

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

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 t̚- in the languages with no vowel-initial nouns, e.g.:

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

The presence vs. absence of initial t̚- 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 t̚ of roots.”

It is clear from wider comparative evidence that the direction of change was the opposite -- the initial t̚- 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 t̥'s exhibit this type of variable distribution. Many noun stems begin with t̥- in all Dusunic languages, as well as in other more distantly related languages:

| | | | | |
|-----|---------------|-------------------|--------------|-----------------|
| (2) | Rungus | Kimaragang | Malay | English |
| | tali | tali | tali | <i>rope</i> |
| | ta'ap | ta'ap | atap | <i>roofing</i> |
| | tambang | tambang | --- | <i>deer</i> |
| | tindal | tindal | --- | <i>dry land</i> |
| | to'un | to'on | tahun | <i>year</i> |

Our concern here is with the “moveable t̥-” illustrated in (1). Where did this t̥- 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 t̥- 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 t̥- 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 t̥-.

§2. The distribution of t̥-

The superficial distinction referred to above between languages with or without the moveable t̥- was based on the presence vs. absence of t̥- 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 t̥-, 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 \underline{t} - in citation forms there are some environments where the \underline{t} -less form is required. Similarly, in some languages which lack the \underline{t} - in citation forms of nouns, a \underline{t} nevertheless appears before vowel-initial nouns in many environments. In other languages, there is no trace of this intermittent \underline{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 \underline{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 \underline{t} - appears in some contexts but not in citation forms, such as Rungus, as “stage II” languages.²

I will refer to languages which show no trace whatsoever of the moveable \underline{t} as “ \underline{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.³ 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 “ \underline{t} -less” language.

§2.1 Moveable \underline{t} - in Kimaragang (a stage III dialect)

Since the “normal” (i.e. citation) form in Kimaragang always has \underline{t} -, I will attempt to describe the environments in which the moveable \underline{t} - does not appear. Following Greenberg’s terminology, I will refer to the \underline{t} -initial form of the noun as the “articulated

²These 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.

³Bisaya combines with Dusunic to form what Ruhlen (1987) has labeled the Dusun-Bisaya group.

form”, and the form without an initial t̚- as the “non-articulated form”.

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

| (3) | Address | Reference | Gloss |
|-----|----------------|------------------|--------------------|
| | aki | taki | <i>grandfather</i> |
| | odu | todu | <i>grandmother</i> |
| | idi | tidi | <i>mother</i> |
| | ama | tama | <i>father</i> |
| | obpinee | tobpinee | <i>sibling</i> |
| | ambaya | tam(ba)baya | <i>friend</i> |

(b) Proper names are frequently vowel-initial, e.g. Uji, Ogidang, Ajah, Ibal (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 t̚- never appears on noun roots in derived verbal forms. That is, verbal morphology requires the use of non-articulated forms. Examples:

| (4) | Citation | Gloss | Verbal form | Gloss |
|-----|-----------------|----------------|--------------------|-----------------------|
| | tasu | <i>dog</i> | mag-asu | <i>to hunt</i> |
| | taddau | <i>sun</i> | mag-addau | <i>dry season</i> |
| | tanak | <i>child</i> | mag-anak | <i>give birth</i> |
| | tuwa | <i>fruit</i> | mong-uwa | <i>bear fruit</i> |
| | talud | <i>canoe</i> | moning-alud | <i>build a canoe</i> |
| | tobpinee | <i>sibling</i> | mi-obpinee | <i>to be siblings</i> |
| | tulun | <i>person</i> | mi-ulun | <i>not related</i> |
| | tagung | <i>gong</i> | mi-agung | <i>play the gongs</i> |

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

| | | | | | | |
|-----|--------|---------------|------------|--------------------|------------|-------------------|
| (5) | taddau | <i>sun</i> | tangaddau | <i>mid-day</i> | sang-addau | <i>one day</i> |
| | tulun | <i>person</i> | tanga-ulun | <i>middle-aged</i> | song-ulun | <i>one person</i> |
| | tinan | <i>body</i> | | | song-inan | <i>one thing</i> |

(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 | <i>to have children</i> |
| | waro tulun sitih | ki-tulun sitih | <i>there are people here</i> |
| | --- | ki-addau | <i>to be sunny</i> |

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

| | | |
|-----|--------------------------|---------------------|
| (7) | tasi-asi | <i>orphan</i> |
| | talun-alun | <i>road</i> |
| | tuwit-uwit | <i>bird species</i> |
| | tundundu (< *tundu-undu) | <i>heart</i> |

(g) The moveable t̚- 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 | <i>return</i> | t-uli-on | <i>place one returns to</i> |
| | odop | <i>sleep</i> | t-odop-on | <i>sleeping place</i> |
| | akan | <i>eat</i> | t-a-akan-on | <i>food; cooked rice</i> |
| | iim | <i>seek</i> | t-iim-on | <i>that which is sought</i> |
| | a-ra'at | <i>bad</i> | t-a-ra'at | <i>that which is bad</i> |

(h) Stage III languages differ concerning the appearance of moveable t̚- 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 t̥- as a noun marker in very recent borrowings.

(i) Moveable t̥- also appears on adjectives used as modifiers within a noun phrase, but not on predicative adjectives. In relative clauses the moveable t̥- 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. *That man is big.*
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 -t̥ before predicate adjectives. However, in Labuk Kadazan, a stage III language fairly closely related to Kimaragang, no t̥ 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 t̥ 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 ṭ 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:

| | | | |
|------|------------------------------|------------|------------|
| (12) | | NOM | ACC |
| | Definite ⁶ | it | dit |
| | Indefinite | ot | dot |

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

- (13) ... ka di-ṭ idi di-ṭ ondu dino. ... *the woman's mother said.*
 say the mother of woman that

In the standard Rungus orthography, this final -ṭ 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:

⁶Forschner 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 *the child stood up*
stand.up that child

iti-t ongo-tangon diti *all these folk-tales*
this pl-folk.tale this

Numerals:

Varo-po, ka, iso-t ondu *there was once, they say, a woman*
exist-yet say one woman

turu-t adau *seven days*
seven day

Particles:

sid koturu=no-t adau *on the seventh day*
at seventh=already day

iso po-t ulun *one more person*
one yet person

However, George Appell (personal communication) informs me that this is a purely orthographic convention⁷, 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 ʈ, as in (15a); but in the vast majority of such cases the moveable ʈ appears, written on the existential predicate as in (15b):

(15) §a. aso ama *has no father*
not.exist father

⁷Presumably this device is intended to preserve the citation form of the noun in all contexts.

| | |
|--|--------------------------|
| §b. <u>varo-t</u> ombolog <i>exist bird</i> | <i>there was a bird</i> |
| aso= <u>po-t</u> anak <i>not.exist=yet child</i> | <i>still childless</i> |
| <u>aso-t</u> andaman=ku <i>not.exist remember=I</i> | <i>I do not remember</i> |

At first glance, it appears that the distribution of moveable ṭ in Rungus is essentially the same as in Kimaragang. However, there are environments where moveable ṭ 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 |
|--|----------------------------|------------------------------|
| <u>ondu</u> it anak <i>female the child</i> | <u>t(ong)ondu</u> it tanak | <i>The child was a girl.</i> |
| <u>ulun</u> oku <i>person I</i> | <u>tulun</u> okuh | <i>I am human.</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) <u>Asu</u> ong asandang <i>dog if carried.on.shoulder</i> | nga oputut-putut; <i>also break</i> |
| <u>onsung</u> ong asandang <i>mortar if carried.on.shoulder</i> | nga opudung gama humago. <i>also cut work hurry</i> |

‘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 ʈ- (tasu and tosung⁸).

In Rungus, no -ʈ occurs after the locative/dative marker sid. In Kimaragang and other stage III languages, however, ʈ-initial (articulated) forms are used in this environment:⁹

| (18) Rungus | Kimaragang | Gloss |
|--|-------------------------------|---|
| sid <u>alun</u> -alun | sid <u>talun</u> -alun | <i>on the road</i> |
| sid <u>isan</u> do tagad ¹⁰ | sid <u>tisan</u> do tagad | <i>at the edge of the field</i> |
| ... sid <u>idi</u> dit anak | ... sid <u>tidi</u> dit tanak | <i>(spoke) to the mother of the child</i> |
| ... sid <u>apad</u> | --- | <i>(went out) to the porch</i> |
| sid <u>azup</u> dino raja | sid <u>tayup</u> ... | <i>in the trap of the king</i> |

Another environment in which Kimaragang requires a ʈ- which does not appear in Rungus involves the plurality marker, ong-. When plural nouns are formed in Kimaragang, moveable ʈ- 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 ʈ- following the prefix. The moveable -ʈ 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 -ʈ between the plural prefix and nominal root in Rungus:

| (19) Rungus | Kimaragang | Gloss |
|-------------|---------------|-----------------|
| ongo-ulun | tongo-tulun | <i>people</i> |
| ongo-uripon | tongo-turipon | <i>slaves</i> |
| anga-anak | tongo-tanak | <i>children</i> |

⁸The source of the first -n- in Rungus onsung is not clear. Many Austronesian roots are reconstructed with an optional nasal infix, which may be a factor here.

⁹In Central Dusun (Bundu dialect) the corresponding locative marker is id, as in id tanak 'to the child' (not *id anak). No such marker exists in Labuk Kadazan, so it is difficult to verify whether ʈ-initial nouns would be used in locative expressions.

¹⁰tagad is a stem which has initial ʈ- in all Dusunic languages.

Finally, no -t appears in Rungus following the (archaic?) possessive marker, dang:

| | | | |
|------|----------------------------|---------------------------------|--|
| (20) | Rungus dang ulun | Kimaragang dang tulun | Gloss <i>belonging to other people</i> |
|------|----------------------------|---------------------------------|--|

§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 “identif[ies] 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,

¹¹Of 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. ʔa 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 t̥- 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 t̥- 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 t̥- 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 t̥- 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 t̥-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 t̥- is powerful evidence in support of membership in the Dusunic family, lack of the moveable t̥- does not, in and of itself, rule out such membership.

§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 t̥-.

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

¹²The only Dumpas data immediately available comes from Hurlbut's 1976 wordlist. In addition to t̥-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.

| | | | |
|---|------------|------------|------------|
| (21) | NOM | ACC | LOC |
| this | iti | diti | siti |
| that (near) | ino | dino | sino |
| that (far) | iloʹ | diloʹ | siloʹ |
| the aforementioned ¹³ | iri | diri | siri |

It seems clear that the original demonstrative forms were *ti, *no, *loʹ 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 Proto-Dusunic had a pre-posed definite article *ti, derived from (and homophonous with) the demonstrative meaning ‘this’. This definite article eroded to t̥- 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 *loʹ 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).¹⁵ This basic pattern is reflected in all of the Sabahan languages under discussion here, as shown in the following tables:¹⁶

(25) **Common noun markers:**

| | NOM | GEN | ACC | DAT/LOC |
|-------------------------|------------|------------|------------|----------------|
| Proto-Philippine | *'i | *na/*nu | --- | *di/*sa |
| Paitanic | | | | |
| Kalabuan | Ø | no' | so' | so' |
| Tombonuo | Ø | nu | nu | so |
| Murutic | | | | |
| Timugon | Ø | ru/nu | ra | ra |
| Tagal | Ø | nu | ra | ra |
| Bisayan | | | | |
| Tatana | Ø | nu | do | do |
| Dusunic | | | | |
| Kimaragang | i/o | di/do | di/do | (s)id |

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

¹⁶These tables are based on a superficial inspection of incomplete data, and should not be taken as definitive.

(26) **Proper noun markers:**

| | NOM | GEN | ACC | DAT/LOC |
|-------------------------|------------|------------|------------|----------------------|
| Proto-Philippine | *si | *ni | --- | *ka ni/*kay |
| Paitanic | | | | |
| Kalabuan | si | ni | ?? | ?? |
| Tombonuo | si | ni | ?? | ?? |
| Murutic | | | | |
| Timugon | i | ri | ri | ?? |
| Tagal | i | ri | ?? | ?? |
| Bisayan | | | | |
| Tatana | i | ni | ?? | ?? |
| Dusunic | | | | |
| Kimaragang | i | di | di | sid+i- ¹⁷ |

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 *t̥i 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 t̥- is never observed on vowel-initial nouns. The few nouns which do carry the moveable t̥- (e.g. tobpina'i

¹⁷ When sid appears before proper names, a proper-name marker appears which is homophonous with the t̥-less form of the definite case marker, i. Forscher 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 \underline{ti} which seems to function primarily as a definite article, as in example (27b), but also has indefinite-specific uses as in (27a):

(27) §a. Ati no \underline{ti} ikam.
 this FOC ART mat
This is a mat.

§b. ... orikot ti addau do molukas.
 arrive ART day of house.warming
... when the day of the house-warming arrives.

§c. Asanang ti ginawo ku do ...
 happy ART heart my COMP
I am happy that ... (Lit: My heart is happy that ...)

§d. Olomi' ti tana do ...
 soft ART earth COMP
The ground is soft because ...

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

(28) §a. Mamatuw oku do ikam.
 weave I(NOM) ACC mat
I am weaving a mat.

§b. Nakaakan ko no do ubat?
 eat(PERF) you(NOM) already ACC medicine
Have you already taken medicine?

§c. Sada' ku nopo onsi nu manuk.
 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 \underline{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

- Clayre, B.M. 1966. A comparison of some dialects of Dusun. *Sabah Society Journal* 3:3-12.
- Dunn, Phyllis and Charles Peck. 1988. Noun phrases in Tatana'. In Peck (ed.).
- Forschner, Traugott. 1978. Outline of Rungus Grammar (2nd edition). Kudat: Protestant Church in Sabah. (mimeo)
- Greenberg, Joseph H. 1978. How does a language acquire gender markers? In Greenberg (ed.), Universals of human language, vol 1. Stanford: Stanford University Press. Pp. 47-82.
- Greenberg, Joseph H. 1981. Nilo-Saharan moveable-k as a Stage III article (with a Penutian typological parallel). *Journal of African Languages and Linguistics* 3:105-112.
- Greenberg, Joseph H. 1989. Proto-linguistic variation: a link between historical linguistics and sociolinguistics. Paper presented at the Berkeley Linguistics Society.
- Hurlbut, Hope. 1976. Dumpas wordlist. Unpublished.
- Hurlbut, Hope. 1988. Verb morphology in Eastern Kadazan. *Pacific Linguistics* B-97.
- King, John Wayne. 1988. Tombonuo clauses. In Peck (ed.).

- 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.
- King, Julie K. and John Wayne King (eds). 1984. Languages of Sabah: a survey report. Pacific Linguistics C-78.
- Lasimbang, Rita. 1988. Kadazan narratives with special reference to oral versus written styles. MA thesis, University of Texas at Arlington.
- Peck, Charles (ed.). 1988. Borneo Language Studies I: Sabah syntax papers. Language Data: Asian-Pacific Series no. 14. Dallas: SIL.
- Prentice, D.J. 1971. The Murut languages of Sabah. Pacific Linguistics C-18.
- Reid, Lawrence A. 1978. Problems in the reconstruction of Proto-Philippine construction markers. In Second International Conference on Austronesian Linguistics: Proceedings, vol 1, pp. 33-66. Wurm and Carrington (eds.). Pacific Linguistics C-61.
- Reid, Lawrence A. 1979. Evidence for Proto-Philippine nominative marking. *Philippine Journal of Linguistics* 10:1-20.
- Ruhlen, Merritt. 1987. A guide to the world's languages, vol.1: Classification. Stanford:Stanford University Press.
- Smith, Kenneth D. 1984. The languages of Sabah: a tentative lexicostatistical classification. In King and King (eds.).
- Spitzack, John A. 1988. Kalabuan clauses. In Peck (ed.).