

The relational noun marker in Tobelo (Northeast Halmaheran)*

Gary Holton

University of Alaska Fairbanks

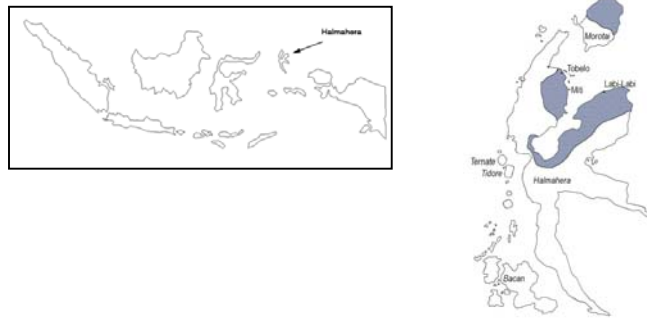
gary.holton@uaf.edu

In the Northeast Halmaheran languages (Tobelo, Galela, Sahu, Modole, Pagu, and Tabaru) all nouns are immediately preceded by an article *o* or *ma*. These articles co-occur freely with other nominal morphology, with the exceptions of first and second person possessive pronouns. Since they occur on all nouns, both articles can function as ligatures within the noun phrases, particularly within possessive constructions and possessive-like attributive constructions. On simple nouns the function of the articles is less clear. This paper draws on a quantitative analysis of Tobelo narrative texts to argue that the articles function as markers of dependency relations in discourse. The article *o* is the default article used new or unimportant referents, while the article *ma* signals a dependency relationship between the following noun and another discourse entity, which may or may not be explicitly mentioned. This notion of discourse dependency not only confirms the intuition of early twentieth century observers but also helps to explain apparent head shift phenomena in complex noun phrases. This paper concludes with a discussion of relationship between Northeast *ma* and Proto-Malayo-Polynesian **ma*, suggesting a possible additional path of contact between North Halmaheran languages and neighboring Austronesian languages.

1. Introduction

Tobelo is the largest of six language varieties (the others being Galela, Tabaru, Modole, Loloda, and Pagu) which together comprise the Northeast Halmaheran (NEH) language (Voorhoeve 1988). All NEH languages contain two noun markers *o* and *ma* whose syntactic status and function is difficult to characterize precisely.

Figure 1. Map showing location of Halmaheran island and the Tobelo language area



* This paper is a preliminary version of a paper in preparation for the proceedings of the tenth International Conference on Austronesian Linguistics. The author wishes to thank Yohanis Labi, Matias Oga, Paltiel Oga and family, Elkana Kukihi and family, Domingus Diba, Frans Diba, Paulina Tindagi, Tobias Tjileni, and Jason Moloku for sharing their knowledge of Halmaheran languages. Fieldwork in Halmahera was supported in part by the University of Hawai'i and the Henry Luce Foundation (grant #P95280F348B164).

While these particles have similar distributional features to definite and indefinite articles in other languages, they do not mark definiteness.¹

- (1) o-kaho i-oara-ilye
 NM-dog 3A-run-DIR
 ‘a/the dog ran by’
- (2) ma-kaho i-oara-ilye
 NM-dog 3A-run-DIR
 ‘a/the dog ran by’

Since they occur on every noun within the NP, both *o* and *ma* may behave as LINKERS or LIGATURES.

- (3) o-hene o-tonaka
 NM-turtle NM-earth
 ‘land tortoise’
- (4) o-gota ma-roehe
 NM-tree NM-body
 ‘tree trunk’
- (5) ahi-kaho ma-ingiri
 1SG.POSS-dog NM-tooth
 ‘my dog’s tooth’

Both *o* and *ma* have scope over the entire NP. That is, whatever their function, alternations between the noun markers occurring on the head noun affect the interpretation of the entire NP. The NP in (6) is marked with *o* occurring on the first noun, whereas, the NP in (7) is instead marked with the second person singular possessive prefix *ani*, also occurring on the first noun.

- (6) [o-karajanga ma-dodobuho] no-goraka i-ni-timono-oka
 NM-work NM-heavy 2SG.A-carry 3A-2SG.U-old-PERF
 ‘you are [too] old to do heavy work’
- (7) dika i-boto [ani-karajanga ma-dodobuho]
 just 3A-finished 2SG.POSS-work NM-burden
 ‘so your heavy work is finished’

In this paper I argue that the various functions of NEH *ma* can be subsumed under a single concept of dependency. I thus characterize *ma* as a RELATIONAL NOUN MARKER, which signals a dependency relationship between the following noun and another discourse referent. This

¹ Where not otherwise indicated, NEH data are from Tobelo. The following abbreviations are used in the glosses: 1, 2, 3, person; A actor; ABL ablative; ALL allative; APPL applicative; ART article; DIR directional; FEM feminine; INC inclusive; LOC locative; MA generic *ma* of unspecified function; MASC masculine; NEG negative; NM noun marker; PL plural; POSS possessive; PRO independent pronoun; REL relativizer; REP repetitive; RDP reduplication; RFLX reflexive; SG singular; U undergoer; V verb formative. Tobelo data are transcribed using a practical orthography. The following symbols differ significantly from IPA conventions: *j* alveo-palatal affricate; *ny* palatal nasal; *ng* velar nasal; *ly* palatal lateral approximant.

referent or HEAD may be explicitly mentioned in the discourse, or the head may be implied. The pathway via which the dependency relationship is established may vary, and a variety of types of dependency relationships exist, including possession, attribution, and antecedence. The main purpose of this paper is to suggest a coherent semantic framework which subsumes all of the types of dependency relationships signaled by the NEH relational noun marker.

The noun marker *o* is analyzed here as a default marker which occurs when no other noun prefix is present. While this paper draws primarily from Tobelo data, the generalizations regarding *ma* hold for all six NEH languages. Indeed, this characterization of *ma* as a relational noun marker is one of the distinguishing features of the NEH languages.

2. Previous characterizations

The NEH noun marker *o* has been characterized as an ‘article’ (van Baarda 1908); ‘indefinite article’ (Wimbish 1991; Sheldon 1991); and a ‘noun marker’ (Fortgens 1928:398). There is no unifying semantic feature or pragmatic function associated with *o*. Instead, the noun marker *o* can be characterized as the default noun prefix which occurs when no other prefix is present.

In contrast to *o*, the noun marker *ma* is generally acknowledged to be multi-functional. While it is possible to associate different functions of *ma* with homophonous *ma* morphemes, most previous analyses have attempted to unite these functions within a single semantic rubric. Visser & Voorhoeve note that the that ‘the presence of a relational particle with such a variety of functions is characteristic of all North Halmaheran languages’ (1987:37). Writing about Pagu in particular, Wimbish remarks that the ‘functions are somewhat complex, involving not only indefinite and definite ... (however, not in the usual straightforward division...), but also involving the property of nuclear’ (1991:17). Though Wimbish refers here to both *o* and *ma*, it is clearly the latter which has such complex functions.

Three primary functions have been distinguished for the noun marker *ma*. Namely, *ma* has been characterized as a possessive prefix; a marker of attribution; and a marker of anaphoric reference. As a possessive prefix, *ma* is usually treated as a non-human possessive pronoun (cf. van Baarda 1891:91). An example of this use occurs in (5) above. As a marker of attribution, *ma* is prefixed to the attribute, indicating the preceding noun has the quality described by that attribute. As described by van Baarda, ‘one also uses this *ma* to indicate that a noun is a modification of quantity or quality of another noun’ (1891). While the existing literature does not attempt to relate these two functions of *ma*, the attributive function is clearly an example of what Ross (1998) has called a POSSESSIVE-LIKE ATTRIBUTIVE CONSTRUCTION.

The characterization of *ma* as a marker of anaphoric reference reflects much more variation in previous approaches. It has long been recognized that NEH nouns tend to be introduced with *o* and then referred to subsequently with *ma* (Hueting 1936: 348). This observation naturally leads to a characterization of *ma* as a definite article. Indeed, the term definite article is found both in the early work of van Baarda (1908) and in later work by Wimbish (1991) on Pagu and Sheldon (1991) on Galela. None of these descriptions deals directly with the problem of tokens of *ma* occurring on nouns which are clearly not definite, such as the use of *ma* with non-referential nouns. Nevertheless, the concept of *ma* as a marker of anaphoric dependency is latent in some of the early descriptions.

Nouns which obligatorily occur with *ma* in all contexts provide a particular challenge for the definite article interpretation. Van Baarda makes the following observation regarding such nouns.

‘One has a number of nouns in Galela which ALWAYS [orig. emph.] have the possessive pronoun *ma* before them, in place of the article *o*, because the concept which it denotes can not be thought of except in relation with another thing.’ (van Baarda 1891:100)

Both Schmidt (1895) and van der Veen (1915) are extremely careful to avoid referring to *ma* as a definite article. Their purpose seems to be to distinguish the anaphoric function of *ma* from the similar anaphoric function typically associated with definite articles. Thus, Schmidt characterizes *ma* as ‘a particle, which expresses anaphoric reference’ (1895:135). This characterization could equally apply to a definite article, but Schmidt’s avoidance of the term article is deliberate. Van der Veen generalizes further, characterizing *ma* as a ‘relation indicator’ for non-human class of nouns (1915).²

3. Distribution of noun markers

The noun markers *o* and *ma* do not occur independently; they must be attached to a word. Since most NEH roots may occur as either nouns or verbs, the presence of a noun marker is taken as criterial for the definition of nouns. In contrast to similar markers in the Philippine languages surveyed by Reid (2002), NEH noun markers occur only immediately preceding a noun.

3.1 Co-occurrence with other word classes

The NEH noun markers may co-occur with many other word and morpheme classes in the NP. In particular, noun markers may co-occur with numerals, as in the following examples.

(8) o-nyawa moi
 NM-person one
 ‘one person’

(9) ma-kaho moi
 NM-dog one
 ‘one dog’

Noun markers also occur with deictic determiners, which specify location within a water-based coordinate system.³

(10) ma-buho gaa i-pali
 NM-pandanus DISTAL.AREAL 3A-roll.up
 ‘roll up that pandanus’

(11) o-kapa danena
 NM-ship SEAWARD.PUNCTUAL
 ‘that ship down there’

Noun markers also co-occur with directional suffixes.

² The term used in the original Dutch is ‘betrekking-aanduiders’.

³ The complete NEH directional system contrasts two orthogonal dimensions, seaward-landward and up-down coast, as well as a distinction between areal and punctual location, yielding a total of at least 13 distinct deictic determiners.

- (12) o-hongana-ika
 NM-forest-ABL
 ‘from the forest’
- (13) o-lyoku-iha to-oiki
 NM-mountains-LANDWARD 1SG.A-go
 ‘I’m going inland to the mountains’
- (14) ma-tiba-uku wo-noa-uku ma-gula
 NM-bamboo-DOWNWARD 3SG.MASC.A-fill-down NM-sugar
 ‘he put sugar down in the bamboo.’

In NEH languages, aspect markers may occur with both nouns and verbs. Noun markers co-occur with aspect markers when they occur on nouns.

- (15) ma-moi-oli
 NM-one-REP
 ‘one more, one again’
- (16) nako o-Galela gaa dau de ma-akere-oka
 if NM-G. DISTAL.AREAL downward with NM-water-PERF
 ‘as for Galela down the coast, they already have a lake’

Noun markers also co-occur with prepositions. There is only one of these, an associative/instrumental preposition borrowed from Malay *dengan*.

- (17) de ma-kakatama n-a-lye-ino
 with NM-tongs 2SG.A-3U-roll-ALL
 ‘roll it up with the tongs’
- (18) ngohi-o to-modeke de o-Matias
 1SG.PRO-also 1SG.A-agree with NM-M.
 ‘I also agreed with Matias’

Noun markers co-occur with interrogative pronouns.

- (19) o-kia no-oiki?
 NM-what/where 2SG.A-go
 ‘Where are you going?’
- (20) o-kia no-olyomo?
 NM-what/where 2SG.A-eat
 ‘What are you eating?’
- (21) o-nagona n-a-make?
 NM-who 2SG.A-3U-find
 ‘whom did you find?’

3.2 Restrictions on co-occurrence

Noun markers are restricted to occurring on nouns. Hence, they do not occur with demonstrative pronouns.

- (22) *o-genanga DISTAL.PUNCTUAL
 *ma-nenanga PROXIMAL.PUNCTUAL

Nor do noun markers occur with independent personal pronouns.

- (23) *o-ngohi 1SG.PRO
 *o-muna 3SG.FEM.PRO

Noun markers also do not occur with possessive pronominal prefixes. Possessed nouns with first, second, or third person human possessor occur with possessive pronoun in lieu of noun marker.

- (24) ahi-kaho (*o/ma-ahi-kaho)
 1SG.POSS-dog
 ‘my dog’
- (25) o-Johanis ai-kaho
 NM-J. 3SG.MASC.POSS-dog
 ‘Johanis’ dog’

However, noun markers do not fill the same templatic slot as possessive prefixes. Evidence for this can be found in the co-occurrence of noun markers with the feminine prefix *ngo*. This prefix occurs optionally with nouns referring to female humans and may co-occur with noun markers or possessive prefixes. Noun markers precede feminine prefix *ngo*, as in (26).

- (26) o-ngo-Rian
 NM-FEM-R.
 ‘Rian’ (woman’s name)

But the possessive prefixes follow the feminine prefix, as in (27).

- (27) ngo-ai-ayo
 FEM-3SG.MASC.POSS-mother
 ‘his mother’

Noun markers do not occur with relative clauses, though relative clauses may be optionally marked by deictic determiner *gena*.

- (28) to-wi-tidingi (gena) wo-oara-oka
 1SG.A-3SG.MASC.U-hit REL 3SG.MASC.A-run-PERF
 ‘the man I hit already ran away’

3.3 Noun markers are not case markers

Both noun markers *o* and *ma* may occur on any noun, core or oblique. In particular, either noun marker may occur on actor or undergoer arguments of both intransitive and transitive verbs. Either *o* or *ma* occurs with dynamic referents of intransitive verbs (S_A macro-role).

- (29) o/ma-momoholehe mo-bata
 NM-young.woman 3SG.FEM.A-jump
 ‘the young woman jumped’
- (30) o/ma-momoholehe mo-hángeru
 NM-young.woman 3SG.FEM.A-sneeze
 ‘the young woman sneezed’

Either occurs with non-dynamic undergoer arguments of intransitive verbs (S_P macro-role).

- (31) *o/ma-momoholehe i-mi-bole*
 NM-young.woman 3A-3SG.FEM.U-tired
 ‘the young woman is tired’

Either noun marker may occur with actor arguments of transitive verbs (A macro-role).

- (32) *o/ma-momoholehe mo-wi-tidingi*
 NM-young.woman 3SG.FEM.A-3SG.MASC.U-hit
 ‘the young woman hit him’

And either may occur with undergoer arguments of transitive verbs (P macro-role).

- (33) *o/ma-momoholehe wo-mi-tidingi*
 NM-young.woman 3SG.MASC.A-3SG.FEM.U-hit
 ‘he hit the young woman’

Having established some of the distributional properties of NEH noun markers, I turn now to an examination of the function of the noun markers *o* and *ma*.

4. Functions of *o*

As the default noun marker, *o* is required when no other noun prefix is present. The noun marker *o* also has a very limited function as a linker with classificatory nouns. In this construction *o* occurs on the second of a sequence of two nouns, where the second noun specifies a category to which the first noun belongs.

- (34) *o-nghohaka o-nauru*
 NM-child NM-male
 ‘male child’

In particular, the N_1 *o*- N_2 construction is standard for numeral classifiers, where N_2 is the classifier noun. More than a dozen numeral classifiers can be identified in NEH languages. Their use is obligatory when counting certain objects. Some Tobelo examples are given below.

- (35) *o-gakana o-gahumu hinóto*
 NM-knife *o*-CL(3-dim) two
 ‘two knives’
- (36) *o-tau o-utu tumidi*
 NM-house NM-CL(long.thin) seven
 ‘seven houses’
- (37) *o-nauoko o-ngai ngimoi*
 NM-fish NM-CL(fish) ten
 ‘ten fish’

Some nouns always occur with *o*. These include globally unique referents,

- (38) *o-wange* ‘sun’

and names (with some exceptions).

Obligatorily possessed nouns cannot occur with *o*.

5.2 Associative possession

The noun marker *ma* marks an associated entity in ASSOCIATIVE constructions, though these are not formally distinguished from possession. In this case *ma* marks the head noun.

- (44) o-hepata ma-daluku
 NM-aren.palm NM-palm.wine
 ‘the wine of the aren palm’
- (45) o-ngotiri ma-lyeha
 NM-canoe NM-keel
 ‘the keel of the canoe’
- (46) o-Tobelo ma-nyawa
 NM-T. NM-person
 ‘a person from Tobelo’

There are no formal grounds on which to distinguish associative possession from regular possession in NEH.

5.3 Attribution

Attributive adjectives are expressed as nouns with *ma* in a possessive-like attributive constructions of Ross’ (1998) type 2. In this case *ma* marks the semantically dependent noun, showing a reversal of headship.

- (47) o-toimi ma-doto
 NM-spear NM-sharp
 ‘sharp spears’
- (48) o-ngotiri ma-dokara
 NM-canoe NM-red
 ‘a red canoe’
- (49) o-harimi ma-eteki
 NM-paddle NM-small
 ‘small paddles’
- (50) o-gota ma-dubuho
 NM-wood NM-heavy
 ‘heavy wood’

Predicative adjectives may be expressed as either nouns or verbs. That is, NEH roots denoting property concepts may occur as either nouns or verbs, as with *pako* ‘large’ in the following examples (Holton 1999).

- (51) o-ode ma-pako
 NM-pig NM-large
 ‘the pig is large’
- (52) o-ode i-pako
 NM-pig 3A-large
 ‘the pig is large’

However, when they occur as nouns, predicative adjectives always occur with *ma* rather than *o*. In addition, nouns referring to abstract qualities occur with *ma* rather than *o*. Thus, *ma-dokara* ‘red, redness’ but not **o-dokara*.

5.4 Relational nouns

Beyond those which are obligatorily possessed or denote adjectives (property concepts), certain nouns always occur with *ma* because they cannot be conceived of ‘except in relation with another thing’ (van Baarda 1891). I refer to such nouns with the label RELATIONAL NOUN. There are several subtypes. The largest category consists of meronymic relational nouns, as in (53). These nouns denote entities which stand in an implicit part-whole relationship with another entity. That is, they are a part of something. For example, *roehe* ‘trunk’ must be the trunk of something, presumably a tree.

- (53) *ma-roehe* ‘trunk [of a tree]’
 ma-timi ‘underside’
 ma-pamuru ‘flames’
 ma-otini ‘handle’

Ordinal numerals form another category of relational nouns. Cardinal numerals occur without noun marker, as in (54).

- (54) *o-kaho hinóto*
 NM-dog two
 ‘two dogs’

But ordinal numerals occur as nouns with *ma*, as in (55).

- (55) *ma-hinóto wo-ruba i-wi-bole*
 NM-two 3SG.MASC.A-trip 3A-3SG.MASC.U-tired
 ‘the second one tripped because he was tired’

Numerals never occur with *o*.

Other relational nouns appear to be highly lexicalized. These include titles, as in (56).

- (56) *ma-koana* ‘king’
 ma-jou ‘God’
 ma-uku ‘district chief’
 ma-pandita ‘priest’

Titles are relational in the sense that the holder of the title holds that title over some other entity. For example, *koana* ‘king’ must be king of some kingdom. A small residue of relational nouns admit no ready semantic explanation.

- (57) *ma-taihe* ‘piglet’
 ma-ngale ‘reason, cause’
 ma-tudili ‘tool for thatching roof’

Additional etymological work may help to provide semantic explanations for the status of forms such as those in (57) as relational nouns.

5.5 Discourse importance

The uses of *ma* discussed so far are determined by lexical or clausal semantics. Over longer stretches of connected discourse, *ma* may be used to indicate that the referent ‘is of special importance to the story at this time’ (Wimbish 1991:102). For example, names usually occur with *o*, but *ma* may be used for emphasis. In folktales initial mentions of major characters may occur with *ma*, as in the following Pagu example.

- (58) genakadau oras moi ma-fen moi i-pudal
 then time one NM-turtle one 3A-appear
 ‘then at that time a certain turtle appeared...’ (Wimbish 1991:102)

In narratives, speakers have control over the use of *ma*. In the following Tobelo excerpt, the first three mentions of the name *Yahe* in the text occur with *o*, as expected. (Recall that names generally occur with the noun marker *o*.) However, the final mention occurs with *ma* in an off-storyline comment.

- (59) nako utu naa naga h-a-huloko
 if other here exist 1INC.A-3U-order
 ‘Now there were others who when asked to work’
- yo-oluku y-a-tebeturu
 3PL.A-not.want 3PL.A-3U-lazy
 ‘didn’t want to because they were lazy,’
- i-ma-howono de o-Yahe
 3-RFLX-wrong with NM-Y.
 ‘but that wasn’t true of Yahe.’
- ((LINES OMITTED))
- jadi ma-lihanga naa naga daku o-Iga
 thus NM-gong here exist upwards NM-I.
- to-una ai-lihanga
 POSS-3SG.MASC.PRO 3SG.MASC.POSS-gong
 ‘so now his gong is up there in Iga.’
- ai-lihanga ma-Yahe gena nenga-oka-o
 3SG.MASC.POSS-gong NM-Y. that there-LOC-also
 ‘yeah that Yahe’s gong is too there.’

The use of *ma* with the name adds a contrastive emphasis, indicating that it is Yahe’s gong which is there, and that the gong is still there today.

5.6 Discourse continuity

Related to the marking of discourse importance is the use of *ma* to mark discourse continuity. In an extended narrative, continued reference is made with *ma*. For example, in the following Tobelo text excerpt both mentions of *gaharu* are non-referential, hence indefinite, but the initial reference is with *o*, while the subsequent reference is with *ma*.

- (60) o-oraha genanga,
 NM-time that
- o= --
 HEST
- o-gaharu yo-ija,
 NM-eagle.wood 3PL.A -buy
- yo-uti o-LabiLabi-ha.
 3PL.A -arrive NM-LabiLabi-landward
 ‘At that time they arrived in Labi-Labi to buy eagle wood.’
- ... ((LINES OMITTED)) ...
- dina mi-a-iha,
 inland 1EXC.A-V-inland
- o-hutu hange,
 NM-night three
- mi-lingiri ma-gaharu.
 1EXC.A-look.for NM-eagle.wood
 ‘We went inland for three nights to look for eagle wood.’

In contrast to discourse importance, which is difficult to characterize quantitatively, discourse continuity of the type exemplified in (60) can be measured using established metrics. One such metric is REFERENTIAL DISTANCE (RD), which counts the number of clauses between a token of a given referent and the nearest preceding token of the same referent, up to a maximum of twenty clauses (cf. Givón 1983). Referents which have been recently mentioned have a low RD, while those which have not been recently mentioned have a high RD. To the extent that the noun marker *ma* is indeed functioning as a marker of anaphoric reference, we expect *ma* to occur preferentially with those referents which have a lower RD.

In order to apply the RD metric I analyzed a small corpus of six Tobelo narrative texts, comprising roughly five thousand lines total. These texts were transcribed from tape recordings of six different speakers of Tobelo, and they are generally representative of modern Tobelo conversational narrative.⁴ The following methodology was applied in order to calculate referential distance. First, for each of these texts in the study corpus I identified those nominal referents which were tracked more or less continuously through the narrative. Referents which were mentioned only in passing were omitted since they did not provide enough tokens with which to examine the alternation between *o* and *ma*. Each token of the tracked referents in the corpus was then coded for one of five lexical forms, either *o* marked; *ma* marked; marked with a possessive pronoun; marked with an independent pronoun; or zero-marked. A lexical zero was considered to be present whenever a pronominal cross-reference marker occurred for which no

⁴ The existing corpus of Tobelo folktales collected in Huetting (1908) was not included in this study. Most of the referents in these stories are anthropomorphized animals which are assigned titles and hence referred to throughout the narrative with full noun phrases and the relational noun marker *ma*. In contrast, modern conversational and historical narratives exhibit much more alternation between the use of *o* and *ma*.

co-referential pronoun or noun phrase occurred in the same clause. First mentions were assigned the maximum RD value of 20. The results of the RD measurements are summarized in Table 2.

Table 2. Referential distance in six short Tobelo narrative texts (recorded 1995)

	<i>ma</i>	<i>o</i>	POSS	PRO	ZERO
Tokens	6	19	3	34	141
Median RD	6.5	19	20	4	1

As expected, the lowest median RD value is found for zero-marked referents. However, the values for nouns marked with *ma* are also quite low. In fact, the median RD for nouns marked with *ma* (RD=6.5) is much closer to that for independent pronouns (RD=4.0) than it is to nouns marked with *o* (RD=19.0). Unfortunately, this result is based on a relatively small number of tokens of referents marked with *ma* (only 6 of 203 total mentions).⁵ Nevertheless, the tendency for *ma* to occur with referents which have been previously introduced into the discourse is confirmed in measurements of referent continuity.

5.7 Contrast with definite article

Distribution of *ma* is similar to that of a definite article. However, many NEH nouns marked with *o* are clearly definite (e.g., names). In particular, *o* does not mark indefiniteness. This is a crucial difference between the NEH noun markers and a system of definite and indefinite articles. Like a definite article, *ma* may signal anaphoric reference. However, the absence of *ma* (and presence of *o*) does not signal the introduction of a new referent which will be tracked through the narrative. Referents introduced with *o* may or may not be important to the discourse, or ‘nuclear’ in the sense of Wimbish (1991). In other words, the choice to mark a noun with *ma* is optional in a way that the use of a definite article is not.

In fact, definiteness may be less important to NEH than referentiality, which can be signaled via choice of person-marking prefix. Non-referential third person singular human actor referents may be indexed via the third person plural actor prefix *yo-*, as in (62).

(61) o-nyawa wo-boa
 NM-person 3SG.MASC.A-come
 ‘a/the man is coming’

(62) o-nyawa yo-boa
 NM-person 3PL.A-come
 ‘Someone is coming’

Only referential third-person undergoer referents are indexed on the verb.

(63) o-pine t-a-ija
 NM-rice 1SG.A-3U-buy
 ‘I bought (the) rice’

(64) o-gota mi-a-tobiki de mi-a-bela-belaka
 NM-wood 1INC.A-3U-break and 1INC.A-3U-RDP-split
 ‘we bucked and split (the) wood’

⁵ Tokens of *ma* on nouns are much more frequent in the texts. However, these tokens occur primarily on referents which are not tracked through the discourse, for example, in attributive constructions and associative possession constructions.

Non-referential third person singular human undergoer referents are not indexed on the verb, so that the verbs in the following examples are coded as if they were intransitive.

- (65) o-pine to-ija
 NM-rice 1SG.A-buy
 ‘I went rice-shopping’
- (66) jadi ngohi to-lye-ua o-hilo
 therefore 1SG.PRO 1SG.A-get-NEG NM-resin
 ‘therefore I didn’t get any resin’

Syntactic processes in NEH are sensitive to referentiality, not definiteness. Thus, while the distribution of the relational noun marker may overlap with that expected for a definite article, there are good reasons for distinguishing *ma* from a definite article. Before continuing with a characterization of *ma*, I consider in the following section the comparative evidence for the origin of *ma*.

6. Origin of the relational noun marker

Comparative evidence indicates that the development of *ma* as a relational noun marker is a NEH innovation. Internal phonological and syntactic evidence point to a recent origin for NEH noun markers. Though written here as prefixes, the noun markers are less phonologically fused than either noun suffixes or verb affixes. When they occur preceding a vowel of the same quality, both the vowel of the noun marker and the initial vowel of the noun are pronounced.

- (67) o-ode [o ode]
 NM-pig
 ‘a/the pig’
- (68) ma-akere [ma akere]
 NM-water
 ‘water’

In contrast, noun suffixes are more fused. Where a vowel-initial morpheme is suffixed to a vowel-final noun, one of the two vowels elides, as indicated in the phonetic transcription of (69).

- (69) o-ngairi-ika [oŋairika]
 NM-stream-ABL
 ‘from the stream’

Verb prefixes are also more fused to the root. Sequences of identical vowels which arise through prefixation are not permitted.

- (70) o-daluku ho-okere [hokere]
 NM-palm.wine 1INC.A-drink
 ‘let’s drink palm wine’
- (71) w-a-ato [wato]
 3SG.MASC.A-3U-say
 ‘he said’

The position of noun markers as prefixes is also unexpected, given the syntactic structure of the NEH noun phrase. NEH NP structures are right-branching, but noun markers occur to the left. The basic NP structure is given in (72), followed by an example from Tobelo.

- (72) NM-N (NM-N) DET NUM
- (73) ma-goduru ma-homoa genanga moi
 NM-youn.man NM-other DISTAL.PUNCTUAL moi
 ‘that one other young man’ (Huetting 1908)

Also, with the exception of a single borrowed preposition (from Malay *dengan*), oblique case marking follows the noun. The NEH noun markers are thus likely to derive from a non-NEH source. Since the languages surrounding NEH are Austronesian, it is likely that the source of the relational noun marker lies in Austronesian languages as well.

In his survey of possessive-like attributive constructions in Oceanic languages, Ross (1998) identifies the source of the linkers in constructions similar to those found in NEH NPs (his type 2) as a noun classifier, as in the following Yabem form.

- (74) m ŋa-kana
 banana LINKER-sweet
 ‘a sweet banana’ (Ross 1998)

But NEH lacks evidence of noun classifiers of this type. We thus need to look elsewhere for a source of NEH *ma*. The Austronesian stative marker **ma* would seem an excellent candidate. Internal evidence for a stative marker as a source for the NEH relational noun marker can be found in the patterning of *ma* with stative verbs.

6.1 Nouns derived from stative verbs occur with *ma*

Many productive processes derive verbs from nouns and nouns from verbs in NEH. The most common nominalization processes are reduplication (75) and initial consonant voicing (76).

- (75) -dowiti ‘carry on shoulder’
 o-do-dowiti ‘stick used for carrying on shoulder’
- (76) -toro ‘pass thru waves’
 o-doro ‘passage thru the waves’

Verbs may be also derived from nouns through the use of the applicative *hi-*.

- (77) -baiti ‘hole’
 -hi-baiti ‘bury’

Many NEH roots occur as either nouns or verbs with no derivational morphology. Such cases may be analyzed as zero-derivation or as instances of categorical indeterminacy.

- (78) -pokoro ‘abdomen’
 -pokoro ‘have dysentery’

However, noun forms corresponding to stative (i.e., intransitive undergoer) verbs always occur with *ma*, as opposed to *o*.

- (79) o-nyawa i-wi-timono
 NM-man 3A-3SG.MASC.U-old
 ‘the man is old’
- (80) o-nyawa ma-timono (*o-timono)
 NM-man NM-old
 ‘the man is old’ / ‘the old man’

6.2 Reflexive/reciprocal verbs occur with *ma*

The noun marker *ma*, or at least a morpheme homophonous with *ma* is found as marker of reflexive and reciprocal verbs, but it cannot co-occur with the undergoer prefix. In particular, reflexive/reciprocal *ma* does not occur with stative verbs. Reflexive derivations of intransitive verbs are valence neutral but admit a more stative, less dynamic interpretation. Thus, while the verbs in both (81) and (82) take actor prefixes, the form with reflexive *ma* denotes the less dynamic ‘sleeping’ as opposed to ‘lay down’.

- (81) wo-idu-ika
3SG.MASC.A-lay.down-ABL
‘he lay down’
- (82) wo-ma-idu-oka
3SG.MASC.A-RFLX-lay.down-PERF
‘he is sleeping’

Reflexive derivations of transitive verbs are valency decreasing. The reflexive prefix *ma-* occurs in lieu of the undergoer prefix. That is, the reflexive and undergoer prefixes cannot co-occur.

- (83) to-mi-ohiki
1SG.A-3SG.FEM.U-wash
‘I washed her’
- (84) to-ma-ohiki
1SG.A-RFLX-wash
‘I washed myself’

The reciprocal prefix *maka-* is derived from *ma-* and exhibits similar valency-decreasing properties.⁶

- (85) ngaro o-hongana-oka ho-maka-make ma wo-na-temo
yet NM-forest-LOC 1INC.A-RECIP-find but 3SG.M.A-1INC.U-speak
‘although we met each other in the forest, he still spoke to us’

In some languages reflexives occur frequently without subject prefix, inviting reanalysis as a noun with *ma* prefix, as in the following examples from Pagu.

- (86) i-a-make de ka ma-idu-oka siadon wo-gila-gilang
3A-3U-see and only RFLX-sleep-PERF until 3SG.MASC.A-RDP-direct
‘when the eagle saw it, it just went to sleep, so he continued on’ (Wimbish 1991:138)

⁶ Some authors analyze the *maka-* as a complex morpheme consisting of the *ma-* reflexive prefix and an additional prefix *ka-*.

- (87) ma-moi-uku w-a-ika gena
 NM-one-down 3SG.MASC.A-V-ABL that
 ma-todokan awi-oti ma-dea-ika
 RFLX-surprize 3SG.MASC.POSS-canoe NM-inside-ABL
 ma-ali-ali ma-dumoi i-lulu
 NM-ring NM-one 3A-roll
 ‘one day when he got to it he was surprised to see a ring rolling around inside of his boat’
 (Wimbish 1991:133)

Since Pagu verbs with the reflexive prefix occur more frequently without person markers, they are more easily interpreted as nouns. Since nearly any NEH root can occur with either verbal or nominal morphology, this yields a possible pathway of reanalysis by which *ma* could be reinterpreted as a noun prefix. Indeed, Schmidt’s (1895) early analysis of NEH languages links the verbal and nominal uses of *ma* and treats the reflexive/reciprocal use of *ma* as primary. In any case, this use points toward an origin of the relational noun marker as a stative marker.

6.3 Stative markers outside Northeast Halmaheran

Indeed, stative markers with identical shape occur widely throughout the region, both in Austronesian and non-Austronesian languages. In particular, a stative *ma* prefix is found in neighboring Tidore, a North Halmaheran (West Papuan) language outside the NEH group. In Tidore *ma* marks third person singular non-human possession, as in NEH, but may also be used for all persons with inalienable nouns (van Staden 2001:126).

- (88) una i-fayaa
 3SG.MASC.PRO 3SG.MASC.POSS-wife
 ‘his wife’ / ‘he has a wife’
- (89) una ma-fayaa
 3SG.MASC.PRO MA-wife
 ‘his wife’ / ‘he has a wife’

It is of particular interest that the category of nouns which admit *ma* as a marker of possession for all persons is precisely the category of inalienable nouns, a class of nouns which cannot be envisaged except in relation to another entity. This use of *ma* in Tidore can be extended to adjectival concepts, yielding a kind of inalienable possessive-like attributive construction, called ‘nominalization’ by van Staden (2001:266).

- (90) mina mi-jang
 3SG.FEM.PRO 3SG.FEM.POSS-beautiful
 ‘she has beauty’ / ‘she is beautiful’
- (91) mina ma-jang
 3SG.FEM.PRO MA-beautiful
 ‘she is a beauty’ / ‘she is beautiful’

The adjectival form with *ma* in (91) describes a more permanent, inherent or inalienable quality, whereas, the form with a possessive prefix is more transitory.

Stative markers are found even more obviously in the Austronesian languages of Halmahera. In Taba and Sawai (South Halmahera-West New Guinea) a similar *ma* is found lexicalized with

stative verbs and dynamic verbs denoting involuntary processes. Bowden (2002) speculates that Taba *ma* may have once been marker of involuntariness. It still occurs in Taba in marginally productive forms such as the following.

- (92) mayas n=fati ma-lol dumik
 smoke 3SG-cover MA-be.big/full be.complete
 ‘smoke covered things ... it was totally covered’ (Bowden 2002:225)

Homophonous noun markers are found in Austronesian languages beyond Halmahera as well. In Kambera (Central Malayo-Polynesian) *ma* occurs as a relativizer (Klamer 1998:316). This leads to a distinction between a paratactic noun plus adjective sequence, as in (93), versus a noun plus relative clause sequence, as in (94).

- (93) na pau rara
 ART mango red/ripe
 ‘the ripe mango’
- (94) pau ma-rara
 mango REL-red/ripe
 ‘a ripe (rather than unripe) mango’
- (95) na pau na ma-rara
 ART mango ART REL-red/ripe
 ‘the mango that is ripe’

In Tagalog *ma* occurs with adjectives denoting transitory states

- (96) ma-bilog ang buwan
 MA-round ART moon
 ‘the moon is full’ (Evans & Ross 2001:282)

The use of *ma* with adjectives in Kambera and Tagalog recalls a similar distribution of *ma* on NEH nouns denoting property concepts.

7. Conclusion

Internal criteria clearly point toward an external source for the NEH relational noun marker, and the evidence for an Austronesian origin of *ma* is rather suggestive. And yet, no similar relational noun markers are found in the other languages of the region. Nevertheless, the connection between *ma* as a relational noun marker and *ma* as a stative marker is quite strong. The source of NEH *ma* is likely to be found in PMP **ma*, originating as a marker of stative verbs, particularly those denoting property concepts, eventually reanalyzed as nouns. The further development of this stative *ma* into a relational noun marker is a NEH innovation.

Essentially, NEH has extended the relational function of *ma* within the possessive and attributive constructions into a general marker of relations between entities. The function of *ma* in discourse is not fundamentally different from the function of *ma* to mark relationships between possessors and possessed, or between entity and attribute. In each case *ma* signals existence of semantic or pragmatic relationship with a previously established entity.

References

- van Baarda, M. J. 1891. *Beknopte spraakkunst van de Galillareesche taal*. Utrecht: Kemink.
- van Baarda, M. J. 1908. *Leidraad bij het bestudeeren van 't Galela'sch dialekt, op het eiland Halmahera*. 's-Gravenhage: Martinus Nijhoff.
- Bowden, John. 2002. *Taba: Description of a South Halmahera Austronesian language*. (Pacific Linguistics 521). Canberra: Pacific Linguistics.
- Evans, Bethwyn & Malcom Ross. 2001. The history of Proto-Oceanic *ma-. *Oceanic Linguistics* 40(2).269-90.
- Fortgens, J. 1928. *Grammaticale aantekeningen van het Tabaroesch, Tabaroesch volksverhalen en raadsels*. *Bijdragen tot de Taal, Land en Volkenkunde* 84.300-544.
- Holton, Gary. 1999. Categoriality of property words in a switch-adjective language. *Linguistic Typology* 3(3).341-60.
- Hueting, Anton. 1908. *O Tobelohoka manga totoade: verhalen en vertellingen in de Tobeloreesche taal*. *Bijdragen tot de Taal, Land en Volkenkunde* 61.1-318.
- Hueting, Anton. 1936. *Iets over de spraakkunst van de Tobeloreesche taal*. *Bijdragen tot de Taal, Land en Volkenkunde* 94.295-407.
- Klamer, Marian. 1998. *A Grammar of Kambera*. Berlin: Mouton.
- Reid, Lawrence A. 2002. Determiners, nouns, or what? Problems in the analysis of some commonly occurring forms in Philippine languages. *Oceanic Linguistics* 41(2).295-309.
- Ross, Malcom. 1998. Possessive-like attribute constructions in the Oceanic languages of Northwest Melanesia. *Oceanic Linguistics* 37(2).234-76.
- Schmidt, P. W. 1895. *Die Sprachlichen Verhältnisse von Deutsch-Neuguinea*. *Zeitschrift für Afrikanische, Ozeanisch und Ostasiatische Sprachen* II(4).
- Sheldon, Howard. 1991. *Galela pronominal verb prefixes*. *Papers in Papuan Linguistics* no. 1, ed. by Tom Dutton, 161-75. (Pacific Linguistics A-73). Canberra: Australian National University.
- van Staden, Miriam. 2001. *Tidore: A Linguistic Description of a Language of the North Moluccas*. Ph.D. thesis, Leiden University.
- van der Veen, Hendrik. 1915. *De Noord-Halmahera'se Taalgroep tegenover de Austronesiese talen*. Leiden: Van Nifterik.
- Visser, L. E. & C. L. Voorhoeve. 1987. *Sahu-Indonesian-English Dictionary and Sahu Grammar Sketch*. Dordrecht: Foris.
- Voorhoeve, C. L. 1988. The languages of the northern Halmaheran stock. *Papers in New Guinea Linguistics*, no. 26., 181-209. (Pacific Linguistics A-76). Canberra: Australian National University.
- Wimbish, Sandra G. 1991. *An introduction to Pagu through the analysis of narrative discourse*. MA thesis, University of Texas at Arlington.

The preceding document was presented at the Tenth International Conference on Austronesian Linguistics (10-ICAL). To properly reference this work, please use the following format:

<LastName>, <FirstName>. 2006. <PaperTitle>. Paper presented at Tenth International Conference on Austronesian Linguistics. 17-20 January 2006. Puerto Princesa City, Palawan, Philippines.
<http://www.sil.org/asia/philippines/ical/papers.html>

For other papers that were presented at 10-ICAL, please visit <http://www.sil.org/asia/philippines/ical/papers.html>.