The one who is getting married tomorrow isn’t me.’
    (Lit. ‘The one who is getting married tomorrow, that person isn’t me.’)

Note that the use of enclitic –nya (third person pronoun ‘him/her/it/that person/etc.’) following the negator bukan (‘not’) in (7) above is optional, as can be seen earlier in (6a). Its presence, though, reveals very clearly an underlying cleft (hence focus) construction. Such constructions, involving yang in topic and cleft constructions, yield strong speaker stance interpretations, often of an emphatic nature.

Let us now summarize the major functions of yang, as well as suggest how the various functions may be related. As noted earlier, not much is known about the origin of yang, although a pronominal link is likely in the form of third person pronominal ia, as suggested in

\(^7\) Here it should be noted that the use of demonstrative *tu* (a reduced form of distal demonstrative itu ‘that’) yields a definiteness effect by virtue of anaphoric reference. In topic constructions, it also has a topic-marking function; its presence helps reduce the pressure on nominalizer yang to double up as topic marker.
Adelaar (1985). More work is needed, both in terms of diachronic evidence and typological
comparison of ‘ligatures’ (or linking particles) across Austronesian languages. What is
evident, however, is that yang, to the extent that it is related to Old Malay definiteness marker
yan, yields a definiteness reading when it introduces a noun phrase (Mees 1969). How the
relative clause marking function of yang emerged is another area for further investigation.
Elsewhere in the literature, it has been argued that relative clause constructions could emerge
via parataxis (as in English I know that. He is coming. > I know that he is coming), or via
apposition of a (more fully specified) clause and a head noun (as in Old Chinese yue zhu zhi
guang ‘the moon emits that light’ > ‘the light that the moon emits’ (Yue 1998:263; see
also Delancey 1986 for Tibetan). Note that the English example involves postnominal
modification, while the Chinese example involves prenominal modification. In the case of
Malay yang, if Adelaar is correct, and yang was derived from the merger of a third person
pronoun ia and a pre-existing –ŋ ligature, we could be seeing an appositive construction
involving a head noun, followed by a nominal formed by a resumptive pronoun ia linked to a
modifying expression via the –ŋ ligature. Perhaps along the lines of [head noun NP] [appositive
nominal ia–ŋ (NP) VP] > [[head noun NP][modifying clause yang (NP) VP]].
Nominalizer yang constructions can occur independently of a head noun, and are
sometimes referred to as ‘headless’ relative clause constructions. Whether the use of
nominalizer yang constructions predates that of relative clause yang constructions is a
question that cannot be resolved without textual diachronic analysis. Nevertheless, it is quite
reasonable to assume that they may have emerged at about the same time, especially given
that the use of ‘headless’ constructions is heavily dependent on contextually retrievable
referents, while headed ones are useful for indefinite reference.

As discussed earlier, the availability of nominalizer and relative clause yang constructions
easily gives rise to the use of yang in factive complementation constructions as well as topic
marking constructions. What we have thus far then is the following pathway: adnominal (or
linking particle) yang > nominalizer and relative clause marker yang > complementizer yang
and topic marker yang (the latter involving preposed nominalized constructions). This
development is schematized in Figure 1.

---

8 Huang (2007), among others, has noted that simplex verbs were common in Old Chinese. These simplex verbs
were derived from nouns. Thus, the noun yu ‘fish’ could also be used to mean ‘to fish’. From Middle Chinese
onwards, it was necessary to use more analytic means such as ‘do fish(ing)’. Here it is worth noting that the
availability of simplex verbs could have facilitated the rise of relative clauses as follows: ‘the moon emits (light),
that light’ > ‘the moon emits , that light’ > ‘the moon emits REL light’.
3. Nominalizer -nya

In addition to yang, there is another nominalizer in Malay of relatively ancient origin. The morpheme in question is enclitic -nya. It occurs with high frequency and serves a wide range of functions. Englebretson (2003) identifies the following major functions for -nya: third person pronoun, possessive pronoun, ‘identifiability’ (i.e. definiteness) marker, nominalizer, and adverbial marker (see (8) to (12) respectively).

(8) Jangan dibuangnya.
    do not PASS:throw:3SG
    ‘Don’t throw it away.’

(9) Potong saja ekornya.
    Cut just tail:3SG-POSS
    ‘Just cut off its tail.’

(10) Kainnya terlalu halus.
    cloth:3SG-DEF excessive fine
    ‘The fabric is too fine.’ (< ‘Its fabric is too fine.’)

(11) Budak 'ni makannya tak ikut masa.
    child this eat:NOMZ(<3SG-POSS) NEG follow time
Lit. ‘As for this child, his eating is not according to schedule.’
Intended meaning: ‘This child is not eating regularly.’

(12) Biasanya dia awal.
usually 3SG early
‘Usually he’s early.’

Note the cline of abstraction in the above examples. In (8), enclitic -nya appears as a third person pronominal object after a transitive verb buang (‘throw away’). In (9), -nya appears as a possessor pronoun following the possessee head noun. In this context, pronominal -nya serves both referential function (as possessor pronoun) and relational marker (as possessive linker). In (10), -nya appears after a noun in a context where the focus is now less on the possessive (or modifying) relationship but rather more on identifiability of the head (i.e. possessee) noun. Englebretson (2003:162) also posits a semantic extension from possessive to identifiability marker for enclitic -nya.

In (11), -nya appears as an enclitic after the verb (makan ‘eat’), and yields an event nominal interpretation (equivalent to English gerundive ‘his eating’/ ‘the eating’). In this context, -nya functions as a nominalizer. Unlike nominalizer yang, which focuses on an argument (e.g. yang makan ‘the one that is eating’; yang di-makan ‘the thing that is being eaten’), nominalizer -nya is able to focus on the predicate (e.g. makan-nya ‘his eating’/* the eater’/* the thing eaten’). Nominalizer -nya thus provides us with a means to comment on events and situations as abstract nominals. As Englebretson (2003:168) puts it: “Because of its close association with nominal expressions, when it [i.e. -nya] is affixed to a lexical item from a different class, this item takes on the characteristics of a noun.”

Also noted in Englebretson (2003:170-171), enclitic -nya often combines with various lexical items to form adverbials such as biasanya (‘normally, usually’), as shown in (12) above. Other examples include akhirnya ‘finally’ (< akhir ‘end’), agaknya ‘presumably, perhaps’ (< agak ‘guess’), and silapnya ‘unfortunately’ (< silap ‘mistake’). Note that these adverbial -nya expressions often reflect speaker moods as well, particularly when found in clause-initial position, as seen in (13), (14) and (15), where enclitic -nya can readily serve as a convenient landing site for prosodic embellishments—such as vowel lengthening and emphatic rise-fall pitch contour—to reflect various shades of speaker’s emotive expressions, ranging from surprise or disbelief, to regret or resignation, or even annoyance or disgust.

(13) Akhirnya:: kita kalah.
end-NYA 1PL lose
‘In the end, we lost.’
(14) *Agaknya:... dia tak suka engkau.*
   guess-NYA 3SG NEG like 2SG
   ‘Perhaps, s/he doesn’t like you.’

(15) *Silapnya:... aku lupa beritahu mereka.*
   mistake-NYA 1SG forgot inform 3PL
   ‘Unfortunately/Regrettably, I forgot to mention it to them.’

Worth noting is that we can still retrieve a nominal interpretation in many such constructions—for example, ‘the usual practice is, he comes early’ in (12); ‘the final analysis is, we lost’ in (13); ‘my guess is, s/he doesn’t like you’ in (14); ‘the mistake was, I forgot to inform him’ in (15). Here I suggest that absence of the copula, which is highly typical for many Austronesian languages (Pustet 2003), may have facilitated a reanalysis of *-nya* nominal expressions into *-nya* adverbial ones. That is, we obtain reinterpretations such as follows: ‘the usual practice (is), he comes early’ > ‘usually, he comes early’. Similar reinterpretations obtain for the other examples as well.9 As noted in Fischer (2007), sentential adverbs are often well-suited to express speaker mood functions. It is not surprising therefore that we find enclitic *-nya* heavily involved in the marking of speaker stance, arguably via a nominalizer > adverbializer > stance marker grammaticalization pathway.

The use of *-nya* as stance marker is in fact very productive, and Englebretson’s analysis of colloquial Indonesian (2003:172-185) reveals at least three major stance functions associated with *-nya*. These include mental/emotional attitude, as in (16),10 speaker assessment of interactional relevance to ongoing discourse, as in (17),11 and marker of evidentiality, as in (18).12

(16) *Bagusnya mereka datang.*
   good-STANCE 3PL come
   ‘It’s so good that they came/are coming/will come.’

---

9 *Agaknya* in (14) is interesting because we see here an extension of *-nya* from third person reference to first person reference (‘my guess’), and from thence to such readings as ‘I suppose’ and ‘maybe, perhaps’. Extensions from third person to first person pronominal usage was reported in Englebretson (2003), and is attested in other languages as well (e.g. Classical Chinese pronoun *zhi*).

10 See also (15) above.

11 See also (16) above.

12 Note that while it’s still possible to retrieve a nominalization interpretation for (17) and (18)—e.g. ‘the good thing is, he came’ (> ‘it’s good that he came’) and ‘my meaning is, we don’t have to worry at all’ (> ‘meaning, we don’t have to worry at all’) —such a link is not available for (19), indicating that enclitic *-nya* has grammaticalized extensively as a stance marker.
(17) *Maksudnya kita tak perlu risau lansung.*
  meaning-STANCE 1PL:INCL NEG need worry completely
  ‘Meaning, we don’t have to worry at all.’

(18) *Nampaknya enak juga.*
  see-STANCE delicious also
  ‘It looks quite delicious, too.’ (Pragmatically: ‘I must say/agree/admit’)

Figure 2 below (based largely on Englebretson 2003:157-186) summarizes the grammaticalization pathways of enclitic *-nya*, highlighting in particular deictic and definiteness functions within the pronominal domain and extensions into speaker mood and stance marking functions via the nominalizer and adverbial pathways.\(^{13}\)

---

![Diagram of grammaticalization pathways of Malay nominalizer *-nya* (enclitic)](https://example.com/diagram.png)

**Figure 2.** Grammaticalization pathways of Malay nominalizer *-nya* (enclitic)

Compared to the grammatical functions of *yang*, those of *-nya* show clear evidence of pronominal usage, and massive extensions into adverbial and stance functions. Indeed, in his corpus analysis of colloquial Indonesian, Englebretson (2003) notes that a significant one-third of *-nya* constructions convey speaker (inter)subjective interpretations. Here I argue

\(^{13}\) Englebretson (2003:8) identifies pronominal uses of *-nya* as a relatively late development. Here I posit a very early stage for *-nya* along the grammaticalization cline because pronominal *-nya* is clearly referential.

\(^{14}\) See (8) for an example of enclitic *-nya* used in cleft constructions.
that the head-final characteristic of enclitic –nya readily facilitates the hosting and encoding of focus and stance functions, a position I continue to pursue in our discussion of head-final uses of punya in the following section (see also Xu 2004, for discussions of sentence-final position being a natural site for in-situ focus realizations).

4. Nominalizer punya: a case of language contact

The use of punya is now largely restricted to colloquial Malay. Classical Malay texts show frequent use of empunya as a lexical noun meaning ‘master, owner’, as in (19), and eventually increasingly as a lexical verb mempunyai, often accompanied by affixes mem-...-i to mark voice and transitivity features, as in (20). In the colloquial register, punya is extensively used as a possessive marker and possessive pronoun, as in (21) and (22) respectively, as well as an epistemic, attitudinal and interactional marker (i.e. speaker mood or stance marker), as in (23).

(19) Dia empunya kebun ini.
3SG owner orchad this ‘S/he’s the owner of this orchard.’

(20) Mereka mem-punya-i ilmu ajaib
3PL PREF-possess-SUF knowledge magic ‘They possess magical powers.’

(21) Jaga-jaga kau punya barang
look.after 2SG GEN stuff ‘Look after your stuff.’

(22) Jangan sentuh aku punya.
do.not touch 1SG GEN ‘Don’t touch mine.’

(23) Ibu (pasti) akan marah punya/mya.
mother certainly FUT be.angry STANCE ‘Mother will (surely) get upset (I can assure you).’

As noted in the literature (e.g. Gil 1999; Yap, Matthews & Horie 2004; Yap 2007), without the use of punya/mya, utterances such as (24) would simply be a factual statement. The
addition of sentence-final punya/mya, however, makes it a ‘trust me’ or ‘I’m telling you’-type assertion. That is, the intersubjective (or interactional) overtone becomes evident when sentence-final punya/mya is added.

Punya also produces an intensifying effect in pre-adjectival position, as shown in (24), where it expresses assertions or exclamations that are laced with strong feelings, including feelings of surprise, incredulity, and/or even annoyance.

(24) **Punya** jauh rumah kau ’ni, Timah.
STANCE far house 2SG this (name of person: female)
Lit. ‘So far, this house of yours, Timah.’
Intended meaning: ‘This house of yours is incredibly far, Timah.’

As discussed in Yap (2007), pre-adjectival uses of stance punya appear to have emerged as an extension of its genitive and associative linking functions, in particular via [begitu + punya + Adjective] constructions, as seen in (25a) and (25b) below. Crucially, the emphatic prosody accompanying the deictic adverb begitu ‘like that’ spreads across linking particle punya and extends over the descriptive adjective jauh ‘far’. When the deictic adverb begitu is elided, the intensifying prosody persists, and linking particle punya readily gets reinterpreted as an intensifier-type stance marker. This intensifier usage of punya is highly productive in colloquial speech.

(25) a. **Begitu** punya jauh::: rumah kau ’ni, Timah.
like that LNK far house 2SG this (name of person: female)
‘So::: far, this house of yours Timah!’

b. **Punya** jauh::: rumah kau ’ni, Timah!
STANCE far house 2SG this (name of person: female)
‘So::: far, this house of yours Timah!’

As reported in Yap, Matthews and Horie (2004), diachronic evidence reveals a gradual expansion in the grammatical functions of empunya as follows: lexical noun/verb > genitive/possessive pronominal > stance marker. As discussed above, the final stage involves the extended use of phonologically reduced punya as stance marker in both sentence-final and pre-adjectival positions. As a sentence-final particle, punya is often further reduced to mya (or mia). Gil (1999) further reports the use of pun and nya in other Malay varieties as well.

The development from nominal to pronominal to pragmatic marker (sentence-final particle) is well attested crosslinguistically, and often involves an intermediate stage involving the extended use of versatile nominalizers (see Yap & Matthews in press for a
recent summary; see also Noonan 1997, in press, this volume for Bodic and Tamangic languages; Simpson in press for Burmese; Rhee in press & this volume for Korean). Nominalizer *punya* is attested in colloquial Malay, as in (26), but not in the standard variety. The use of *punya* as a relative clause marker, as seen in (27), is marginal in colloquial Malay, and is more typical of Bazaar (or market variety) Malay, pointing to the possible influence of language contact, in particular with dialects of Chinese origin.

(26) *(Yang)* nak jahit *punya* letak 'kat sini.
NOMZ want sew PRON:ones put LOC(<‘near’) here
‘The ones you want to sew (or mend), put (them) here.’

(27) *(Yang)* nak jahit *punya* baju letak 'kat sini.
NOMZ want sew REL clothes put LOC(<‘near’) here
‘The clothes that you want to sew (or mend), put (them) here.’

Diachronic evidence in fact implicates an important role for language contact in the rise of nominalizer and relative clause marker *punya* (Yap et al. 2004). Corpus analysis of classical Malay texts reveals that genitive and possessive pronominal uses of *punya* rose sharply in the 19th century, coinciding with a massive influx of immigrant workers from southern China. The rise of these constructions may have facilitated the rise of nominalizer and relative clause uses of *punya*, given that such constructions were not attested in Classical texts but are evident in colloquial and Bazaar Malay. Arguably, these constructions in turn facilitated the rise of sentence-final stance uses of *punya* (and mya/mia); this stance marker usage is still in evidence in contemporary colloquial Malay. The examples in (28a-b) illustrate how nominalizer *punya* can easily be reanalyzed as a sentence-final particle (i.e. stance marker *punya*) via elision of the subject and verb in the matrix clause.

(28) a. *Aku pasti [dia akan datang *punya]*.
1SG certain 3SG FUT come NOMZ
‘I’m sure that he will come.’

---

15 It should be noted that native speakers often prefer to add nominalizer *yang* in constructions such as (26), giving rise to ‘double nominalization’ phenomena. Interestingly, non-native speakers using the Bazaar variety tend not to use nominalizer *yang*.

16 It is possible that a similar development has taken place at an even earlier date, particularly in places where Chinese traders have made contact with local Malays along the coast, not only on the Malayan peninsula but among the many islands of Indonesia as well, such as Ambon for example. Inclusion of Ambonese and other Malay texts in future analysis will help provide a fuller picture.
b. Dia akan datang punya.
3SG FUT come SFP
‘He will come (I assure you).’

Structural parallels between Malay punya and Cantonese ge3 constructions, the latter representing the Chinese dialects, are highlighted in (29) below. Among them are the genitive, possessive pronominal, relative clause, nominalizer, and stance usage respectively. With particular reference to (29c), note that nominalizers punya (Malay) and ge3 (Cantonese) retains the assertive force of the matrix clause (‘I know’) even when this higher clause is elided. Such force is encoded prosodically with the sentence-final nominalizers as host, which facilitates their reanalysis as stance markers.

(29) Cantonese:                  Colloquial Malay:
  a. nilgo3 (hai6) nei5 ge3 co3 this:CL COP 2SG GEN mistake this 2SG GEN doing
     ‘This is your mistake.’ ‘This is your doing!’

  b. nilgo3 (hai6) nei5 ge3 this:CL COP 2SG GEN
     ‘This is yours.’ this 2SG GEN
     ‘This is yours.’

  c. waam4 faam1 NGO5 maa5 ge3 je5 return:back 1SG buy REL stuff
     ‘Give me back the stuff that I bought.’
     ‘Give me back the stuff that I bought.’

  d. waam4 faam1 NGO5 maa5 ge3 return:back 1SG buy NOMZ
     ‘Give me back the one I bought.’
     ‘Give me back the one I bought.’

  e. (NGO5 zil) nei5 seong2 lai4 ge3 1SG know 2SG want come SFP
     ‘I know you want to come.’ (assertive)
     ‘I know you want to come.’ (assertive)

Chronologically, given that prenominal modification was attested even in Old and Middle Chinese (via zhi and she constructions, among others), and also given that prenominal punya constructions were a much more recent development in Malay, there is strong reason to believe that the southern Chinese dialects were the donor languages. Typologically, too, it has been shown that prenominal modification is extremely rare among SVO languages.
(Greenberg 1966; Dryer 1992), with the Chinese dialects making heavy use of topicalization strategies to mitigate the often high processing costs of using prenominal modification constructions within an SVO configuration (Matthews & Yeung 2001; Kwan 2005). Indeed, Malay, an SVO language, has long relied on postnominal yang constructions (see §2). All things being equal, the emergence of prenominal punya would have been largely unmotivated, particularly in the case of the lengthier relative clauses, had it not been for a sociological development with sufficient force to overcome the high cognitive processing constraints. This sociological force, I argue, came in the form of contact with Chinese dialects on a massive scale.

Figure 3 summarizes the grammaticalization pathways for Malay punya. Not discussed earlier but included for the sake of comprehensiveness is the emergence of subordinator punya pasal, via the merger of genitive punya and a general head noun pasal (used with a wide range of meanings including ‘reason, matter, affair, business, problem, fault, etc.’), as shown in (30) below.

(30) a. *Semua ni dia punya pasal.*
   all this 3SG GEN fault
   ‘All this is his fault.’

   b. *(Sebab) dulu takut gemuk punya pasal,  
      because previously fear fat GEN reason
      sekarang tinggal tulang ‘aja.
      now remain bone only
   ‘Because previously (she) was afraid of being fat, she’s now all bones.’

---

**Lexical verb empunya**

**Lexical noun** → **Genitive (em)punya** → **Cause/Reason subordinator punya pasal**

**empunya** → **Possessive pronominal punya**

**Relative clause marker punya**

(headless relative clause & Bazaar Malay)

**Headless relative clause context**

**Nominalizer punya** → **Stance punya/nya**

---

**Figure 3.** Grammaticalization pathways of Malay empunya

5. **Semantic niches of nominalizers yang, -nya and (em)punya**
In this section, we focus on the similarities and differences between the three nominalizers: *yang*, *-nya* and *(em)punya*. Table 1 displays their functional distribution and highlights their structure in terms of head-initial vs. head-final configuration.

<table>
<thead>
<tr>
<th>Function</th>
<th><em>yang</em></th>
<th><em>-nya</em></th>
<th><em>empunya/punya/mya</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun/Verb</td>
<td>Head-initial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genitive</td>
<td>‘Head-medial’</td>
<td></td>
<td>(see footnote 17)</td>
</tr>
<tr>
<td>Possessive Pronominal</td>
<td>Head-final enclitic</td>
<td>Head-final enclitic</td>
<td>‘Head-medial’</td>
</tr>
<tr>
<td></td>
<td>(possessor-referring)</td>
<td>(possessee-referring)</td>
<td></td>
</tr>
<tr>
<td>Definiteness marker</td>
<td>Head-initial</td>
<td>Head-final enclitic</td>
<td>(low focus)</td>
</tr>
<tr>
<td></td>
<td>(high focus)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Nominalizer/ Gerundive</td>
<td></td>
<td></td>
<td>Head-final enclitic</td>
</tr>
<tr>
<td>e.g. <em>makan</em> <em>nya</em> ‘his eating (habits)’</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative Clause marker</td>
<td>Head-initial</td>
<td>Head-final (mixed)</td>
<td><em>yang .... punya N</em></td>
</tr>
<tr>
<td>Nominalizer</td>
<td>Head-initial</td>
<td>Head-final (mixed)</td>
<td>*(yang) .... punya (<em>N)</em></td>
</tr>
<tr>
<td>Complementizer</td>
<td>Head-initial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinator</td>
<td></td>
<td>Head-final (+head noun)</td>
<td><em>punya pasal</em></td>
</tr>
<tr>
<td>Epistemic/Stance marker</td>
<td>Head-final enclitic</td>
<td>Head-final</td>
<td>(adverbial complex)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(sentence-final particle)</td>
<td></td>
</tr>
</tbody>
</table>

Overall, *yang* and *-nya* show very little functional overlap, each morpheme having found their own semantic niches over the course of time. Where they sometimes overlap, i.e. in contexts where they function as identifiability (or definiteness) marker, *yang* is generally a stronger focus marker, often used for contrastive identification, as noted earlier in §2 and illustrated again in (31) below, while *-nya* is less emphatic, and typically appears in contexts where the identity of the referent is already given or known, either in prior discourse or from shared knowledge, as in (32).

(31) *Aku nak yang biru, (bukan yang merah).*
1SG want DEF blue NEG DEF red
‘I want the red one, (not the blue one).’
(32) *Aku kena bayar sekarang, tapi *duinya tak cukup.*

ISG must pay now but money-DEF NEG sufficient

Lit. ‘I have to pay now, but the money is insufficient.’

Intended meaning: ‘I have to pay now, but I don’t have enough money (with me).’

As seen from Table 1, structural constraints contribute to the functional differences between *yang* and *-nya*. Consistently, *yang* appears in head-initial position (e.g. *[yang tumbang ke arah selatan]* ‘the one(s) that fell toward the south’), while *-nya* appears in enclitic position—that is, at the right edge of a phrase boundary, often yielding different nominalization effects (e.g. *[tumbang-nya] ke arah selatan* ‘the falling (inclusive of the way/direction it fell) was toward the south’). Crucially in both cases, a predicative gets treated as an argument (i.e. nominal), but the former type of nominalization focuses on the external argument (*the one who VPs*), while the latter type focuses on the event or situation that the predicate describes (*the VP-ing itself*).

Given their morphosyntactic differences, it is not surprising that we find different types of constraints on the uses of *yang* and *-nya*. For example, *yang* appears in head-initial relative clause and factive complementizer constructions (e.g. *ikan yang dijualnya* ‘the fish that he sells’; *aku tahu (yang) dia kalah* ‘I know (that) he lost’), while *-nya* is not found in these environments. On the other hand, enclitic *-nya* readily gives rise to epistemic uses (e.g. *Mahalinya ikan yang dia jual* lit. ‘So expensive, the fish that he sells’ or ‘The fish that he sells is so expensive!’).

Whereas *yang* and *-nya* tend to confine themselves to mutually exclusive functional domains, there is some overlap between the grammatical uses of *empunya* with those of *yang* and *-nya*. The lexical uses of *empunya* (as noun as well as verb) are still evident; particularly interesting for our present purpose is the shift from head-initial to ‘head-medial’ to head-final position for *(em)punya* as it grammaticalizes from lexical noun to more grammatical functions such as genitive, possessive pronominal, nominalizer and stance marker. This shift came at some cost in terms of processing efficiency, particularly in the case of relative clause marker *punya*, which emerged from language contact situations and for a period of time survived as part of Bazaar (or market-place variety) Malay. Head-medial genitive *punya* had to compete with the canonical [NP-possessee + NP-possessor] construction in Malay. However, once available, possessive pronominal *punya* basically came for free, and this paved the way for other head-final *punya* constructions.

Also worth noting is that *punya* has largely succeeded in avoiding competition with enclitic *-nya* in head-final environments. When used in possessive pronominal contexts,

---

17 In formal syntax, a more appropriate term would be ‘head-initial *punya* with the possessor noun in specifier position and the possessum in complement position’. Here we use the term ‘head-medial’ simply in linear terms and as a convenient label to highlight different structures in stage-like fashion.
enclitic -nya points to the possessor, while head-final punya points to the elided possessee. Greater functional overlap is found in the epistemic domain, and it is worth noting that this corresponds to the later stages of grammaticalization. Nevertheless, even at this late stage, there still are subtle differences: epistemic -nya is more explicit, since it expresses the semantic values of its lexical host (e.g. agaknya ‘probably, presumably’; from agak ‘guess’), while epistemic punya is much more subtle, given its origin in ‘headless’ environments, though it is no less impressive in its ability to express fine shades of speaker mood given its enviable access to sentence-final prosodic cues. In short, the Malay language welcomes both types of epistemic markers since they provide native speakers with a choice of explicit vs. implicit ways to express their subjective and intersubjective moods (e.g. Nampaknya dia tak akan datang ‘Looks like he’s not coming’ and Dia tak akan datang punya ‘He’s not coming, (believe me)’, sometimes with the socio-interactive implication ‘So let’s not wait any longer’). Competition among the Malay nominalizers is therefore not necessarily unduly eliminative; rather, it often constitutes a subtle negotiation of semantic and functional turf and their boundaries.

There is also evidence of facilitation. Recall that the reanalysis from possessive pronounal punya to nominalizer punya requires considerable facilitation from nominalizer yang, yielding yang...punya ‘double nominalization’ constructions. In other words, we obtain a nominalization that is marked both head-initially and head-finally, a situation which allows one of them—in this case, punya—to be reanalyzed as a sentence-final stance marker. In this regard, it is interesting to see how the Malay language bows, not only to pressure from language contact to replicate a head-final nominalizer along the lines of the southern Chinese dialects, but also to pressure from within the language to rely on a canonical head-initial nominalizer.18

6. Conclusion

In this paper we have examined three nominalizers in Malay, namely yang, -nya and punya, in terms of their functions, structural differences, and to the extent possible, their diachronic development as well. The analysis presented here highlights some robust grammaticalization patterns that are consistent with typological observations reported elsewhere, among these being a nominal/pronominal > nominalizer > stance marker development. The notion of stance or speaker mood discussed in this paper is a broad one, and has included emphatic and focus constructions (particularly with respect to yang constructions), and epistemic, attitudinal and interactional/intersubjective markers (in the case of enclitic -nya and sentence-final punya constructions).

---
18 Double marking (or redundancy) is not uncommon in bilingual code-mixing contexts (see Chan 1998).
Intensifier stance constructions, often with exclamative effect, also emerged in the case of punya constructions, although via adnominal (i.e. genitive/associative/relative clause) rather than nominalizer extension. This intensifier-type punya construction should prove interesting for future studies on the typology of non-nominalizer type exlamatives, which include among others the ‘scalar’ or ‘intensifier’ demonstrative constructions like those involving English that (e.g. he’s that crazy, I can’t believe it), Classical Chinese zhi (e.g. Han zhi guang ye ‘the Han River so broad!’ (< ‘the Han River that broad’; see Yue 1998:265), and Czech tak (e.g. tak velký ‘so big’ (exclamative reading); see Landman & Morzycki 2003; cited in Sio & Tang 2007:72). As noted earlier, the intensifier usage of punya can be traced back to the elision of a deictic adverb with scalar values, namely begitu ‘like that, to that extent’ (e.g. begitu punya nakal lit. ‘like that punya mischievous’ > punya nakal ‘so mischievous’), leaving genitive linker punya to assume the pre-adjectival intensifier function, and thus essentially behave much like English that, Classical Chinese zhi and Czech tak in exclamative expressions. Further investigations into these intensifier-type exlamatives will contribute to our understanding of the distinctions between nominalizer vs. non-nominalizer type stance constructions.

In this paper, we have not yet delved into the contribution of the copula (if any) to the development of nominalizers into stance markers. While Austronesian languages generally make minimal use of copulas, with some languages such as Tagalog showing no recognizable copula form (Pustet 2003:39), Malay does have two copula forms: adalah and ialah. The former comprises of the existential verb ada (‘be, have’) plus focus particle lah, and the latter is formed from the combination of third person pronoun ia (‘s/he, it’) plus focus particle lah. One wonders, of course, if focus particle lah might not have once been an erstwhile copula, particularly given that copulas are known to sometimes grammaticalize further into focus markers (e.g. Mandarin shì, Chaoshou si and kai, and potentially Classical Chinese ye; see Xu & Matthews, this volume; Pulleyblank 1995; Heine & Kuteva 2003:95-96). To express speaker stance (e.g. emphasis, focus, assertion), lah appears with clauses nominalised by yang to form cleft constructions such as Dia lah [yang nak nikah esok] ‘She’s the one that’s

19 The term ‘adnominal’ refers to a linker between a modifying expression and its head noun, and is used extensively in linguistic descriptions related to languages with prenominal modification, which are typically verb-final languages such as Japanese, Korean, and Tibeto-Burman languages, except of course the rare case of Sinitic languages among the SVO type. Previous studies have identified general nouns (sometimes referred to as formal nouns) meaning ‘person’, ‘thing’ or ‘place/locaton’ and demonstratives (e.g. Chinese zhi) as sources of adnominals and nominalizers (e.g. Delancey 1986; Horie 2000; Rhee in press; LaPolla 1994; Yap, Choi & Cheung 2007).

20 There is debate on whether Classical Chinese ye has copula functions. Its status as a focus marker is non-controversial; likewise, its status as topic marker and sentence-final particle (e.g. Yue 2004). Less clear is whether its focus function may have had a copula origin.
getting married tomorrow’, as seen in (5) in §2 earlier. However, lah does not appear in stance –nya and stance punya constructions (e.g. [Nampaknya] (*lah) nak hujan ‘ni ‘It looks like it’s going to rain’; Esok (*lah) tok hujan punya ‘Tomorrow won’t rain, trust me’). Where lah appears in the presence of –nya and punya, we obtain referential rather than expressive or stance readings. For example, ini lah tempat-nya ‘This is his place/ This is the place’, yields a possessive pronominal or definiteness marker usage of enclitic –nya. Likewise, ini lah dia punya ‘This is his (one)’ yields a pronominal one interpretation for punya. More elaborate constructions such as Ini lah yang dia beli punya ‘This is the one that he bought’ also yields a referential use of the ‘double nominalizer’ yang...punya. Essentially, where they co-occur, it is lah that independently conveys expressive functions such as speaker’s emphatic stance, while –nya and punya participate in referential functions such as pronominal and determiner/definiteness marking related to referent identification. Interestingly, the copula forms adalah and ialah cannot readily replace lah in these constructions.

One reason has to do with register. Recall that grammatical uses of punya tend to be restricted to the colloquial register, while the copula forms adalah and ialah are more commonly seen in formal writing. Another reason has to do with the status of these copulas. At least at this stage in time, they are primarily used as linkers in equative constructions, for identificational and specification purposes, hence strongly tied to referential rather than expressive functions. This would explain why, even though focus marker lah is frequently used to produce cleft constructions, the copulas adalah and ialah do not participate in these constructions (e.g. Dia lah/*adalah/*ialah yang curi mutiara aku ‘He’s the one who stole my pearl’).

Let us summarize what we have thus far. We see that focus marker lah can co-occur with nominalizer yang, and in the absence of lah, as discussed earlier in §2, focus effects persist as prosodic stress typically associated with lah can readily be attached to nominalizer yang (e.g. Dia (lah) yang patut kau awasi ‘He’s the one you need to beware of’). Focus marker lah cannot co-occur with enclitic –nya nor sentence-final punya to yield a stance construction. Instead, stance reading comes via other sources. In the case of enclitic –nya, we see a link to nominalizer and adverbial origins (see §3). The stance effect ranges from epistemic to attitudinal to interactional or intersubjective (e.g. Rasanya, tak patut kita buat ‘giu ‘My gut feeling is, (we) shouldn’t do it that way’). In the case of sentence-final punya, we see a potential reanalysis of the nominalizer as sentence-final particle when the higher matrix clause involving a mental or perceptual verb is elided (e.g. Aku tahu dia akan jual (punya) >

21 Adalah and ialah appear to be of relatively recent origin, arguably due to contact with colonizing languages such as English and Dutch.

22 Emphasis, where necessary, is done through prosody, using stress and vowel lengthening, for example, which usually falls on the constituent preceding the copula.

23 Interestingly, this constraint applies to copulas of verbal origin (adalah) as well as pronominal origin (ialah).
Dia akan jual punya ‘I know he will sell’ > ‘He will sell, I know it’) (see §4). The stance
effect is typically assertive.

A tentative conclusion that emerges from these observations is that copulas (such as
adalah and ialah) are not designed for stance work, but focus markers (such as lah) are. And
if it turns out that focus marker lah was once a copula, a tempting question to posit is whether
copulas that engage in stance work in other languages have in fact been reanalyzed as focus
morphemes as well. Such a view is not entirely new, given that considerable attention has
been given to the reanalysis of Mandarin shi from pronominal to copula to focus marker, the
latter in shi...de constructions. What is worth further investigating is whether copula
morphemes have likewise been reanalyzed as focus morphemes in sentence-final
combinations such as Japanese –n(o) desu and Korean –n kesia, as well as Chantyal –wa hin
among the Tibeto-Burman languages. If the answer to this is yes, we would be able to have a
common analysis for copula-dependent nominalizations in both verb-final and non-verb-final
languages.

While more research is needed for a fuller understanding of how copulas evolve and
contribute to stance constructions, the present paper has shown how three different
nominalizers in Malay give rise to a wide range of constructions that capture various shades
of speaker mood. Their different stance effects can be traced back to differences in
etymological origin and preferred structural configuration, and in the case of punya can be
attributed to language contact as well. Their development from referential to expressive
marker is consistent with a general tendency for linguistic items to often grammaticalize into
pragmatic markers (e.g. Traugott 1982, 1991, 1995), and thus points to strong cognitive
motivations underlying a process of semantic extension to meet pragmatic needs.

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASSOC</td>
<td>Associative</td>
</tr>
<tr>
<td>CL</td>
<td>Classifier</td>
</tr>
<tr>
<td>COMP</td>
<td>Complementizer</td>
</tr>
<tr>
<td>DEF</td>
<td>Definite</td>
</tr>
<tr>
<td>DEM</td>
<td>Demonstrative</td>
</tr>
<tr>
<td>DIST</td>
<td>Distal</td>
</tr>
<tr>
<td>FOC</td>
<td>Focus</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative</td>
</tr>
<tr>
<td>NEG</td>
<td>Negative</td>
</tr>
<tr>
<td>NOMZ</td>
<td>Nominalizer</td>
</tr>
<tr>
<td>PN</td>
<td>Pronoun/Pronominal</td>
</tr>
<tr>
<td>PROX</td>
<td>Proximal</td>
</tr>
<tr>
<td>REL</td>
<td>Relativizer</td>
</tr>
<tr>
<td>SFP</td>
<td>Sentence Final Particle</td>
</tr>
<tr>
<td>SG</td>
<td>Singular</td>
</tr>
<tr>
<td>TOP</td>
<td>Topic marker</td>
</tr>
</tbody>
</table>
References


Chan, Hok-shing Brian. 1998. How does Cantonese-English code-mixing work? In M. Pennington (ed.), *Language in Hong Kong at century’s end.* Hong Kong: Hong Kong University Press.


Noonan, Michael. (this volume). Aspects of the historical development of nominalizers in the Tamangic languages.


Rhee, Seongha. (this volume). Nominalization and stance-marking in Korean.


Yap, Foong Ha. 2007. On Native and Contact Grammaticalization: The Case of Malay empunya. Manuscript, Department of Linguistics and Modern Languages, Chinese University of Hong Kong.
Yap, Foong Ha, Pik-ling Choi & Kam-siu Cheung. 2007. Delexicalizing di: How a Chinese noun evolves into a nominalizer, with attitude. Manuscript, Department of Linguistics and Modern Languages, Chinese University of Hong Kong.


