Asu vs. Tasu:
On the origins of Dusunic moveable t-

Paul R. Kroeger
Stanford University/SIL
14-May-90
(pre-publication version)

§1. Overview

The Dusunic languages are a group of Philippine-type languages spoken in the
Malaysian state of Sabah, on northern Borneo. A curious fact about many Dusunic
languages is that, whereas verbs and adjectives may begin with virtually any phoneme,
there are no nouns which begin with vowels. Moreover, vowel-initial nouns in those
Dusunic languages which allow them always have cognate forms beginning with t- in the
languages with no vowel-initial nouns, e.g.:

(1)  Rungus   Kimaragang
     asu      tasu       dog
     onsi     tonsi      meat
     anak     tanak      child
     ulun     tulun      person
     ulu      tulu       head
     apuy     tapuy      fire
     avan     tawan      sky

The presence vs. absence of initial t- in such forms has long been used as a rough
typological parameter for classifying Dusunic languages and dialects. For example
Clayre (1966), in discussing the phonological contrasts among Dusunic dialects, states:

“Another difference is the tendency for R[ungus], B[engkoka]¹ and sometimes
L[otud] to drop the initial t of roots.”

It is clear from wider comparative evidence that the direction of change was the
opposite -- the initial t- is an innovation, of which no trace is found in cognate forms in

¹Clayre used the term Bengkoka for the dialect of Rungus spoken in the Pitas district, on the Bengkoka
peninsula. King (1984a) reports that the Bengkoka dialect is 90% cognate with Rungus proper.
Philippine languages or even in the closely related Murutic and Paitanic languages of Sabah. On the other hand, not all initial t-’s exhibit this type of variable distribution. Many noun stems begin with t- in all Dusunic languages, as well as in other more distantly related languages:

(2) | **Rungus** | **Kimaragang** | **Malay** | **English** |
--- | --- | --- | --- | --- |
  tali | tali | tali | rope |
  ta’ap | ta’ap | atap | roofing |
  tambang | tambang | --- | deer |
  tindal | tindal | --- | dry land |
  to’un | to’on | tahun | year |

Our concern here is with the “moveable t-” illustrated in (1). Where did this t- come from? If it represents an innovation, as it must, what motivated this development and why did it fail to occur in some Dusunic languages?

The answer proposed in this paper is that the moveable t- represents the last vestiges of an old definite article -- a “stage III article”, in Greenberg’s terms. I will begin by examining the distribution of t- more closely in two representative languages. In section 3, I will discuss the general development and properties of stage II and stage III articles as described by Greenberg (1978, 1981, 1989), citing a number of examples from other language families around the world. Finally, I will discuss specific possible sources of the t-.

§2. The distribution of t-

The superficial distinction referred to above between languages with or without the moveable t- was based on the presence vs. absence of t- in citation forms of nouns, i.e. in the forms likely to be elicited on a wordlist or listed in a dictionary. However, the actual distribution of the moveable t-, both within any particular language and across
languages, is more complex than can be ascertained merely by comparing wordlists.

In every language which has the \textit{t} in citation forms there are some environments where the \textit{t}-less form is required. Similarly, in some languages which lack the \textit{t} in citation forms of nouns, a \textit{t} nevertheless appears before vowel-initial nouns in many environments. In other languages, there is no trace of this intermittent \textit{t}. Thus the simple binary classification of Dusunic dialects referred to above actually masks a three-way distinction. I will refer languages which have the \textit{t} in citation forms, such as Kimaragang, as “stage III” languages, for reasons that will be made clear below. I will refer languages in which \textit{t} appears in some contexts but not in citation forms, such as Rungus, as “stage II” languages.\footnote{These labels are meant to be suggestive, rather than precise --- they may actually exaggerate the degree of difference between the two classes, as we will see shortly.}

I will refer to languages which show no trace whatsoever of the moveable \textit{t} as “\textit{t}-less” languages. Most languages of this type which have been classified as Dusunic appear to be peripheral, both genetically and geographically, with respect to the main body of the Dusunic family. The best studied example is Tatana, a member of the quite distinct Bisaya sub-group.\footnote{Bisaya combines with Dusunic to form what Ruhlen (1987) has labeled the Dusun-Bisaya group.} Another example is Dumpas, which is lexically intermediate between Dusunic and Paitanic. The apparent exception is Lotud, which is clearly Dusunic in its lexicon and phonology (and relatively central geographically). But, as we will see below, it may be quite misleading to refer to Lotud as a “\textit{t}-less” language.

\section*{§2.1 Moveable \textit{t} in Kimaragang (a stage III dialect)}

Since the “normal” (i.e. citation) form in Kimaragang always has \textit{t}, I will attempt to describe the environments in which the moveable \textit{t}- does not appear. Following Greenberg’s terminology, I will refer to the \textit{t}-initial form of the noun as the “articulated
form”, and the form without an initial t- as the “non-articulated form”.

(a) The moveable t- is not used in kinship terms of address or other vocative forms, though it is required in the corresponding terms of reference. Examples:

(3)

<table>
<thead>
<tr>
<th>Address</th>
<th>Reference</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>aki</td>
<td>taki</td>
<td>grandfather</td>
</tr>
<tr>
<td>odu</td>
<td>todu</td>
<td>grandmother</td>
</tr>
<tr>
<td>idi</td>
<td>tidi</td>
<td>mother</td>
</tr>
<tr>
<td>ama</td>
<td>tama</td>
<td>father</td>
</tr>
<tr>
<td>obpinee</td>
<td>tobpinee</td>
<td>sibling</td>
</tr>
<tr>
<td>ambaya</td>
<td>tam(ba)baya</td>
<td>friend</td>
</tr>
</tbody>
</table>

(b) Proper names are frequently vowel-initial, e.g. Uji, Ogidang, Ajah, Ibali (sometimes pronounced Tibal), etc. Non-articulated forms of common nouns may occur as proper names, even though the articulated forms are required in normal usage.

(c) The moveable t- never appears on noun roots in derived verbal forms. That is, verbal morphology requires the use of non-articulated forms. Examples:

(4)

<table>
<thead>
<tr>
<th>Citation</th>
<th>Gloss</th>
<th>Verbal form</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>tasu</td>
<td>dog</td>
<td>mag-asu</td>
<td>to hunt</td>
</tr>
<tr>
<td>taddau</td>
<td>sun</td>
<td>mag-addau</td>
<td>dry season</td>
</tr>
<tr>
<td>tanak</td>
<td>child</td>
<td>mag-anak</td>
<td>give birth</td>
</tr>
<tr>
<td>tuwa</td>
<td>fruit</td>
<td>mong-uwa</td>
<td>bear fruit</td>
</tr>
<tr>
<td>talud</td>
<td>canoe</td>
<td>moning-alud</td>
<td>build a canoe</td>
</tr>
<tr>
<td>tobpinee</td>
<td>sibling</td>
<td>mi-obpinee</td>
<td>to be siblings</td>
</tr>
<tr>
<td>tulun</td>
<td>person</td>
<td>mi-ulun</td>
<td>not related</td>
</tr>
<tr>
<td>tagung</td>
<td>gong</td>
<td>mi-agung</td>
<td>play the gongs</td>
</tr>
</tbody>
</table>

(d) Other kinds of derivational morphology also require non-articulated forms:

(5)

<table>
<thead>
<tr>
<th>taddau</th>
<th>sun</th>
<th>tangaddau</th>
<th>mid-day</th>
<th>sang-addau</th>
<th>one day</th>
</tr>
</thead>
<tbody>
<tr>
<td>tulun</td>
<td>person</td>
<td>tanga-ulun</td>
<td>middle-aged</td>
<td>song-ulun</td>
<td>one person</td>
</tr>
<tr>
<td>tinan</td>
<td>body</td>
<td></td>
<td></td>
<td>song-inan</td>
<td>one thing</td>
</tr>
</tbody>
</table>
(e) There are two existential markers in Kimaragang, the free form waro and the proclitic ki-. The articulated form always appears after waro, but after ki- the non-articulated form is more common.4

(6) waro tanak           ki-anak             to have children
    waro tulun sitih   ki-tulun sitih       there are people here
    ---                 ki-addau              to be sunny

(f) In lexicalized reduplicated forms, the moveable t- appears only word-initial and never at the internal word boundary:

(7) tasi-así             orphan
    talun-alun            road
    tuwit-uwit               bird species
    tundundu (< *tundu-undu)   heart

(g) The moveable t- is used productively as a nominalizing prefix. The -on in the following examples signals that the nominalized argument is the Patient or Location:

(8) uli return           t-uli-on          place one returns to
    odop sleep            t-odop-on         sleeping place
    akan eat              t-a-akan-on        food; cooked rice
    iim seek               t-iim-on           that which is sought
    a-ra’at bad            t-a-ra’at         that which is bad

(h) Stage III languages differ concerning the appearance of moveable t- on vowel-initial loan words. From the Malay word akal, originally borrowed from Arabic, Labuk Kadazan has akal corresponding to Kimaragang takal ‘cleverness; trick’. Similarly the Malay word emas ‘gold’ appears in Penampang Kadazan as amas and in Kimaragang as tamas5. Both of these loans seem fairly old and well-established; it would be interesting

4This is an area of variation among stage III languages. For example, Penampang Kadazan regularly uses articulated forms with ki-, as in ki-tanak (C. Miller, p.c.).
5This loan word seems to be displacing an older term for gold, bulawan, in most Dusunic languages.
to compare the productivity of moveable t- as a noun marker in very recent borrowings.

(i) Moveable t- also appears on adjectives used as modifiers within a noun phrase, but not on predicative adjectives. In relative clauses the moveable t- appears on the complementizer when it precedes a predicate adjective (all adjectives are vowel-initial, beginning with the stative prefix o-/a-):

(9) Agayo ilo kusay. big that man
     kusay t-agayo big man / large male (e.g. water buffalo)
     kusay do-t agayo man/male who is big

This is another area of variation among stage III languages. In Kimaragang relative clauses, as in the third line of example (9), the complementizer carries the moveable -t before predicate adjectives. However, in Labuk Kadazan, a stage III language fairly closely related to Kimaragang, no t occurs in these environments. Rungus, on the other hand, a stage II language, follows the Kimaragang pattern:

(10) Labuk Kadazan Rungus Gloss
tulun do otulid ulun dot otopot person who is true/honest

In Penampang Kadazan, another stage III language, the moveable t in such environments attaches to the adjective:

(11) Penampang Kadazan Kimaragang Gloss
tondu’ do t-avanus tongondu dot alanji woman who is beautiful
§2.2 Moveable \( \dot{t} \) in Rungus (a stage II dialect)

Forschner (1978) reports that the basic form of the nominative and accusative case markers in Rungus are as follows:

\[
\begin{array}{ccc}
\text{NOM} & \text{ACC} \\
\text{Definite}^6 & \text{it} & \text{dit} \\
\text{Indefinite} & \text{ot} & \text{dot}
\end{array}
\]

However, he notes that the final \( \dot{t} \) in these forms is optionally deleted, especially before consonants. Moreover, \( \dot{t} \) never occurs before proper names or vocative terms. A preliminary inspection of written Rungus folk-tales indicates that the alternation between \( \dot{i} \) and \( \text{it} \), \( \dot{di} \) and \( \text{dit} \) etc. is virtually categorical. The articulated forms (ending in \( \dot{t} \)) always occur before vowel-initial nouns (excluding vocatives and proper names) and almost never before consonant-initial nouns. Note the following example:

\[
\begin{array}{l}
... \text{ka di-}t \quad \text{id} \quad \text{di-}t \quad \text{ondu} \quad \text{dino.} \\
\quad \text{say the mother of woman that}
\end{array}
\]

In the standard Rungus orthography, this final \( \dot{t} \) is not restricted to case markers but can appear on a variety of other classes of words as well. It is written as the final element in demonstratives, numerals, predicates etc. when they occur before vowel-initial nouns. Some examples are listed below:

\footnote{Forschner refers to these clitics as “articles” rather than case markers; in fact, as discussed below, the two functions are combined in a portmanteau which indicates both definiteness and case. In emphasizing the case-marking function, I have replaced Forschner’s labels for them, i.e. “subject” vs. “object” forms, with the labels nominative and accusative. Forschner does not use the terms “definite” and “indefinite”, referring instead to “substantive” vs. “suppositioned” subject, and known or specific vs. general or collective object.}
(14) **Demonstratives:**

- mingkakat ino-t anak \( \rightarrow \) the child stood up
- stand.up that child
- iti-t ongo-tangon diti \( \rightarrow \) all these folk-tales
- this pl-folk.tale this

**Numerals:**

- Varo-po, ka, iso-t ondu \( \rightarrow \) there was once, they say, a woman
- exist-yet say one woman
- turu-t adau \( \rightarrow \) seven days
- seven day

**Particles:**

- sid koturu=nowt adau \( \rightarrow \) on the seventh day
- at seventh=already day
- iso po-t ulun \( \rightarrow \) one more person
- one yet person

However, George Appell (personal communication) informs me that this is a purely orthographic convention\(^7\), and that in examples like those in (14) the /t/ is invariably pronounced as the initial segment of the noun, rather than the final segment of the preceding word. If this is correct, then in all of the contexts exemplified in (14) the use of /t/ in Rungus and Kimaragang is virtually identical. However, I will continue to follow the standard orthographic convention in citing Rungus examples, since I am concerned here only with the presence of /t/ rather than with its status as onset vs. coda.

Vowel-initial nouns may occur in existential constructions without the moveable \( t \), as in (15a); but in the vast majority of such cases the moveable \( t \) appears, written on the existential predicate as in (15b):  

(15) §a. aso ama \( \rightarrow \) has no father
- not.exist father

\(^7\)Presumably this device is intended to preserve the citation form of the noun in all contexts.
§b. varo-t ombolog
exist bird

aso=po-t anak
not.exist=yet child

aso-t andaman=ku
not.exist remember=I

there was a bird

still childless

I do not remember

At first glance, it appears that the distribution of moveable t in Rungus is essentially the same as in Kimaragang. However, there are environments where moveable t would be required in Kimaragang but does not appear in Rungus. The most obvious is, of course, in citation forms (note the examples in (1) above). Predicate nominals always use the citation form, e.g.:

(16) **Rungus** | **Kimaragang** | **Gloss**
--- | --- | ---
ondu it anak | (ong)ondu it tanak | The child was a girl.

female the child

ulun oku | tulun okuh | I am human.
person I

Another context where the citation forms occur is when a noun is used in its generic sense, referring to the entire class of things it denotes. The following example is taken from a Rungus folk-tale:

(17) Asu ong asandang nga oputut-putut;
dog if carried.on.shoulder also break

onsung ong asandang nga opudung gama humago.
mortar if carried.on.shoulder also cut work hurry

‘If you sling a dog over your shoulder, the strap will keep breaking; if you sling a mortar over your shoulder, the strap will cut you.’ (???)
Again, the Kimaragang equivalents to the initial nouns in each clause would appear with the moveable \( t^- \) (\( \text{tasu} \) and \( \text{tosung} \)).

In Rungus, no \( t^- \) occurs after the locative/dative marker \( \text{sid} \). In Kimaragang and other stage III languages, however, \( t^- \)-initial (articulated) forms are used in this environment:

\[
\begin{align*}
\text{Rungus} & \quad \text{Kimaragang} & \quad \text{Gloss} \\
\text{sid alun-alun} & \quad \text{sid talun-alun} & \quad \text{on the road} \\
\text{sid isan} & \quad \text{sid tisan} \, \text{do tagad} & \quad \text{at the edge of the field} \\
\ldots \text{sid idi} \, \text{dit anak} & \quad \ldots \text{sid tidi} \, \text{dit tanak} & \quad \text{(spoke) to the mother of the child} \\
\ldots \text{sid apad} & \quad \ldots & \quad \text{(went out) to the porch} \\
\text{sid azup} \, \text{dino raja} & \quad \text{sid tayup} \ldots & \quad \text{in the trap of the king}
\end{align*}
\]

Another environment in which Kimaragang requires a \( t^- \) which does not appear in Rungus involves the plurality marker, \( \text{ong}^- \). When plural nouns are formed in Kimaragang, moveable \( t^- \) appears both on the nominal root which follows the prefix and on the plural form as a whole (i.e. before the prefix). In Rungus there is no \( t^- \) following the prefix. The moveable \( t^- \) is written on articles etc. which precede the plural form in the same environments where it would appear before simple (singular) vowel-initial nouns. Thus the significant difference here is the non-occurrence of \( t^- \) between the plural prefix and nominal root in Rungus:

\[
\begin{align*}
\text{Rungus} & \quad \text{Kimaragang} & \quad \text{Gloss} \\
\text{ongo-ulun} & \quad \text{tongo-tulun} & \quad \text{people} \\
\text{ongo-uripon} & \quad \text{tongo-turipon} & \quad \text{slaves} \\
\text{anga-anak} & \quad \text{tongo-tanak} & \quad \text{children}
\end{align*}
\]

\[8\text{The source of the first \( -n^- \) in Rungus \( \text{onsung} \) is not clear. Many Austronesian roots are reconstructed with an optional nasal infix, which may be a factor here.}\]

\[9\text{In Central Dusun (Bundu dialect) the corresponding locative marker is \( \text{id} \), as in \( \text{id tanak} \) \text{‘to the child’} \text{(not \( \text{id anak} \)). No such marker exists in Labuk Kadazan, so it is difficult to verify whether \( t^- \)-initial nouns would be used in locative expressions.}\]

\[10\text{\( \text{tagad} \) is a stem which has initial \( t^- \) in all Dusunic languages.}\]
Finally, no - kad appears in Rungus following the (archaic?) possessive marker, dang:

(20)  

<table>
<thead>
<tr>
<th>Rungus</th>
<th>Kimaragang</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>dang ulun</td>
<td>dang tulun</td>
<td>belonging to other people</td>
</tr>
</tbody>
</table>

§3. Life cycle of an article

Greenberg (1978) identifies three stages in the grammaticalization of articles. Stage I represents the definite article, e.g. the in English, which “identifies an element as previously mentioned in discourse.” In stage II, the article spreads to include non-definite but specific (i.e. referential) uses as well. Greenberg refers to this stage as the “non-generic” article, because it marks elements as being either previously mentioned or at least specific. More precisely: “We may define Stage II as the stage in which we have an article which includes, along with possibly other uses, both definite determination and non-definite specific uses.” (Greenberg 1978, p. 62)

The indefinite article a(n) in English can be ambiguous between the indefinite-specific and generic uses. For example, the sentence “John wants to marry an actress” could mean either that John is in love with a particular woman who happens to be an actress (the specific/referential reading) or that John hopes that whatever woman he marries will be an actress, but as yet has no particular actress in mind (the generic reading). Thus stage II articles combine the use of English the and the specific/referential use of English a(n).

In stage II, the articulated form of the noun (i.e. the combination of noun plus stage II article) has replaced the unarticulated form (the bare noun stem) as the norm. The articulated form usually appears as the citation form in wordlists etc. and appears much more frequently in text, e.g. 90% of the time in text counts. It becomes easier to define the article in terms of the environments where it is not used, and these environments fall
into two classes.

On the one hand, “inherently determined” elements such as proper names, vocatives, place names and nouns modified by a demonstrative or possessive pronoun are automatically definite. Such elements would not normally occur with a (stage I) definite article (e.g. *the Thomas Edison,11 *the Singapore, *the my hat, etc.), and do not normally acquire an article in stage II. On the other hand, generic uses of the noun also fail to take the stage II article. Greenberg classifies these generic uses into four main types: (1) negations; (2) predication; (3) adverbial and locative uses; (4) generic verb objects, in particular incorporated objects, and dependent genitives in compounds.

With respect to these patterns, we have seen that Kimaragang and Rungus both preserve unarticulated forms in certain inherently determined contexts, e.g. proper names and vocative forms. Both languages also preserve unarticulated forms in incorporated nouns and other morphologically bound forms, Rungus more consistently than Kimaragang. But whereas Rungus preserves the unarticulated forms in many other generic contexts, e.g. locatives, predicate nominals, topicalized generic nouns and some negated existentials, in Kimaragang the articulated form has been generalized more widely and now occurs in these contexts.

Stage II articles still function in at least some contexts to distinguish the specific and generic uses of a noun. In stage III, this function is lost and the former article becomes merely a fossil. It may generalize even further, occurring on every noun in virtually every context. In this case it may be reanalyzed as a noun-class or gender marker, or acquire a new grammatical function as a nominalizing affix (as in the Kimaragang examples in (8)). On the other hand, the article may disappear entirely,

11 Of course, some languages, e.g. Portuguese, use the definite article before proper names.
leaving only a few sporadic traces, either as segmental material or phonological alternations (e.g. tone or consonant mutation).

Stage III articles may preserve some features of the “non-generic” article from which they arrive. For example, in Hausa all nouns end in vowels, which are almost always long. Forms with short vowels occur in two main environments: proper names and adverbial expressions. The same word may take a long final vowel as a common noun but appear with a short vowel as a proper name, e.g. dawakii ‘horses’ vs. Dawaki ‘name of town’; wutaa ‘fire’ vs. Pa wuta ‘in the fire’. The final vowel length is a stage III article, derived from an old suffixed definite article.

In the Penutian languages of North America, most nouns end in -s; however, incorporated nouns never carry this -s. In some languages, the -s has become a nominalizing suffix, e.g. Wintun hara ‘to go’, hara-s ‘the act of going’. Similarly, there is an initial k- which appears on nouns in Nilo-Saharan languages. Its distribution is quite sporadic, and even closely related dialects may differ as to whether the k- is present in particular lexical items, e.g. ini vs. kini ‘skin’; umu vs. kumu ‘fly’ in different dialects of the Moru language. Some dialects may even have free variation, e.g. Keliko kari or ari ‘blood’. In the Sara language, this k- is now used as a nominalizer, as in usa ‘to eat’, k-usa ‘act of eating’. Greenberg suggests that this k- was already a stage III article in proto-Nilo-Saharan.

It seems that the moveable t- in Dusunic is now undergoing the transition from stage II to stage III article. In Rungus it still seems to be functioning as a stage II article, as can be seen especially in the use of non-articulated forms for locatives, predicate nominals and generic topics. In Kimaragang, although the t- fails to appear in some environments (especially in vocatives and incorporated nouns), one should probably
classify it as an early stage III article.

This suggests that in Proto-Dusunic there was a definite article which was the source of the moveable t- in the daughter languages. This article probably developed at some stage after the separation of the Dusunic and Bisaya subgroups, since there is no trace of a pre-nominal t- in Tatana. Similarly, Dumpas was classified by Smith (1980) as a Dusunic language, but King (1984b) argues that is genetically Paitanic rather than Dusunic. The fact that Dumpas lacks the moveable t-\(^{12}\) would seem to support King’s claim. But one cannot rule out the possibility that the Proto-Dusunic article was simply lost in some daughter language. Thus, while the presence of the moveable t- is powerful evidence in support of membership in the Dusunic family, lack of the moveable t- does not, in and of itself, rule out such membership.

\section*{§4. The Proto-Dusunic definite article}

Greenberg (1978, p. 61) states that “... the most common origin of the definite article is the demonstrative, a development of which there are numerous and well-attested examples ...”. Similarly, Reid (1978, p. 57) observes, “... the historical source of articles in many languages is from demonstratives ...”. Thus it is natural to look for a demonstrative source for the moveable t-.

The determiners in most, if not all, of the modern Dusunic languages are as follows:

\footnote{\(^{12}\)The only Dumpas data immediately available comes from Hurlbut’s 1976 wordlist. In addition to t-less citation forms, we find the phrase: mongintaaun do umo ‘burn off a field’, where stage II languages such as Rungus would say ... dot umo.}
It seems clear that the original demonstrative forms were *ti, *no, *lo’ and *ri.

To these the modern nominative and accusative case markers, i and di, were affixed to derive the modern demonstrative forms. Dusunic *ti may be cognate with the demonstrative tu ‘this’ found in Paitanic and some Murutic languages. Reid (1978) notes the presence of parallel /i/-grade and /u/-grade case markers in the Philippine languages.14

My hypothesis is that ProtowDusunic had a pre-posed definite article *ti, derived from (and homophonous with) the demonstrative meaning ‘this’. This definite article eroded to t- and became generalized to a stage II article, either before the divergence of the core Dusunic languages or independently in each. As the old article became grammaticalized, the function of marking definiteness was taken over by a contrast between /i/-forms and /o/-forms of the modern nominative and accusative case markers, which now serve a dual function as case markers and articles.

One surprising aspect of this hypothesis is that the article developed from the demonstrative meaning this rather than that. Both Greenberg and Reid point out that the most distant spatial demonstrative is the normal source for definite articles. Thus we might have expected the Dusunic article to be something like *lo’ rather than *ti.

However, the demonstratives ti and tu occur as post-posed markers of definiteness in various Murutic and Paitanic languages:

---

13Hurlbut (1988) labels this form in Labuk Kadazan as referring to objects which are out of sight.
14Reid posits both *tu and *na (> Dusunic *no ?) as Proto-Austronesian demonstratives.
(22) **Timugon Murut**; Prentice, 1971, p. 64:

magiraq i barigil rirun ra sapuk-ti

*borrow subj Barigil you(SG) obj blowpipe-the*

‘Barigil will borrow the blowpipe from you.’

**Kalabuan** (Paitanic); Spitzack, 1988, p. 106:

Piison ku papan tu so’ basi

*split I board this with machete*

‘I split this board with a machete.’

**Tombonuo** (Paitanic); John King, 1988, p. 158:

Tinaakan no Amjan si Amad nu bosu tu

given AGT Amjan subj Amad obj knife spec

‘Amad was given the knife by Amjan.’

Other demonstratives can also appear as markers of definiteness in these languages:

(23) **Timugon Murut**; Prentice, 1971, p. 43:

mariuq lalaing-no

*bathe child-the*

‘The child will bathe.’

**Kalabuan** (Paitanic); Spitzack, 1988, p. 120:

Raragang no matiru

*infant that sleeping*

‘The baby is sleeping.’

Now it is crucial to our hypothesis that the definite article in Proto-Dusunic must have been pre-posed. It might be argued that the existence of cognate post-posed articles in Murutic and Paitanic makes it doubtful that the Dusunic article was actually pre-posed. But the Murutic and Paitanic evidence is less of a problem than it seems. The common ancestor of these three sub-groups was much earlier than the stage of Dusunic we are discussing here. Moreover, Greenberg points out that the position of demonstratives with respect to their head noun is often much freer than the position of articles.
Greenberg cites two examples of demonstratives which had relatively free word order in the proto-language but stabilized in different configurations as articles developed in the various daughter languages. First, Latin ille / illa *that* gave rise to pre-posed articles in Western Romance (e.g. French le / la) but suffixed articles in Rumanian. Second, the Niger-Congo noun class markers are stage III articles derived from Proto-Niger-Congo demonstratives which registered (i.e. agreed with) the class membership of the head noun. In some languages these markers are prefixed, while in other languages they are suffixed.

In some Dusunic languages demonstratives are often doubled, appearing both before and after the head noun. The same pattern can also be found in languages of the other families discussed here:

(24) **Kimaragang** (Dusunic):
Isai minomoli ditih tubat ditih?
*who bought this*(OBJ) *medicine this*
‘Who bought this medicine?’

**Tombonuo** (Paitanic); John King, 1988, p. 158:
Inobaran ku yo ikau nu itu tuturan tu.
told *I ... you*(SUBJ) OBJ SPEC story SPEC
‘I told you the story.’

These considerations make the placement of the definite article before the head noun in Proto-Dusunic seem far less of a problem.

The hypothesis sketched out thus far seems to require us to assume that the Proto-Dusunic case markers were distinct from the definite article. In fact, this seems very likely on independent grounds. We have already noted that in modern Dusunic languages the case markers double as articles. However, this is not the case in Murutic and Paitanic languages, as seen in examples (21) and (23) -- case marking and definiteness are marked
by separate functors, with case markers occurring before the head noun and determiners normally post-posed.

Reid (1978) suggests that Proto-Philippine had three case markers, nominative, genitive and dative/locative, with distinct forms for proper vs. common nouns in each category. Objects were marked with either GEN or DAT case (both patterns are relatively frequent in modern Philippine languages). This basic pattern is reflected in all of the Sabahan languages under discussion here, as shown in the following tables:

(25) **Common noun markers:**

<table>
<thead>
<tr>
<th>Language</th>
<th>NOM</th>
<th>GEN</th>
<th>ACC</th>
<th>DAT/LOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proto-Philippine</td>
<td>*'i</td>
<td>*na/*nu</td>
<td>---</td>
<td>*di/*sa</td>
</tr>
<tr>
<td>Paitanic</td>
<td>Ø</td>
<td>no’</td>
<td>so’</td>
<td>so’</td>
</tr>
<tr>
<td>Kalabuan</td>
<td>Ø</td>
<td>nu</td>
<td>nu</td>
<td>so</td>
</tr>
<tr>
<td>Tombonuo</td>
<td>Ø</td>
<td>ru/nu</td>
<td>ra</td>
<td>ra</td>
</tr>
<tr>
<td>Murutic</td>
<td>Ø</td>
<td>ru/nu</td>
<td>ra</td>
<td>ra</td>
</tr>
<tr>
<td>Timugon</td>
<td>Ø</td>
<td>nu</td>
<td>do</td>
<td>do</td>
</tr>
<tr>
<td>Tagal</td>
<td>Ø</td>
<td>ru/nu</td>
<td>ra</td>
<td>ra</td>
</tr>
<tr>
<td>Bisayan</td>
<td>Ø</td>
<td>nu</td>
<td>do</td>
<td>do</td>
</tr>
<tr>
<td>Tatana</td>
<td>Ø</td>
<td>nu</td>
<td>do</td>
<td>do</td>
</tr>
<tr>
<td>Dusunic Kimaragang</td>
<td>i/o</td>
<td>di/do</td>
<td>di/do (s)id</td>
<td></td>
</tr>
</tbody>
</table>

---

15 Reid suggests that definite objects may have taken DAT marking while indefinite objects took GEN marking, as in some modern languages (e.g. Tagalog).

16 These tables are based on a superficial inspection of incomplete data, and should not be taken as definitive.
Only Dusunic has distinctive definite vs. indefinite forms of the case markers.

One plausible explanation is that the old proper noun forms were generalized as definite common noun markers. This might also explain the fact that only Dusunic among the Sabahan languages has a non-zero nominative marker for common nouns. If this is correct, the case markers of Proto-Dusunic may have been much like those of modern Tatana. However, confirmation of this hypothesis must await a more systematic reconstruction of Proto-Dusunic phonology and morphology.

§5. Evidence from Lotud

The hypothesis that Proto-Dusunic had a definite article *ti gains significant support from some Lotud materials recently made available to me by Doris Blood. Lotud appears to be a strictly t-less dialect, i.e. one in which the moveable t- is never observed on vowel-initial nouns. The few nouns which do carry the moveable t- (e.g. tobpina’i

---

17 When sid appears before proper names, a proper-name marker appears which is homophonous with the t-less form of the definite case marker, i. Forschner refers to this marker as the “personal article”. It is often used with vocatives and terms of address as well as proper names.
sibling) are almost certainly borrowings from Central Dusun or Penampang Kadazan.

Lotud has a preposed particle ti which seems to function primarily as a definite article, as in example (27b), but also has indefinite-specific uses as in (27a):

\[(27) \quad \text{§a. Ati no ti ikam.}\]
\[\text{this FOC ART mat}\]
\[\text{This is a mat.}\]

\[\text{§b. ... orikot ti addau do molukas.}\]
\[\text{arrive ART day of house.warming}\]
\[\text{... when the day of the house-warming arrives.}\]

\[\text{§c. Asanang ti ginawo ku do ...}\]
\[\text{happy ART heart my COMP}\]
\[I am happy that ... (Lit: My heart is happy that ... )}\]

\[\text{§d. Olomi’ ti tana do ...}\]
\[\text{soft ART earth COMP}\]
\[The ground is soft because ...}\]

As seen in these examples, ti occurs equally before vowel-initial and consonant-initial nouns. It never occurs before generic nouns, whether vowel-initial or consonant-initial:

\[(28) \quad \text{§a. Mamatuw oku do ikam.}\]
\[\text{weave I(NOM) ACC mat}\]
\[I am weaving a mat.}\]

\[\text{§b. Nakaakan ko no do ubat?}\]
\[\text{eat(PERF) you(NOM) already ACC medicine}\]
\[Have you already taken medicine?}\]

\[\text{§c. Sada’ ku nopo onsi nu manuk.}\]
\[\text{fish my only meat of chicken}\]
\[My viand is chicken meat.}\]

The fact that subjects in Philippine-type languages are normally definite might lead one to identify ti as a nominative case marker. However, examples (27a) and (28c)...
form a minimal pair showing that the relevant distinction here is specific vs. generic rather than grammatical function, since in both cases the nouns of interest occur as predicate nominals.

A great deal of work remains to be done in analyzing the functions of this particle, the related particles *tu* and *di*, and of the Lotud system of NP marking in general. But preliminary observations make it seem plausible that the function of *ti* in modern Lotud is very similar to that of the non-generic article *ti* proposed for Proto-Dusunic.

**References**


King, Julie K. 1984(a). The Rungus language. In King and King, eds.

King, Julie K. 1984(b). The Dumpas language. In King and King, eds.


Smith, Kenneth D. 1984. The languages of Sabah: a tentative lexicostatistical classification. In King and King (eds.).