VOICE IN AUSTRONESIAN

edited by

MARIAN KLAMER

Voice in Austronesian

EDITORIAL

The present volume is the thirty ninth of the Series *NUSA*, *Voice in Austronesian*.

Series *NUSA* volume 1 (1975) through volume 11 (1981) appeared nonperiodically, but from volume 12 (1982) to volume 29 (1987) *NUSA* was issued periodically three times a year. Beginning 1988 the Series do not appear periodically, as it was in 1975 through 1987.

Starting with volume 25 (1986), *NUSA* is type-set differently. There is only one column on each page and the right margin is adjusted. We hope that the improvement will bring *NUSA* more to your expectations.

The Series focuses on works about Indonesian and other languages in Indonesia. Malaysian and the local dialects of Malay will be accepted, but languages outside regions will be considered only in so far as they are theoretically relevant to our languages. Reports from field work in the form of data analysis or texts with translation, book reviews, squibs and discussions are also accepted. Papers appearing in *NUSA* can be original or translated from languages other than English.

Although our main interest is restricted to the area of Indonesia, we welcome works on general linguistics that can throw light upon problems that we might face. It is hoped that *NUSA* can be relevant beyond the range of typological and area specializations and at the same time also serve the cause of deoccidentalization of general linguistics.

Contributions from linguists in the field are solicited. Guest editors are also welcome.

VOICE IN AUSTRONESIAN

Edited by

MARIAN KLAMER

1996
Badan Penyelenggara Seri *NUSA*
Universitas Katolik Indonesia Atma Jaya
Jakarta

TABLE OF CONTENTS

INTRODUCTION

Since Bloomfield (1917) the 'focus' system of the Austronesian language Tagalog has been part of traditional linguistic knowledge as a complex — and exotic — system of so-called 'multiple passivization'. In this system, each transitive clause contains one argument which is placed 'in focus', so that the orientation of the clause is towards this particular argument. Traditionally, it is assumed that the choise of a focus is determined by discourse considerations: focussed NPs are old information.

Which argument is 'in focus' is formally expressed in two ways, by (1) a set of verbal affixes which differ according to the function of the focused nominal (e.g. Agent, Patient, Instrument, etc.) and (2) a set of function and focus markers on NPs. The Tagalog sentences in (1) (taken from Blust, this volume) exemplify the system (more constructions are possible). In these sentences, the verbal affixes *um, in, án* signal whether either the Agent, or the Patient or the Location NP is 'in focus' while the focused NP itself is marked with the special focus morphemes *si* (for proper names) or *ang* (for other nominals):

(1) a. *B-um-ilí*   *si* *Maria ng*   *tinapay sa*   *tindahan*
       buy-Ag.Focus   FOCUS Mary PATIENT   bread LOCATION store
       'MARY is buying/bought some bread at the store'

    b. *B-in-ilí*   *ni* *Maria* *ang* *tinapay sa tindahan*
       buy-Pat.Focus/perfective AGENT Mary   FOCUS bread LOCATION store
       'Mary bought THE BREAD at the store'

    c. *B-in-ilh-án*   *ni Maria*   *ng tinapay*   *ang* *tindahan*
       buy-perfective-Loc.Focus AGENT Mary PATIENT bread FOCUS store
       'Mary bought some bread at THE STORE'

'Focus' or 'voice' systems similar to the Tagalog system can be found in Austronesian languages throughout Taiwan, the Phillipines, the northern parts of Borneo and Sulawesi and in remote locations such as Micronesia (Chamorro).

The analytical questions raised by such focus/voice systems are the subject of ongoing debate (cf. the references given in the papers by Richards and Voskuil in this volume) and gave occasion to a workshop with the topic 'Voice in Austronesian', organized by Jan Voskuil at Leiden University. The workshop was part of the 28th Annual Meeting of the Societas Linguistica Europea and took place in Leiden on 1 September 1995. On the workshop, ten papers were presented on voice and/or focus systems in a variety of Austronesian languages ranging from the Western part of the archipelago to the East.

The fully developed focus/voice system of the Western Malayo Polynesian languages

such as Tagalog with its alleged 'multiple passive' is reduced the more one gets to the East. Borneo and Sulawesi seem to be a transitional area, having languages with both full and reduced focus/voice systems, while in languages like Sasak and Kambera spoken on, respectively, the islands of Lombok and Sumba more to the east, the complex formal properties of the focus/voice system are (almost) lost. The Oceanic languages form a separate group with a different voice system altogether.

Four of the papers that were presented at this workshop were submitted for the present *NUSA* Volume. The title of this volume, *Voice in Austronesian*, has no other pretensions than to reflect the common theme of the papers in this volume — all four address a topic that relates to Voice in Austronesian.

Norvin Richards and Jan Voskuil, the authors of the final two papers of the present volume, both address related aspects of the focus/voice system in Tagalog. We can distinguish three distinct topics that are relevant in the discussion of the Tagalog system:

(2) a. The syntactic properties of the "focussed" nominals: are they subjects or topics?
  b. The nature of the morphemes marking those nominals: are they case markers or topicalization markers?
  c. The status of the verbal affixes signalling the focussed NPs: are they passive morphemes or lexical derivational affixes?

In his paper, Richards examines the syntactic properties of the "focussed" NPs, that is topic (2a). He presents a number of arguments why the "focussed" NP undergo topicalization which is *not* case-driven and similar to topicalization in V2 languages like Icelandic and German. With respect to (2b) he therefore concludes that the morphemes *ang* and *si* that mark "focussed" NPs cannot be subject markers of nominative (e.g. Guilfoyle, Hung and Travis 1992) or absolutive case (e.g. Maclachlan and Nakamura 1994) but rather are topicalization markers, thus agreeing in spirit with e.g. Shibatani (1988).

Voskuil briefly takes up this issue by noting an argument in favour of the "subject marking" contra the "topic-marking" function of *ang* and *si*, namely that the range of possible topics in Tagalog is much smaller than in V2 languages like German, Icelandic or Dutch because Tagalog PPs and adjunct NPs are not possible topics. However, the focus of Voskuil's paper is on the status of the verbal affixes that mark the focussed NPs (2c). Traditionally, these affixes are considered to mark many different kinds of passive voices. In contrast to this, Voskuil argues that Tagalog has only a single passive alternation. The different verbal affixes are not part of the passive-non-passive system of the language but have a lexical derivational function. His analysis posits two levels: a "lexical" and a "voice" level. The affixes are part of a lexical word-formation process and reflect the semantic relation between the verb and its "deep object", while the voice system promotes the object to subject. His argument is based on a comparison of affixes in Indonesian and Dutch having a similar function as the Tagalog "Locative focus" affix *-an* (cf. (1c) above).

The reconstructed ancestor of this very same Tagalog affix is the Proto-Austronesian 'locative' suffix *-an, discussed in the paper by Robert Blust, the first paper in the volume. Comparing data from a wide variety of languages, Blust argues that Proto Austronesian *-an cannot be classified as 'locative' under any universal definition of this term. He also raises some objections to Starosta, Pawley and Reid's (1982) claim that the affixes in question were originally nominalizing with a later developed verbalizing function.

As noted above, the complex focus/voice system that we find in many Western Malayo Polynesian languages becomes less complex the further we get to the east. In the Central Malayo-Polynesian languages Leti (Van Engelenhoven 1995), Rotinese (Jonker 1915), and Kambera have no passive constructions. The paper by Marian Klamer discusses the lack of a passive construction and verb form in Kambera and the two alternative strategies that the language employs as a functional equivalent to the passive in other languages. Kambera relative structures are shown to be related to passive-like structures because the marker *pa* that introduces an object relativization can be reanalysed as a passive morpheme in the language, which suggests that passives may be structurally more related to relatives than the standard accounts of these structures often suggest.

The papers are ordered according to alphabet.

REFERENCES

ENGELENHOVEN, A. VAN. 1995. *A description of the Leti language (as spoken in Tutukei)*. Dissertation, Leiden University.

GUILFOYLE, E., H. HUNG, AND L. TRAVIS. 1992. 'Spec of IP and Spec of VP: two subjects in Austronesian languages.' *Natural Language and Linguistic Theory* 10. 375-414

JONKER, J.C.G. 1915. *Rottineesche spraakkunst*. Leiden: Brill.

MACLACHLAN, A. AND M. NAKAMURA. 1994. 'Case checking and specificity in Tagalog.' ms., McGill University.

SHIBATANI, M. 1988. 'Voice in Phillipine languages.' In M. Shibatani (ed.) *Passive and voice*, 85-142. Amsterdam and Philadelphia: John Benjamins.

STAROSTA, STANLEY, ANDREW K. PAWLEY AND LAWRENCE A. REID. 1982. 'The evolution of focus in Austronosian.' In: Amran Halim, Lois Carrington and S.A. Wurm (eds). *Papers from the Third International Conference on Austronesian Linguistics,* Vol. 2:145-70.

Marian Klamer
September 1996

NOTES ON THE SEMANTICS OF PROTO-AUSTRONESIAN *-an 'locative'*

Robert Blust
University of Hawaii

Many of the attested Austronesian languages of Taiwan, the Philippines and some other areas have complex systems of affixation which serve both case-marking (or 'focus-marking') and nominalizing functions. In their case-marking functions these affixes have prototypical uses which permit a fairly ready semantic characterization. However, the range of functions performed by 'focus-marking' affixes is surprisingly broad, and close attention to the comparative evidence shows that in the common ancestor of most attested focus languages some uses must have diverged sharply from those considered typical of their category. This paper examines some uses of Proto-Austronesian *-an 'locative' which cannot readily be classified as locative under any universal definition of this term. Lastly, it raises objections to the claim that the affixes in question had exclusively nominalizing functions in Proto-Austronesian, and implies that the line between inflection and derivation is less clear-cut in these languages than in languages such as English.

1. Introduction

The 'focus' systems of Austronesian languages such as Tagalog have been generally known to linguists at least since the work of Bloomfield (1917). Early interpretations of these systems found them highly exotic, as they appeared from an Indo-European perspective to offer multiple possibilities of passivization.

Wolff (1971) compared the focus-marking morphology of modern Formosan and Philippine languages, and on the basis of systematic correspondences of phonemic form and grammatical function he inferred a similar system for their immediate common ancestor, Proto-Austronesian (PAN). The core of Wolff's reconstructed focus system includes the following affixes: 1. *-um- 'Actor Focus (AF)'; 2. *-en 'Direct Passive', 3. *-an 'Local Passive', 4. *i- 'Instrumental Passive'. The perfective marker *-in- could co-occur with any of these voice markers in combinations that need not concern

us here, beyond noting that *-en is expressed with a zero allomorph in stems infixed with *-in-. For purposes of the following discussion I will adopt a slightly different terminology, using 'Patient Focus' (PF) for *-en, 'Locative Focus' (LF) for *-an, and 'Instrumental Focus' (IF) for *i-, which we now know was *Si- in PAN, but *i- in Proto-Malayo-Polynesian, the hypothetical ancestor of all non-Formosan AN languages (Blust 1995). An introduction to the operation of focus systems will be useful before going further. Given the central purpose of this paper such an introduction will necessarily be brief, informal and incomplete.

Focus affixes are attached to the verb stem, and signal the particular relationship which the focused NP bears to the verb. The focused NP is marked by a particle which typically distinguishes common from personal nouns (in Tagalog /ang/ vs. /si/); non-focused NPs are marked by a different particle: in Tagalog these are /ni/ for personal nouns, /sa/ for locative nouns and /nang/ (conventionally written /ng/) for most others. Consider the following sentences:

(a) *B-um-ilí   ang   lalake   ng   tinapay sa   tindahan*
    buy-AF FOCUS man   PATIENT bread LOCATION store
    'THE MAN is buying/bought some bread at the store'

(b) *B-um-ilí   si   Maria   ng   tinapay sa   tindahan*
    buy-AF FOCUS Mary PATIENT bread LOCATION store
    'MARY is buying/bought some bread at the store'

(c) *Bilh-ín   ni   Maria   ang tinapay sa   tindahan*
    buy-PF   AGENT Mary FOCUS bread LOCATION   store
    'Mary is buying THE BREAD at the store'

(d) *B-in-ilí   ni   Maria   ang   tinapay   sa tindahan*
    buy-PF/PERFECTIVE AGENT Mary FOCUS bread   LOCATION store
    'Mary bought THE BREAD at the store'

(e) *B-in-ilh-án   ni   Maria   ng   tinapay   ang   tindahan*
    buy-perfective-LF AGENT Mary PATIENT bread   FOCUS store
    'Mary bought some bread at THE STORE'

To summarize: in (a) *ang lalake* 'the man' is the focused NP, and hence the actor (signalled by -*um*-); in (b) *si Maria* 'Mary' plays a similar role. However, in (c) and (d) it is *ang tinapay* 'the bread' which is focused, and hence the patient (signalled by

suffixed -*in* in the non-perfective and by infixed -*in*- without an overt focus affix in the perfective), while in (e) it is *ang tindahan* 'the store' which is focused, hence the location of the action of buying (signalled by -*an*). The choice of focus appears to be governed by discourse considerations: focused NPs are old information (hence definite), and the focus apparatus serves as a means of tracking reference above the level of the sentence.

Focus systems are found in a more-or-less solid block among the AN languages of Taiwan, the Philippines and northern portions of Borneo and Sulawesi. However, they also appear in scattered locations elsewhere, most notably in Chamorro of the Marianas Islands (western Micronesia), a language with no close subgrouping ties, and in Malagasy, a languages whose closest relatives in southeastern Borneo have lost almost all traces of their earlier focus morphology.

Many languages which do not have a 'focus' syntax nonetheless contain either formal or functional traces of such a system. Formal traces of an earlier focus system appear in morphemes which have lost a tightly integrated grammatical function and appear instead as derivational affixes used in nominalization processes, or as relatively unintegrated particles which have little or no grammatical function. To cite two representative examples, Kelabit of northern Sarawak has a reflex of *-an 'LF', but it no longer functions to mark location as subject, and Malay has a reflex of *si, but it no longer functions to mark a personal subject as actor.

Kelabit has evolved a syntax with a simple active-passive contrast, the active marked by ŋ-, with a variety of surface realizations determined by stem-initial segment, and the passive by -*en* (in the perfective by -*in*-, also with several surface realizations). During fieldwork in the 1970's I recorded a few instances of stems suffixed with -/an/ which may function verbally: *bala* 'news, fame', *mala* 'to say, tell' vs. *bela-an* 'saying, telling', *kedha* 'able to withstand pain', *ŋedha* 'to suffer' vs. *kedaan* 'suffering'. However, in most cases -*an* forms nouns of location: *dalan* 'road, path', *nalan* 'to walk' vs. *delan-an* 'path made by repeated walking over the same course'; *guta* 'wading across a river' vs. *geta-an* 'fording place'; *irup* 'what is drunk; way of drinking', *m-irup* 'to drink' vs. *rup-an* 'watering hole for animals in the jungle'.

Like Kelabit, Malay has evolved a syntax with a simple active-passive contrast. Like Tagalog, Malay uses *si* as a marker of personal names, as in *si Ahmad* 'Ahmad'. However, unlike the situation in Tagalog, Malay *si* has no syntactic function. In particular, it cannot be used to distinguish actors from patients: *si Ahmad ber-kata apa ke-pada kamu?* 'what did Ahmad say to you?'; *kamu ber-kata apa ke-pada si Ahmad?* 'what did you say to Ahmad?'. Without cognate forms from other languages the origin of such a particle would be quite obscure, as its function is minimal (Macdonald and Soenjono 1967:125 suggest that it carries a mildly denigrating connotation, especially

when used with descriptive terms to form nicknames, as with *si gemuk* 'Fatso', or *si kurus* 'Bones').

As devil's advocate, one could argue that the derivational function of affixes such as *-an* or the grammatically unintegrated use of particles such as Malay *si* represents the original situation in AN, and that languages such as Tagalog have integrated these less tightly bound morphemes into a coherent system of verbal case-marking. In effect this is what Starosta, Pawley and Reid (1982) have done: because agents (the actors of non-AF, or passive verbs) and possessors carry the same morphological marking in many AN languages, non-AF verbs can be interpreted as nouns: *kapatíd ni Juan* 'John's sibling'; *b-in-ilí ni Juan ang tinapay* 'John bought the bread' or 'John's buying/what John bought was the bread'. Where agency or genitive is marked by an overt morpheme rather than solely by structural relations that morpheme can be glossed either 'of' or 'by', as with Tagalog *ni* or its cognates in most other AN languages. Starosta, Pawley and Reid have exploited this ambivalence to argue that focus systems are historically secondary, and that the functions of the focus affixes in PAN were exclusively nominalizing. I will not enter into details here, but it is clear from several lines of evidence that languages such as Tagalog represent the original grammatical type, and that cognate morphemes in languages such as Kelabit or Malay represent the remains of what was once a more highly integrated morphological system which functioned to signal case-marking and discourse reference. Moreover, it is clear from virtually all extant focus systems that a given affixed stem could function either verbally or nominally, dependent upon context. Simple application of the Comparative Method, and argument from parsimony favors the interpretation that PAN also had a focus system in which the focus affixes functioned either verbally or nominally in given contexts.

My purpose in the following discussion is to draw attention to some of the more striking non-canonical uses of the 'locative focus' marker in early AN proto-languages (and in many of their descendants). By 'non-canonical' I mean uses that cannot under any rule-governed definition be construed as locative, and which in some cases appear far more transparently to involve other case relations (such as instrumental). While the focus of my attention will be on PAN *-an* 'LF' and its reflexes, many of the observations that I will make can be applied to the other oblique focus affixes as well.

2. PAN *-an

Granted that it is descriptively convenient to characterize *-an* and its reflexes in modern languages as signalling a locative relationship between focused NP and the verb, it is important to realize that this characterization selects a frequently recurrent type of relationship to represent a semantically much broader and more diffuse category. In the following section I will try to sketch some idea of how broad and diffuse this category really is.

2.1. *-an* as source. Some Formosan aboriginal languages reflect *RiNaS, and others *RiNaS-an for various native pheasant species which within the ethnographic present were valued for their long tail feathers:

(1)   Bunun /Linas/,
         Saisiyat (Taai) /Lilaɪan/,
         Thao /lhizash-an/ 'pheasant'

Both Kavalan and Thao, two languages which probably have not had a common ancestor since the dispersal of the PAN speech community on Taiwan around 4000 BC reflect the shorter form in the meaning 'long tail feathers of a bird': Kavalan *linas* 'long and beautiful tail feather' (Tsuchida 1994), Thao *lhizash* 'long feathers of a bird or fowl'. The simplest explanation of these cognate distributions is a hypothesis that PAN *RiNaS meant 'long tail feather' and that *RiNaS-an meant 'pheasant'. The essential correctness of the foregoing interpretation is corroborated by another comparison with two subparts:

(2) a.   PAN *waNiS 'tusk of a boar' :
         Thao /wazish/ 'the Formosan wild boar: *sus scrofa taivanus* (Swinhoe)',
         Bunun /vanis/ 'tooth, tusk',
         Tsou /hisi/ 'tooth', /hisno tacvohi/ 'large tusks of a boar' (Szakos 1993),
         Rukai (Budai) /valisi/ 'tooth, tusk',
         Paiwan /aLis/ 'tooth, fang, tusk',
         Puyuma /wali/ 'tooth, tusk of a boar',
         Taokas /walis/ 'armlet' (Tsuchida 1982).

(2) b.   PAN *waNiS-an 'wild boar' :
         Saisiyat /waliɪan/,
         Rukai (Tona) /vaLisanə/ 'wild boar'

The PAN words for 'pheasant' and 'wild boar' are notable for their morphological parallelism: in each word the name of an animal is formed from the name of an economically valued product obtained from that animal, suffixed with *-an. In effect the pheasant was called 'source of the long tail feathers' (valued as head ornaments) and the wild boar 'source of the tusks' (valued as armbands). In these cases rather than signalling the place in which an action occurred, *-an signalled the 'place' from which an animal product was obtained.[1]

2.2. *-an as instrument. A number of AN languages both within and outside Taiwan reflect PAN *CapeS, Proto-Malayo-Polynesian (PMP) *tahep 'to winnow'. Reflexes of both *tahep and *tahep-an are found in a number of Philippine languages and in Kayan of central Borneo:

(3) a.   *tahep:
         Ilokano /taép/ 'chaff, glume, husk',
         Tagalog /tahíp/ 'up and down movement of rice grains being winnowed on a
         flat basket',
         Cebuano /tahúp/ 'chaff of cereals; separate the husk from husked grains',
         Kayan /tap/ 'to winnow a tray of grain'

(3) b.   *tahep-an:
         Ilokano /taep-án/ 'to winnow',
         Tagalog /táhip-an/ 'the flat basket used in winnowing',
         Cebuano /taph-an-án/ 'ground cereals to be winnowed; place where
         winnowingis done',
         Kayan /tap-an/ 'tray for winnowing rice'

In addition, a number of other languages in Borneo reflect *tahep-an with loss of the morpheme boundary in the meaning 'winnowing basket', and form the verb 'to winnow' in an innovative manner, through the use of homorganic nasal substitution:

(3) c.   Kenyah, Murik /tapan/ 'winnowing basket' : /napan/ 'to winnow'

---

[1] Chen (1988), who is generally quite thorough, mentions neither pheasant feathers nor boar tusks as raw materials valued by the Formosan aborigines. However, Chai (1967:68) contains a photograph of a Tsou man adorned with a headdress that appears to contain the tail features of Swinhoe's blue pheasant or a similar species, and the etymology of the Taokas word for 'armlet' speaks for itself.

These comparisons clearly support the reconstruction of a verb-base *tahep 'winnow' and a nominal derivative *tahep-an 'winnowing basket'. What is of interest in the present context is the use of *-an in the latter form, since *tahep-an seems clearly to be an instrument rather than a location in the sense in which that term is normally understood. The reflex of *tahep-an in Cebuano Bisayan (with double suffixation) shows the less marked, or more commonly expected semantics ('place where winnowing is done'), but this appears to be a secondary regularization of the semantics to conform with the usual semantic outcome of nominalizations with -an.

Universal semantic roles in case grammar are not explicitly defined, but it is at least intuitively unexpected that a winnowing basket would be conceived as the (stationary) place where winnowing is done rather than as the instrument with which winnowing is done, since the instrument can be moved from place to place. To date I have found no clear parallels for this usage, but its value lies in its distribution over a number of languages.

2.3. *-an as state. In the two previous examples PAN *-an or its reflex in various modern languages was shown sometimes to mark source and sometimes instrument rather than serving its expected function of marking location. Apart from these apparent cross-overs of case-marking function *-an also appears in a few cases to usurp the function of PAN *ma- 'stative'.

Word classes in PAN, PMP and many attested AN languages include verbs, nouns and various particles, but no distinct class of adjectives. Rather, verbs are distinguished as dynamic (eat, walk, swim, dream) or stative (afraid, sleep, living/alive). The latter were originally marked with the prefix *ma-, which is retained in many modern languages. The Long Semado dialect of Lun Dayeh in northern Sarawak illustrate this:

(4)   dynamic verbs:
         /ŋe-lak/ 'to cook (rice)',
         /ŋe-buen/ 'to smell something',   /ŋ-abet/ 'to bite'

(5)   stative verbs:
         /me-buen/ 'smelly, malodorous',
         /m-ulun/ 'living, alive',
         /me-berat/ 'heavy',
         /me-birar/ 'yellow',
         /me-kapal/ 'thick, as a board'

PAN *tiaN 'abdomen', PMP *tian is widely reflected throughout island Southeast Asia

and the Pacific (6a). Next to PMP *tian we also find *tian-an 'pregnant' (6b):

(6) a.   PAN *tiaN, PMP *tian:
          Thao /tiaz/ 'abdomen above the navel'
          Puyuma /tial/ 'abdomen'
          Ilokano /tián/ 'abdomen, belly; womb, uterus'
          Tagalog /tiyán/ 'abdomen'
          Sasak /tian/ 'belly'; /be-tian/ 'pregnant' (lit. 'having a belly')
          Banggai /tia/ 'womb; belly'
          Palauan /díil/ 'abdomen; womb'
          Elat /tian/ 'belly'
          Roviana /tia-/ 'abdomen'
          Lau /i:a/ 'belly'
          Sa'a /ie-/ 'belly, stomach, bowels, womb'

(6) b.   PMP *tian-an:
          Banggai /tian-an/ 'pregnant'
          Makasarese /tian-aŋ/ 'pregnant'
          Lau /i:an-a/ 'pregnant'
          'Are'are /iana/ 'pregnant'
          Mota /tiana/ 'pregnant'

In some languages, as Banggai and Lau it can be argued that the original morpheme boundary has shifted one segment to the left (hence /tia-nan/, /i:a-na/), but it is clear from the overall comparative picture that the PMP word for 'pregnant' consisted of the word for 'abdomen; womb' suffixed with *-an. Contrary to what we might expect from the general pattern for lexical items describing a state or condition, then, the PMP word for 'pregnant' was not marked with the stative prefix *ma-, but rather with the 'locative' suffix *-an, and its conceptual structure was something like 'in the womb'.

    Somewhat similar, but different in detail is a cognate set of more restricted distribution which is found in several of the languages of western Indonesia. Next to PMP *bulu 'body hair; fur; feather; down; floss on plant stems; color; type, kind' there is evidence for *bulu-an 'hairy; kind of hairy fruit, rambutan : *Nephelium lappaceum*':

(7)   PWMP *bulu-an:
          Sundanese /bulu-an/ 'hairy, shaggy, feathered'
          Old Javanese /wulw-an/ 'kind of hairy fruit and its tree: *Nephelium* rambutan'

          Sasak /bulu-an/ 'hairy; a hairy fruit, the rambutan, *Nephelium lappaceum*'
          Tae' /bulu-an/ 'a fruit, the rambutan, *Nephelium lappaceum*'
          Wolio /bulu-a/ '(head of) hair; kind of fruit tree, rambutan, *Nephelium lappaceum*'

What is noteworthy about this comparison is that it parallels *tian-an in its use of *-an to signal a state or condition ('hairy'), but parallels *RiNaS-an and *waNiS-an in using *-an to form the name of a natural species (as does its Malay equivalent /rambutan/, from /rambut/ 'hair'). However, unlike the Formosan animal terms which are based on the names of culturally valued parts of the animal (long tail feathers, tusks), the *bulu-an was not a plant valued for its hairy skin. On the contrary, it is a succulent fruit valued for its edible flesh (similar to that of a lychee); the hair-like appendages covering its skin are simply a visually striking means of identification. Another example of this type is the name of the well-known durian (< *duRi 'thorn' + -an).

2.4. Unclassified. One other well-known idiosyncratic use of *-an is seen in the word for 'gold'. A number of languages reflect PAN *bulaw 'golden colored', while some of these and others also support PAN *bulaw-an 'gold':

(8) a.   *bulaw:
          Kavalan /bulaw/ 'golden color of ripening rice'
                         (loan from an unknown source)
          Ilokano /boláw/ 'cock with dark brownish-yellow or drab plumage'
          Tagalog /buláw/ 'reddish; reddish-gold; golden orange
                         (said of young pigs and roosters)'
          Maranao /bolao/ 'blond, as hair'
          Ngaju Dayak /bulaw/ 'gold'

(8) b.   *bulaw-an:
          Kavalan /berawan/,
          Itbayaten /voxawan/,
          Maranao /bolao-an/,
          Bare'e /wuyawa/ 'gold'
          Tae' /bulaan/ 'gold; noble, precious, lofty, holy'
          Paulohi /hulawane/,
          Numfor /brawen/ 'gold'

The morphological relationship between these forms appears to be beyond dispute.

What is curious is the manner in which *-an is used to form a concrete noun from an unaffixed stem which was either an abstract noun or an attributive. This usage has no clear parallels among the other cases cited here.

3. Verbalization vs nominalization?

As noted briefly above, Starosta, Pawley and Reid (1982) have proposed that the focus-marking affixes of Austronesian languages originally had exclusively nomina-lizing functions. The attested verbal functions of these affixes in scores of languages throughout Taiwan, the Philippines, northern portions of Borneo and Sulawesi, and in such outlying areas as Chamorro of western Micronesia and Malagasy, are said to have arisen through syntactic analogy with equational constructions. Although proportional analogy is an established part of the theory of language change, and has demonstrated its value in the solution of many particular problems in both phonology and morphology, syntactic analogy is far less commonly recognized. The attempt to use such a concept in accounting for the evolution of focus in Austronesian languages is thus simultaneously an attempt to explain a particular set of observations and to elaborate *de novo* the theoretical machinery essential to that explanation.[2]

REFERENCES

ANDERSON, J.M. 1992. 'Linguistic representation: structural analogy and stratification.' In: Werner Winter (ed) *Trends in Linguistics Studies and Monographs* 67. Berlin: Mouton de Gruyter.

ANTTILA, R. 1977. 'Analogy.' In: Werner Winter (ed) *Trends in linguistics State-of-the-Art Reports*. The Hague: Mouton.

---

[2] Lehmann (1992:228ff) maintains that for analogy to take place 'some linguistic set is necessary', and this may be phonological, morphological, syntactic or semantic. Most other writers on analogy, however, restrict their discussions to phonology and morphology. Anttila (1977) treats analogy in considerable detail, but does not apply the concept to syntax, Skousen (1989,92) argues that language behavior can be predicted with far greater success by analogical models than by structural models, but he makes no direct reference to syntax, and Anderson (1992) speaks of 'structural analogy', but appeals to it as a principle mediating levels of representation in linguistic descriptions rather than as an extension of proportional analogy from the well-attested domains of phonology or morphology. Finally, in a detailed discussion of principles of diachronic syntax Lightfoot (1979) is silent on analogy as a possible mechanism of syntactic change.

BLOOMFIELD, L. 1917. *Tagalog texts with grammatical analysis*. [University of Illinois studies in language and literature 3.2-4.]. Urbana, Illinois: University of Illinois.

BLUST, R. 1995. 'The position of the Formosan languages: method and theory in Austronesian comparative linguistics.' In: Paul J-K Li, et al., (eds). *Austronesian Studies Relating to Taiwan* [Symposium Series of the Institute of History and Philology, Academia Sinica, No. 3.]: 585-650. Taipei: Academia Sinica.

CHAI, CHEN-KANG. 1967. *Taiwan aborigines: a genetic study of tribal variations*. Cambridge, Massachusetts: Harvard University Press.

CHEN, CHI-LU. 1988 [1968]. *Material culture of the Formosan aborigines*. Taipei: Southern Materials Center, Inc.

HIMMELMANN, N. 1987. *Morphosyntax und Morphologie --- Die Aufrichtungsaffixe im Tagalog*. [Studien zur Theoretischen Linguistik, vol. 8.]. München: Wilhelm Fink.

LEHMANN, W.P. 1992. (3rd edition). *Historical linguistics: an introduction*. London and New York: Routledge.

LIGHTFOOT, D.W. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.

MACDONALD, R. R. AND S. DARDJOWIDJOJO. 1967. *A student's reference grammar of modern formal Indonesian*. Washington, D.C.: Georgetown University Press.

SKOUSEN, R. 1989. *Analogical modeling of language*. Dordrecht: Kluwer Academic Publications.

SKOUSEN, R. 1992. *Analogy and structure*. Dordrecht: Kluwer Academic Publications. STAROSTA, S., A.K. PAWLEY AND L.A. REID. 1982. 'The evolution of focus in Austronesian.' In: Amran Halim, Lois Carrington and S.A. Wurm (eds). *Papers from the Third International Conference on Austronesian Linguistics*, Vol. 2:145-70.

SZAKOS, J. 1993. 'Tsou-Deutsches Wörterbuch.' Computer print-out.

TSUCHIDA, S. 1994. 'Kavalan-English vocabulary.' Computer print-out.

TSUCHIDA, S. (compiler). 1982. *A comparative vocabulary of Austronesian languages of Sinicized ethnic groups in Taiwan, part I: West Taiwan*. [Memoirs of the Faculty of Letters, University of Tokyo, No. 7.]

WOLFF, J.U. 1971. 'Verbal inflection in Proto-Austronesian.' In: Andrew B. Gonzalez (ed). *Parangal kay Cecilio Lopez: essays in honor of Cecilio Lopez on his seventy-fifth birthday*:71-91. [Philippine Journal of Linguistics Special Monograph Issue No. 4.] Quezon City: Linguistic Society of the Philippines.

Department of Linguistics
University of Hawaii
Moore Hall 569
1890 East-West Road
Honolulu, HI 96822, USA
[blust@hawaii.edu]

KAMBERA HAS NO PASSIVE*

Marian Klamer

Vrije Universiteit/HIL/Royal Academy of Arts and Sciences

The passive construction does not seem to occur in a number of Central Malayo-Polynesian languages in Eastern Indonesia such as Leti, Rotinese and Kambera. In Kambera the functional equivalent of the passive voice in other languages is expressed by using two strategies. The first is a process of focusing/inversion where the syntactic status of the subject and object does not change, the pronominal marking and morphological shape of the verb remains the same, but the adjunct NP that is coreferent to the object is focused while the subject adjunct NP is defocused. The second strategy uses an object relativisation as a nominal predicate. I will argue that the latter structure resembles a 'canonical' passive most closely both in form and function and suggest a structural account of the insight that Kambera provides us with, namely that (object) relative structures are related to passive-like structures because a relative marker can be reanalysed as a passive morpheme.

1. Introduction

Voice distinctions like the active/passive one are a very significant feature of the Western Malayo-Polynesian languages. Strikingly, however, the passive construction does not seem to occur in a number of Central Malayo-Polynesian languages in Eastern Indonesia. Examples are Leti (Jonker 1932, Van Engelenhoven 1995) and Rotinese (Jonker 1915) both spoken in the Timor area, and Kambera, spoken on the island of Sumba.

The aim of the present paper is descriptive: I will discuss the formal strategies that Kambera, as one of the languages lacking a passive, employs as functional equivalents to the passive construction and/or verb form in other languages.

One of the main functional motivations of the passive construction is to defocus the Agent (e.g. Shibatani 1985). In many languages this has the effect of reducing the valency of the verb. The standard structural account for this is to assume (in generative terms) that the external (AGENT) theta-role and structural (objective) case are absorbed by the passive morphology of the verb. Since surface subjects are obligatorily present in syntax, this will force the remaining NP, the PATIENT, to become the subject, i.e. the NP with the internal theta-role moves to subject position because the only way it can survive the case filter is by moving to a position where it can get (nominative) case.

Lacking both a syntactic construction and a morphological verb form for the passive, Kambera employs two other constructions to defocus an Agent or foreground a Patient. The first is to change the order of the NP constituents within a clause, a type of adjunct NP-inversion or focusing. An illustration is given in (1), where the distinction between the two sentences lies in the fact that the object NP *na lau* 'the sarong', postverbal in (1a), has 'changed places' with the subject NP *nyuna* 'she'[1] in (1b) without a change in the morpho-syntactic form of the verbal cluster *na-tinu-nya*:

(1) a. *Ka   nyuna$_j$ na-$_j$ tinu   -nya$_k$ na   lau$_k$*
      CNJ   she   3SN-   weave   -3SD   ART   sarong
      'So that she weaves the sarong (lit. she she-weaves-it the sarong)'

    b. *Ka   na   lau$_k$ na-$_j$ tinu   -nya$_k$ nyuna$_j$*
      CNJ   ART   sarong   3SN-   weave   -3SD   she
      'So that the sarong was woven (by her) (lit. the sarong she-weaves-it she)'

This focusing process is distinct from so-called topicalisation/left-dislocation in Kambera. The latter process involves moving arguments to non-argument positions at the periphery of the clause (cf. section 2 below), putting contrastive stress on the dislocated NP and an intonational break between the dislocated NP and the rest of the clause. In what follows I will use the term left-dislocation (not topicalisation) for this process, reserving the term 'topic' for the information status of an NP in discourse (cf. section 3 below).

The second construction used to background the subject and foreground the object is

---

a relative construction as the nominal predicate of a matrix clause. Kambera has two types of relativisations: subject relativisations that are marked with the proclitic *ma-* and object relativisations marked with the proclitic *pa-*.[2] In object relativisations the object of the embedded verb is the head of the relativisation, the relativised noun that is the obligatorily missing argument (a 'gap') in the relative clause. More than focusing, it is the object relativisation that is functionally most related to a passive in other languages, as we will see below. An illustration of its use is given in (2):

(2)   *Pa-   palu   -ka*
        RM-   hit   -1sA
        'I was hit (lit. (the one) hit (was) I)'

This paper is organised as follows. In section 2 I provide some basic information about Kambera sentence structure and about how the language marks subjects and objects on the verb. In section 3 I discuss how focusing is used to convey voice distinctions. Next, I will discuss how relativisations can be used as nominal predicates and thus be the functional equivalent to passives in other languages (section 4). I will conclude by drawing some general conclusions in section 5.

2. General information on the language

Kambera is a head-marking language (Nichols 1986) in the sense that it has rich marking on the head of the clause — the verb — of pronominal, aspectual and modal clitics. Definite verbal arguments are marked for person, number and case[3] by pronominal clitics on the verb. Definite NPs are crossreferenced on the verb, they are optional and can be used for disambiguation or emphasis. I assume that the pronominal clitics which crossreference these NPs have argument status while the coreferent NPs

---

[2]   In what follows the notions 'subject' and 'object' are used as syntactic, not semantic, notions. The thematic content of the head of a subject relativisation may be AGENT, THEME or POSSESSOR, in an object relativisation it may be the PATIENT, RECIPIENT, BENEFICIARY, MALEFICIARY, GOAL that is relativised. Cf. Klamer 1994: 303-326.

[3]   Abbreviations for pronominal elements: N=Nominative, G=Genitive, D=Dative, A=Accusative; 1s, 2s, 3s =first, second, third person singular; 1p, 2p, 3p = first, second, third person plural.

---

are adjuncts.[4] Kambera word order facts suggest that the verb forms one syntactic constituent together with its clitics and adverbs. I will use the term 'nuclear clause' to refer to this constituent. A nuclear clause may on its own constitute a complete sentence and is thus the core of every clause. If a clause contains (definite) NPs, I assume them to be adjoined to the nuclear clause. These adjunct NP positions are called 'focused position' in the diagram in (3). They are distinct from the 'left-dislocated position' because they are adjacent to the nuclear clause, within the scope of the conjunction, whereas the topicalised position is not:[5]

(3)

                    Sentence
                   /   \
Left dislocated position   Sentence
                           /   \
                    Conjunction   Clause
                                    /   \
                Focused position   Nuclear clause   Focused position
                 for adjoined NPs   *for adjoined NPs*
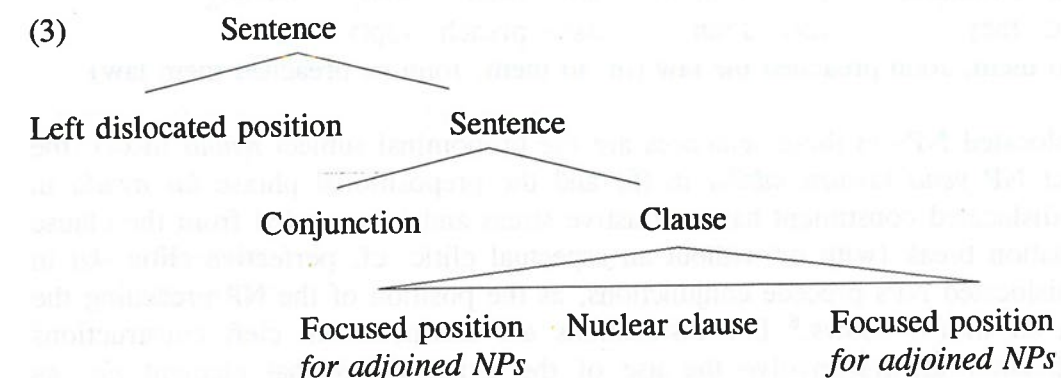
Sentence (1b) above had the definite NP *na lau* in focused position, following the conjunction *ka*. A focused NP does not necessarily get contrastive stress and is not separated from the nuclear clause by an intonational break or a pause. In contrast to this, consider sentences with left-dislocated NPs like (4)-(6):

---

[4]   In this section and the next I will be concerned with definite subject and object NPs only and will not consider indefinite NPs. Definite subject and object NPs are always crossreferenced on the verb and can all be considered adjuncts. The grammatical status of indefinite arguments in Kambera is much less clear cut: indefinite subjects may or may not be crossreferenced in addition to being expressed by an NP. Indefinite objects are never crossreferenced. This, in addition to the restricted number of positions for indefinite NPs in the sentence, seems to suggests that most indefinite subject and objects NPs have argument status, but the details are still unclear.

[5]   In (3) the notion 'sentence' is roughly equivalent to CP, 'clause' to IP and 'nuclear clause' to VP (with subject, negation and aspectual markers).

(4)   *Nyuna$_j$ -ka   na   ma- mayilu -nya$_k$*
           he   -PRF   ART   RM- be first -3sD
           '(It was) he$_j$ who came before him$_k$'

(5)   [*Yena   tawuru mbiha*]$_{NPj}$,   *ka   u- ngàndi -ya$_j$ nú..*
           this one   ring   be magic   CNJ   2sN take   -3sA   DEI
           '(About) this magic ring, you take it (with you)..'

(6)   [*Lai nyuda$_j$*]$_{PP}$,   *i   Yohani   na- peka   -nja$_j$ hukung*
           LOC   they   ART John   3sN- preach -3pD   law
           'To them, John preached the law (lit. to them, John he preached them law)'

The left-dislocated NPs in these sentences are the pronominal subject *nyuna* in (4), the direct object NP *yena tawuru mbiha* in (5) and the prepositional phrase *lai nyuda* in (6). A left-dislocated constituent has contrastive stress and is separated from the clause by an intonation break (with or without an aspectual clitic, cf. perfective clitic *-ka* in (4)). Left dislocated NPs precede conjunctions, as the position of the NP preceding the conjunction *ka* in (5) shows.[6] Left-dislocations are distinct from cleft constructions because the latter always involve the use of the existential verbal element *jia*, as illustrated in (7):

(7)   *Jia   -ya$_j$ [na tawuru mbiha]$_j$*
           EXIST   -3sA ART ring   be magic
           'It is the magic ring (that)....'

Now that we have briefly considered the difference between the two types of position full NPs can occur in, namely focused positions and left-dislocations, we will turn to the structure of the Kambera nuclear clause, the core of every sentence. Consider (8):

(8)   *Nyungga   nda   ku- hili   beli   -ma -nya -pa*
           I   NEG   1sN- again   return   -EMP -3sD -IMPF
           'I am not going back to him again'

---

[6]   The perfective aspect marker *-ka* and the conjunction *ka* have distinct structural and functional properties. Though their homophony may suggest a common ancestor form, synchronically they are distinct grammatical items.

In this sentence, the head of the clause is the verb *beli* 'return', which is modified by the adverb *hili* 'again'. The (optional) subject NP *nyungga* 'I' is crossreferenced on the verb by the nominative proclitic *ku-*. Enclitics are the emphatic marker *-ma*, the pronominal object clitic *-nya* and the aspectual marker *-pa*.[7]

In (9) the structure of the nuclear clause in (8) is given. When we compare the diagram in (9) below to the one in (3) above, we observe again that crossreferenced (adjunct) NPs are not part of the nuclear clause. In (9) the crossreferenced pronoun *nyungga* 'I' occurs in sentence-initial focused position. It precedes the negation *nda*.

(9)   Clause

     Focused position   Nuclear clause
     Full pronoun 'I'

               Modifier   Nucleus
               NEGATION

     Dependent   Modifier   Dependents   Modifier
     Pronom. clitic   Head   *Modal cl.*   *Pronom. cl.*   *Asp. cl.*
     '1s'-   -EMP   -'him'   -IMPF

           Modifier   Head
           Adverb   *Verb*
           'again'   'return'
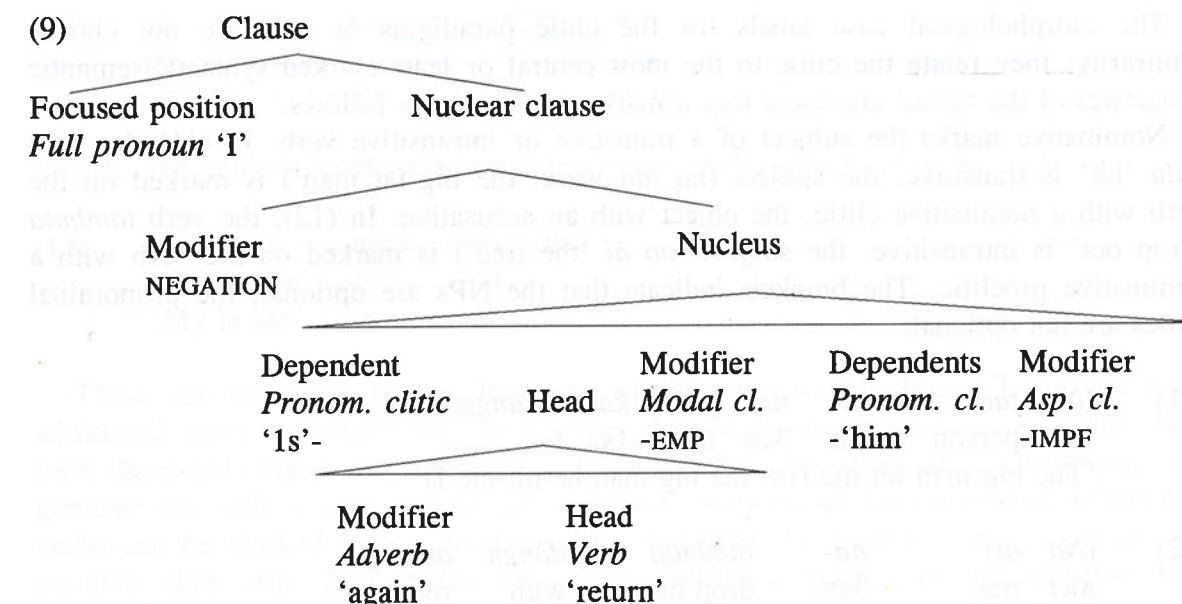
Let us now look at the pronominal clitics more closely. Kambera pronominal clitics are nominative, genitive, accusative or dative,[8] the clitic paradigms are listed in (10):

---

[7]   Enclitics must appear in this order, i.e. modal-pronominal-aspectual. Apart from ordering restrictions like these, Kambera clitics are subject to other idiosyncratic restrictions as well, cf. Klamer (1994, 1996b).

[8]   A verbal argument can also be marked with a combination of these clitics, cf. Klamer 1994; to appear; 1996a,b.

(10)

	NOMINATIVE	ACCUSATIVE	DATIVE[9]	GENITIVE
1s	ku-	-ka	-ngga	-nggu
2s	(m)u-	-kau	-nggau	-mu
3s	na-	-ya	-nya	-na
1p(inc)	ta-	-ta	-nda	-nda
1p(exc)	ma-	-kama	-nggama	-ma
2p	(m)i-	-ka(m)i	-ngga(m)i	-mi
3p	da-	-ha	-nja	-da

The morphological case labels for the clitic paradigms in (10) are not chosen arbitrarily: they relate the clitic to the most central or least marked syntactic/semantic properties of the verbal argument that it marks, which are as follows.

Nominative marks the subject of a transitive or intransitive verb. In (11) the verb *palu* 'hit' is transitive, the subject (*na tau wutu* 'the big/fat man') is marked on the verb with a nominative clitic, the object with an accusative. In (12), the verb *tambuta* 'drop out' is intransitive, the subject (*na ài* 'the tree') is marked on the verb with a nominative proclitic. The brackets indicate that the NPs are optional; the pronominal clitics are not optional.

(11) (*Na tau   wútu*) **na-** palu -ka (*nyungga*)
      ART person be fat **3sN-** hit -1sA I
      'The big man hit me (lit. the big man he-hit-me I)'

(12) (*Na ài*)   **na-** tambuta *dàngu amung*
      ART tree   **3sN-** drop out with   root
      'That tree is uprooted (lit. that tree it is dropped out with root)'

In (13) the ditransitive (applicative) verb *kei(ng)* 'buy X for Y' has two object arguments. Its indirect object is crossreferenced on the verb with the dative clitic *-nja* while the direct object (*rí* 'vegetable') is not marked on the verb because it is indefinite.

---

[9]   Superficially, it seems that the dative paradigm can be derived from the accusative by prenasalising the accusative form. In fact, this prenasalisation is neither a phonological nor a productive morphological process, evidence is given in Klamer 1994. Here I will just assume that the dative clitics in Kambera form indeed a separate paradigm with its own structural and functional properties.

(13) (*I   Ama*) na- kei -**nja** *rí*
      ART father 3sN- buy -3pD vegetable
      'Father buys vegetables for them (lit. Father he buys them vegetables)'

In most cases it is just the indirect object of a ditransitive verb that is marked on the verb. However, it is also possible to cliticize both the indirect object *and* the direct object — in that order — on the verb. This is illustrated in (14), where *-ngga* 'me' marks the indirect object, and *-nya* 'it' the direct object:[10]

(14) (*I   Ama*) na- kei -**ngga** -**nya**
      ART father 3sN- buy -1sD   -3sD
      'Father buys it for me (lit. Father he buys me it)'

Finally, the genitive marks a nominal possessor, such as *-nggu* 'my' in (15):

(15) *Na uma* -**nggu** (*nyungga*)
      ART house -1sG   I
      'My house'

These are the most typical functions of the pronominal clitics. There are some additional functions which are, in a sense, derived from the basic uses as they have been discussed. Those that are relevant to the present discussion are the following. The genitive not only marks a nominal possessor, subjects of transitive and intransitive verbs can be marked with a genitive clitic, too. This is illustrated in (16), where the genitive clitic *-mu* '2sG' marks the subject (semantic AGENT) of the transitive verb *pameti* 'kill', making the clause nominal. Kambera expresses the subject of a clause with a nominal predicate by an accusative clitic that is marked on the NP that functions

---

[10]   Observe that the direct object is marked with a clitic from the dative paradigm here, whereas in (11) it was marked with an accusative; the latter being the unmarked way to mark a direct object, cf. (i):

(i)   *Na- palu -ka*
      3sN- hit -1sA
      'He hit me'

In (14) a dative clitic is used instead of an accusative because of the idiosyncratic Kambera restriction that the second postverbal slot in a clitic cluster may only be filled with a dative clitic.

as the nominal predicate. Kambera does not have an overt copular verb. This is illustrated by (17).

(16) *Pa. meti   -bia   -mu   -ngga*
      CAU.kill -just -2SG -1SD
      '(You) just kill me (i.e. I don't care if you kill me)'

(17) [*Tau   mini*]   *-ya*
      person male   -3SA
      'He/it's a man'

3. Voice effects through focusing/inversion

The diagrams (3) and (9) show that a conjunction is the initial element of a clause. Adjoined nominal constituents appear within the scope of the conjunction, outside the scope of the negation, in positions either preceding or following the nuclear clause. It is these adjoined NP-positions that I will be refering to in this section.

When an NP is focused in Kambera, this has effects that resemble the voice distinctions in other languages. This is illustrated by the sentences in (18). In (18a) the subject NP *na tau wutu* precedes the nuclear clause — the unmarked position for subject NPs. The object NP, following the nuclear clause, also occupies its unmarked position. In contrast to this, (18b) has reversed the order of the NPs: the object NP *nyungga* is fronted and the subject NP *na tau wutu* now follows the nuclear clause. Note again that the morphological marking on the verb does not change.

(18) a. *Ba   na   tau   wútu   na-   palu -ka   nyungga*
         CNJ ART person be fat   3SN- hit -1SA   I
         'The big man hit me (lit. the big man he-hit-me I)'

      b. *Ba   nyungga   na-   palu -ka   na   tau   wútu*
         CNJ I   3SN- hit -1SA ART person be fat
         'I was hit by a big man (lit. I he-hit-me a big man)'

If the function of a passive is to focus a Patient/logical object and defocus an Agent/logical subject, Kambera focusing as illustrated in (18b) is functionally similar to a passive construction in other languages. The basic position for (adjunct) subject NPs is the position preceding the nuclear clause, while the object NP basically follows it.

Reversing the order of the NPs results in an interpretation like a voice alternation where the object is focused and the subject defocused. However, Kambera word order facts are more complex than this, as shown in (19):

(19) a. (*I   Miri$_j$*)   *na-$_j$ kataku -ya$_k$*   (*na   hamayang$_k$*)
         ART Lord   3SN- accept -3SA   ART prayer
         'The Lord accepted the prayer'

      b. *Na   hamayang$_k$*   *na-$_j$ kataku -ya$_k$ i   Miri$_j$*
         ART prayer   3SN- accept -3SA   ART Lord
         'The prayer was accepted by the Lord'

      c. *Na   hamayang$_k$ (i   Miri$_j$)   na-$_j$ kataku -ya$_k$*
         ART prayer   ART Lord   3SN- accept -3SA
         'The prayer was accepted (by the Lord)'

      d. *Na-$_j$ kataku -ya$_k$ i   Miri$_j$ (na   hamayang$_k$)*
         3SN- accept -3SA   ART Lord   ART prayer
         'The prayer/it was accepted by the Lord'

Sentence (19a) illustrates the basic constituent order. The subject NP *i Miri* 'the Lord', is not in focus and therefore precedes the nuclear clause. That is, if it occurs at all. because, as was mentioned above, full definite NPs that are crossreferenced on the verb are optionally used for disambiguation or emphasis (cf. the brackets in (19)).

It is a striking property of Kambera texts that in transitive sentences definite object NPs occur much more often than subject NPs, though grammatically both are possible. Indeed, in most transitive clauses a full subject NP does not occur at all. A possible explanation for this could be the 'discourse topic' status of the subject. Discourse topicality is determined by the information status of an NP in discourse (Foley and Van Valin 1985): if its referent is known, the NP is the topic; if it is new, it is not the topic.

In Kambera, as in many languages, the subject is considered the discourse topic. Thus being known from the context it does not need to be made explicit by an additional NP every time it occurs. An object NP, on the other hand, usually presents new information to the discourse so is much more often fully expressed.

Let us suppose, however, that a sentence has two full adjunct NPs, a subject and an object NP. The order of the NPs in sentence (19b) is the reverse of (19a). But such a reversal does not exhaust the possibilities, as the sentences (19c,d) show: NPs can also

both precede or *both* follow the nuclear clause. What, then, is the difference between (19c,d) and what determines the respective positions of these adjunct NPs?

In this respect it is relevant not only to consider whether the NP is the discourse topic, but also whether the NP is 'in focus', i.e. is particularly salient and/or contrastive compared to the other NP in the sentence. In (20) I summarise the information that Kambera texts provide. The 'topic'/'focus' status of both subject and object NPs of simple transitive clauses is given on the left-hand side, the corresponding sentential position of the NPs is given on the right-hand side. Optionality of NPs is again indicated by brackets.

(20)		Topic	Focus	Sentential position of the NPs
a.	Subject	+		
	Object:	−		(NP_S) nuclear clause (NP_O)
b.	Subject:	−	−	
	Object:	+	+	NP_O nuclear clause NP_S
c.	Subject:	+	−	
	Object:	−	+	NP_O (NP_S) nuclear clause
d.	Subject	+	+	
		−	+	nuclear clause NP_S (NP_O)
	Object:	+	−	

In (20a) the unmarked word order is given. The subject is the discourse topic and neither the subject nor the object has particular saliency. Both adjunct NPs are optional, the subject NP precedes the nuclear clause, the object NP follows it.

The object is focused by moving the object NP to the position preceding the nuclear clause, as shown in (20b,c). When the focused object NP is *also* the discourse topic, the subject NP is defocused and follows the nuclear clause the verb, as in (20b). However, when the focused object NP is *not* the discourse topic, it is fronted while the subject NP stays in its place and is optional, as in (20c). A summary of (20d) is that the focused position for a subject NP is the one that follows the nuclear clause.

The observations made in this section can be summarized as follows. Kambera definite NPs are adjuncts and the inversion/focusing of NPs involves varying the order of adjunct NPs which is determined by discourse grammatical properties (topicality) and functional needs (focus). This has effects similar to the voice distinctions induced by passivization in other languages.

4. Voice distinctions by relativisations

The second construction used to background the subject and foreground the object involves relativisations where the object of the embedded verb is the head of the relativisation. In other words, the relativised noun is the object argument that is obligatorily missing in the relative clause. The head noun of such a relativisation is the object of the embedded verb, as *na kabela* in (21) and *da tau* in (22):

(21)   [*Na kabela*   [*na pa- piti-na   na tau   nuna*]]$_j$,   *na$_j$-ruhak*
       ART sword   ART RM- take-3SG   ART person that.one   3SN-be broken
       '[The sword that was taken by that man]$_j$ it$_j$ is broken'

(22)   [*Da tau*   [*da pa- ita -nggu la anda*]]
       ART person   ART RM- see -1SG LOC street
       'The people I saw/that were seen by me on the street'

Among other things, these sentences show that the external syntax of Kambera relative clauses is nominal — they have the structural characteristics of nouns given in (23):

(23)   a. Relative clauses are coreferent to a pronominal clitic in the matrix clause
          (cf. *na-* in (21)).
       b. Relative clauses can be marked for definiteness by the (definite) articles
          (singular *na* in (21), plural *da* in (22)).[11]
       c. The subject of relative clauses is marked with a genitive enclitic (-*na* in (21), -*nggu* in (22)), parallel to the marking of nominal possessors (cf. (15) above).

Note also that Kambera does not mark relative clauses as finite (vs. non-finite). The internal syntax of relative clauses is verbal: the lexical head is a verb, they may contain PPs like *la anda* 'on the street' in (22) and adverbs[12] like *hina* 'recently' in (24):

---

[11]   Kambera does not have indefinite articles. Indefinite NPs are bare, as are indefinite relative clauses.

[12]   There are reasons to distinguish a separate category for adverbs in Kambera (Klamer 1994:116-119). Adverbs only occur in verbal contexts.

(24) [Na njara   [pa-   [hina   kei   -na]]]
       ART horse RM-   recently   buy   -3sG
       'The horse which he recently bought'

Thus, we can represent a Kambera relative construction as in (25). The first nominal constituent (DP₁) *da tau* 'the people' is the head of the relativisation. In apposition to this constituent, there is a second nominal constituent (DP₂). This constituent consists of a functional head (D⁰) (which is the article *da* here) and a clausal complement (CP), which is here the relative clause *pa-ita-nggu la anda* 'that I saw on the street'.

(25)   DP₁
                      /   \
          DP₁   DP₂
          da tau |
          'the people'   D'
                          /   \
                      D⁰   CP
                      | /   \
                      da
                      ART pa-   ita -nggu la   anda
                           RM-   see -1sG   LOC   street
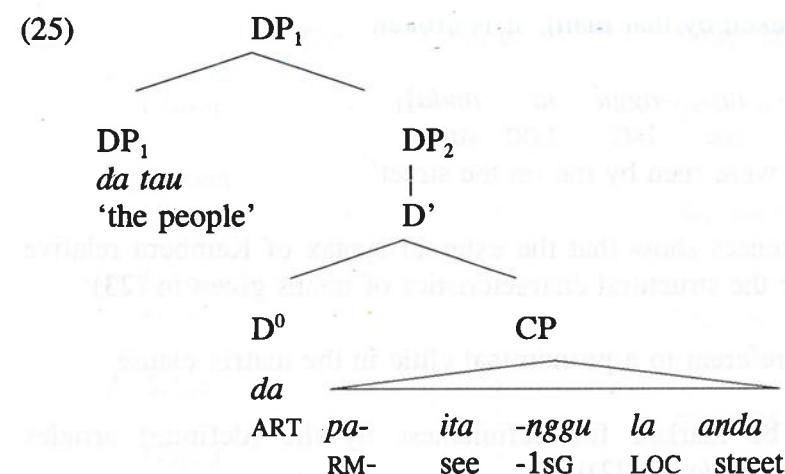
The first and second DPs are considered to be in apposition instead of the second being embedded in the first, because in principle inversion of twe two DPs is possible. Given the limits of this paper I will just assume that (1) there are reasons to consider the relative marker *pa-* as a complementizer: the functional head of the relative clause (CP), i.e. a C (see also Klamer 1994:305-326) and (2) that relativisations in Kambera are nominal structures, even when they lack an article, as in (24), or when they are headless, as in (27):

(26) a. *Na   pa- ngangu -na*
        ART   RM- eat   -3sG
        'What he ate'

   b. *Na   pa-   ita   -nggu*
      ART   RM-   see   -1sG
      'What I saw'

   c. *Na   pa-   palu   -na*
      ART   RM-   hit   -3sG
      'What he hit'

   d. *Na   pa-   bohu -nggu*
      ART   RM-   steal -1sG
      'What I stole'

Some headless relativisations are lexicalised as nouns with a specific interpretation, others are not, as shown in (27):

(27)   *pa-ngangu*   'food (=entity eaten)'
       pa-ita   'vision (=entity/person seen)'
       *pa-palu   *'entity/person hit'
       *pa-bohu   *'entity stolen'

And, like nouns, headless relativisations can modify other nouns, as shown in (28a,b):

(28) a. *meu   rumba*
        cat   grass   'wild cat'
        uhu   wei
        rice   pig   'pig's fodder'

   b. *tau   pa-   palu*
      person   RM-   hit   'a hit person/a person who is hit'
      ndui   pa-   bohu
      money   RM-   steal   'stolen money/money that is stolen'

If relativisations are nominal structures, we also expect to find them as nominal predicates. This is indeed the case. Above, it was mentioned that Kambera has no overt copula or other equivalent of the copular verb *be*. When a clause has a nominal predicate the predicate is simply juxtaposed to the subject. Subjects of nominal

predicates are expressed with an accusative clitic, as illustrated by (17) above.[13] The sentences in (29) below contain nominal predicates that consist of headless relativisations. In (29a) the matrix subject is marked by the clitic -ka '1sA'. This clitic is coreferent with the full NP *nyungga* (which is optional). The embedded subject is marked with a genitive enclitic -na '3sG'. (29b) is similar, except that here the embedded genitive subject is absent. This construction is stripped down further in (29c), a headless relativisation predicating the subject -ka.

(29) a. *Nyungga$_j$* [*pa- palu -na   nyuna*] *-ka$_j$*
        I   RM- hit -3sG   he   -1sA
        'I (am) (the one) who was hit by him'

     b. *Nyungga$_j$* [*pa-   palu*] *-ka$_j$*
        I   RM-   hit   -1sA
        'I was hit (lit. I (am) (the one) hit)'

     c. [*Pa-   palu*]   *-ka*
        RM-   hit   -1sA
        'I was hit (lit. I (am) (the one) hit)'

     d. *Nyungga$_j$* [*pa- palu   la   anda*] *-ka$_j$*
        I   RM- hit   LOC   road   -1sA
        'I was hit on the road (lit. I (am) (the one) hit on the road)'

Apart from the semantic evidence, is there any structural reason to interpret an accusative clitic like -ka in (29c) as the subject of the matrix clause? After all, with its accusative case it is marked like an object. Couldn't it be marking the object of the embedded verb instead? Now consider (29d). The grammaticality of this sentence shows that -ka cannot be considered the object of the embedded transitive verb *palu* 'hit'. The embedded clause contains a PP that occurs *between* the verb and the clitic -ka. As a sentential adjunct, a PP can never occur between a verb and its object clitic

---

[13]   In Kambera, a subject of a non-verbal predicate is marked with an accusative clitic. This is related to the fact that non-verbal predicates are descriptive predicates. With their non-active argument they resemble objects of transitive verbs and subjects of some intransitive verbs (cf. Klamer 1994:137-152).

argument. Therefore, -ka cannot be the clitic used to mark the object of *palu*.[14]

In conclusion, sentences like (29a-d) show that the logical object of the nominalised embedded verb, in a sense, has 'become' the subject of the nominal predicate in the matrix clause. The matrix clause contains a nominal predicate which is made up of an object relativisation. The structure of e.g. (29c) can thus be represented as in (30):

(30) [SENTENCE [NOMINAL CONSTITUENT   [CLAUSE *pa- palu*]]   *-ka*]
                                                     RM- hit   -1sA

What we observe is thus that the logical object of a verb that is embedded in a relative clause can eventually 'become' the grammatical subject of the matrix clause. In this respect, the use of relativisations as nominal predicates as it has been discussed in this section is more similar to a passive than the inversion/focusing discussed in section 3 above, because in the latter process the logical object does not 'become' a grammatical subject, whereas in the former construction it does.

5. Conclusions and discussion

In this paper I have discussed two strategies that Kambera uses to express the functional equivalent of the passive voice in other languages, i.e. putting the PATIENT/logical object in focus and defocus the AGENT/logical subject. The first strategy is a process of focusing/inversion where the syntactic status of the subject and object does not change, the pronominal marking and morphological shape of the verb remains the same but the adjunct NP coreferent to the object is focused while the subject adjunct NP is defocused. The second strategy is to use an object relativisation as a nominal predicate. This structure resembles a canonical passive most closely both in form and function. Both a canonical passive verb and a transitive verb in a Kambera object relativisation have a reduced valency compared to their base verb. They cannot have (or: no longer have) an overt object. In both structures, the logical object of the base verb has, in a sense, 'become' the syntactic subject in the derived structure. The insight that Kambera provides us with is that relative structures are at least functionally related to passive-

---

[14]   Considering -ka as the object marker would also cause other descriptive/ analytical problems. If the construction *pa-palu-ka* were not an object gap relativisation — which it clearly *could* not be if it contained the object clitic -ka — how should the morpheme *pa-* that is attached to the verb *palu* then be analysed and described? What would it be if it were not the same morpheme as the morpheme *pa-* in all the other object relativisations?

like structures. To account for this striking relatedness, I would like to suggest that relative constructions which function as nominal predicates are open to reinterpretation as passive structures in Kambera. This is particularly the case for simple 'short' relative clauses, such as the ones in (29) with no modifying phrases between the relative marker and the remainder of the sentence (an example of such a phrase is *hina* 'recently' in (24)). In what follows I will sketch a possible structural account of the observed reanalysis of the 'short' passives.[15]

Consider a structure for Kambera relatives in standard generative terms as in (31) where the relative morpheme *pa* is a complementizer (C). The relative clause contains a case-marked trace which has as its antecedent an empty operator in the non-argument position Spec CP:
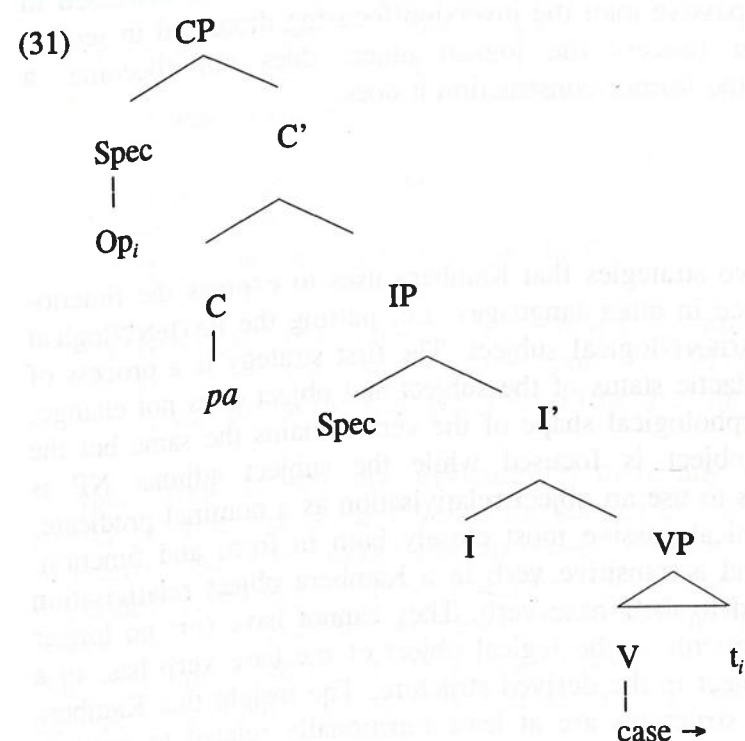
(31)

In 'short' relative constructions there is no overt evidence to interpret the morpheme *pa* as a complementizer (C) instead of an inflectional head (I), i.e. there is no overt evidence to build the structure in (31) with an IP that is empty. Because the evidence for the C-interpretation of *pa* is lacking, the morpheme can be reinterpreted as an inflectional head — a case-absorbing 'passive' affix of the verb (cf. Baker, Johnson and Roberts 1989 for arguments for the base-generation of the passive morpheme under I). As a result of the reanalysis, the whole interpretation of the structure changes. Since the 'passive' prefix now absorbs case, case can no longer be assigned to t_i. This has the result that the antecedent of t_i can no longer be in a non-argument position such as Spec-CP, but must be in a case-marked argument position such as Spec-IP. To make the Spec-IP position available, we need a functional I. To cater for this, *pa* gets a new status in the configuration, it is base-generated under I instead of C, resulting in the structure in (32):

(32)
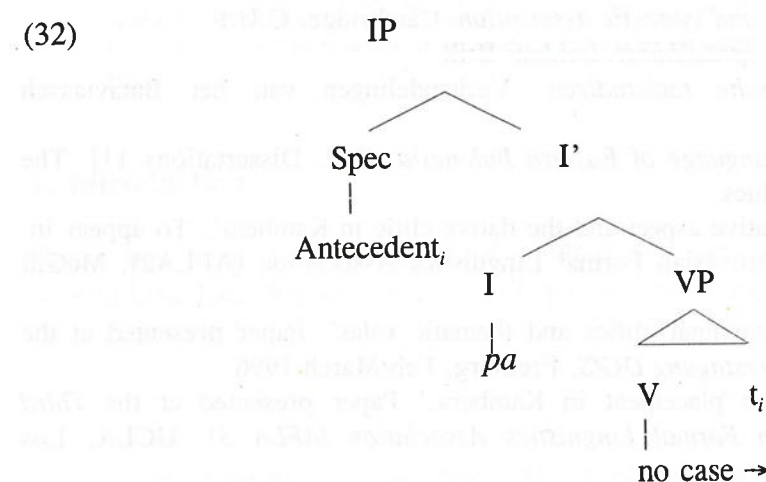
After the reinterpretation of *pa* the CP projection has become superfluous and is removed.

Within the limits of this paper it is impossible to deal with the important details and consequences of an approach like this. But it suggests that a structural account of the insight that Kambera provides us with, namely that relative structures are related to passive-like structures because the relative marker *pa* can be reanalysed as a passive morpheme, is possible. In standard generative accounts of relative clauses and passives there is no hint of similarity between the two structures. Yet, given the Kambera facts described in this paper they seem to have more in common than we expect. Furthermore, *all* languages seem to employ relative structures and non-verbal

---

[15] Thanks are due to Marcel den Dikken for discussion on this point and for suggesting this structural account of the reanalysis.

predicates while, apparently, a number of languages do *not* have specific passive verb forms or passive structures. Therefore it may very well be that the passive as it has been analysed in the literature is a language(-family) specific construction. What Kambera shows us is that an account of passive-like structures with universal claims *must* somehow be related to an account of relativisations in the world's languages.

REFERENCES

BAKER, M., K. JOHNSON & I. ROBERTS. 1989. 'Passive arguments raised'. *Linguistic Inquiry* 20:219-251.

ENGELENHOVEN, A. VAN. 1995. *A description of the Leti language (as spoken in Tutukei)*. Dissertation, Leiden University.

FOLEY, W.A. AND R.D. VAN VALIN. 1985. 'Information packaging in the clause.' In: T. Shopen (ed). *Language typology and syntactic description*. Cambridge: C.U.P.

JONKER, J.C.G. 1915. *Rottineesche spraakkunst*. Leiden: Brill.

JONKER, J.C.G. 1932. *Lettineesche taalstudieën*. Verhandelingen van het Bataviaasch Genootschap LXX.

KLAMER, M. 1994. *Kambera: a language of Eastern Indonesia*. [HIL Dissertations 11]. The Hague: Holland Academic Graphics.

KLAMER, M. (to appear). 'Continuative aspect and the dative clitic in Kambera'. To appear in: Proceedings of the Second Austronesian Formal Linguistics Association (AFLA2), McGill University, Montréal.

KLAMER, M. 1996. 'Kambera pronominal clitics and thematic roles'. Paper presented at the *Workshop on Inflection, 18. Jahrestagung DGfS*, Freiburg, Feb/March 1996.

KLAMER, M. 1996. 'Optimal clitic placement in Kambera.' Paper presented at the *Third conference of the Austronesian Formal Linguistics Association (AFLA 3)*, UCLA, Los Angeles, April 1996.

SHIBATANI, M. 1985. 'Passives and related constructions: a prototype analysis.' *Language* 61:821-848.

Vakgroep Taalkunde
Fac. der Letteren
Vrije Universiteit
P.o. Box 7161
1007 MC Amsterdam
The Netherlands
[klamerm@jet.let.vu.nl]

SUBJECTS IN TAGALOG AND ICELANDIC*

Norvin Richards
MIT

A number of researchers have investigated the syntactic properties of a class of Tagalog nominals marked with *ang* and *si*, variously known as 'topics', 'foci', 'triggers', or 'subjects'. Many Austronesianists (e.g., Payne 1982, De Guzman 1988, Gerdts 1988, Maclachlan and Nakamura 1994) regard these as markers of absolutive case, and see Tagalog as an ergative language. Others (Guilfoyle, Hung, and Travis 1992, Richards 1990, 1993) claim that *ang* and *si* mark nominative case, and describe Tagalog as a nominative-accusative language. In this paper I argue from data on binding, weak crossover, quantifier float, and topic-drop, among other phenomena, that the closest syntactic analogue to the Tagalog topic is the preverbal XP in V2 languages such as Icelandic. Tagalog topicalization is thus argued to be unrelated to any syntactic notion of Case.

1. Introduction

There has been a fair amount of controversy over the right way of thinking about alternations like that shown in (1) (Tagalog), involving a phenomenon I will refer to as "topicalization"[1]:

---

[1]　　Throughout this paper I will be using Schachter's (1976) terminology for describing the Tagalog topicalization system; thus, verbs will be marked AT for Actor-Topic, GT for Goal-Topic, DT for Direction-Topic, etc.; similarly, T stands for Topic, A for Actor (roughly, the logical subject), G for Goal (roughly, the direct object), and so forth. In (1a), for example, the verb is in the Actor-Topic form, because the actor *lalaki* 'man' has been topicalized, while in (1b) the verb bears Goal-Topic morphology that signals the topicalization of the goal *bigas* 'rice'. I will also follow Schachter in using Li for the Tagalog "linker", about which I will have nothing to say here.

(1) a. *Bumili   ang   lalaki   ng   bigas*
        AT-bought   T   man   G   rice
        'The man bought rice'

     b. *Binili   ng   lalaki   ang   bigas*
        GT-bought   A   man   T   rice
        'A man bought the rice'

There are two major camps on this question that I am familiar with, both of which represent the difference between (1a) and (1b) as one involving voice. For those who regard Tagalog as a Nominative-Accusative language (e.g., Guilfoyle, Hung, and Travis 1992, Richards 1990, 1993), (1a) is in the active voice, while (1b) represents something like the passive. This account takes *ang* and *si*[2] to be markers of nominative case. Others (including Payne 1982, De Guzman 1988, Gerdts 1988, and Maclachlan and Nakamura 1994) claim that Tagalog is an Ergative-Absolutive language; on this view, (1b) is the active voice, while (1a) is an antipassive. For those pursuing this view, *ang* and *si* are markers of absolutive case. I will try to argue here that both of these approaches are mistaken and that Tagalog topicalization has nothing to do with case, thus agreeing in spirit with Shibatani (1988). Rather, I will claim that the alternations in (1) are more similar to those in (2) than they are to phenomena involving voice (Icelandic, from Rögnvaldsson and Thráinsson 1990, 3):

(2) a. *Ég   hef   aldrei   hitt   Maríu*
        I   have never   met   Maria

     b. *Maríu   hef   ég   aldrei   hitt*
        Maria   have I   never   met

In Icelandic, as in a number of other languages, some element must move to a position preceding the verb (the so-called "verb-second" or "V2" phenomenon). The question of what drives this movement is a difficult one, and one I do not propose to go into here. It is typically assumed, however, that the driving force behind V2 phenomena has nothing to do with case. Not only do V2 phenomena have no effect on the

---

[2]   These two markers appear to have the same grammatical function; *si* is used with proper names, and *ang* with other nouns.

morphological expression of case in V2 languages, but the movements in question have a very different syntactic character from those we associate with case-driven movement, as we will see. I will try to show here that Tagalog topicalization is syntactically similar to the movement of elements to the pre-verbal position in V2 languages (I will focus particularly on Icelandic, although the claims made here should generalize to any V2 language). According to this theory, *ang* and *si* are not markers of a case at all; rather, they mark the same feature that drives *Maríu* to move to the pre-verbal slot in (2b).[3]

A variety of recent papers (Branigan 1992, Jonas 1992, Harley 1995) have argued for the presence of an A'-specifier above the position in which the subject typically receives case, which is typically occupied by the subject in many languages; I will refer to this functional projection as πP, which is Branigan's (1992) name for it. Branigan (1992), Jonas (1992), and Harley (1995) all take this position to be that occupied by the element preceding the verb in V2 clauses. I will try to argue here that the Tagalog topic also occupies Spec πP. Thus, I will be assuming that Tagalog and Icelandic share the structure in (3):[4]

---

[3]   A natural question arises at this point; if *ang* and *si* are not Case-markers, what is the status of the other morphemes typically taken to mark Case in Tagalog (*ng*, *ni*, *sa*, and so on)? I will have to leave this question an open one for now. It may be that these are true Case-markers, or they may be reanalyzable along the lines developed in this paper as marking roles in another licensing system which bears no relation to Case.

[4]   Verb movement is not shown here; I assume, for example, that the Tagalog verb is base-generated under V° and raises to some higher position (I° or π°).

(3)

π P tree diagram (Icelandic):

```
            πP
          /    \
        NP      π'
               /  \
             π°    IP
                  /  \
                NP    I'
                     /  \
                   I°    VP
                        /  \
                      NP    V'
                           /  \
                         V°    NP
    Maríu hef   ég    e   hitt   e        (Icelandic)
```

```
    e nakilala   ko    e        si Maria   (Tagalog)
```

'I met Mary'
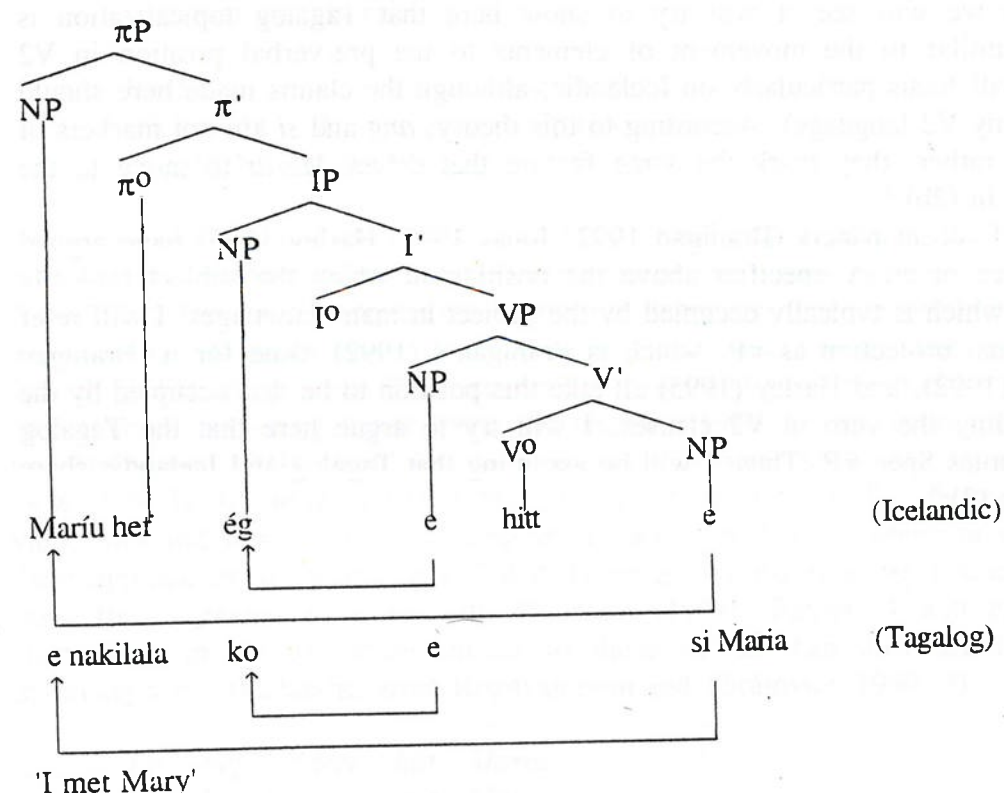
According to this story, the only relevant syntactic difference between Icelandic and Tagalog is that movement to Spec πP in Icelandic is overt, while the Tagalog equivalent involves movement at LF. Of course, there are some obvious morphological differences between the Icelandic and Tagalog cases. For example, Tagalog verbs typically bear morphology indicating which nominal has been topicalized (the infixes *-um-* and *-in-* in (1) are examples of this); I assume that this is an overt instantiation of π°. Icelandic verbs, on the other hand, bear no such morphology. Syntactically, however, Tagalog and Icelandic are quite similar in the relevant respects, as we will see.

2. Tagalog = Icelandic

To begin with, Tagalog topicalization behaves like A'-movement for Binding Theory, which is surprising on a theory in which topicalization is case-driven:

(4) a. *Nagmamahal si Juan sa   kanyang   sarili*
        AT-loves   T Juan Loc his   self
        'Juan loves himself'

     b. *Minamahal   ni Juan ang kanyang   sarili*
        GT-loves   A Juan T   his   self
        'Juan loves himself'

Topicalization of an anaphor (by hypothesis, movement of the anaphor to a position c-commanding its binder) violates neither Condition C nor Condition A. This is not typical of A-movement:

(5)   * *I believe himself to seem to John to be smart*

On the other hand, movement to the preverbal position in a language like Icelandic does behave this way (Steingrímur Karason, p.c.):

(6) a. *Jón elskar   sjálfan sig*
        John loves   himself

     b. *Sjálfan sig elskar   Jón*
        himself   loves   John

Syntactically, then, Tagalog topicalization looks more like movement to the preverbal slot than it does like case-driven movement. Case-driven movement of a reflexive to a position c-commanding its binder is ruled out by the binding theory, as seen in (5). Such a configuration presumably violates Condition C, since the name *John* is A-bound. Icelandic and Tagalog topicalization, on the other hand, both fail to trigger this effect, which suggests that topicalization is not case-driven in either language.

Another parallel between Tagalog and Icelandic topicalization has to do with the behavior of extraction. Adherents of the πP-based explanation of V2 phenomena will have to provide an explanation for the facts in (7) and (8) (Icelandic, (7) adapted from Rögnvaldsson and Thráinsson 1990, 14; (8) from Eirikur Rögnvaldsson and Höskuldur

Thráinsson, p.c.):

(7) a. [$_{CP}$ *Hvern* [$_{\pi P}$ *hefur* [$_{IP}$ *María kysst*]]]?
            whom   has   Mary   kissed

     b.*[$_{CP}$ *Hvern* [$_{\pi P}$ *María hefur* [$_{IP}$ *kysst*]]]?
            whom   Maria   has   kissed

(8) a. [$_{CP}$[$_{\pi P}$ *Steingrímur hefur* [$_{IP}$ *gefið Maríu bókina*]]]
        Steingrimur   has   given   Mary   the-book

     b. [$_{CP}$[$_{\pi P}$ *Bókina hefur* [$_{IP}$ *Steingrímur gefið Maríu*]]]
        the-book   has   Steingrimur   given   Mary

     c. [$_{CP}$ *Hverjum* [$_{\pi P}$ *hefur* [$_{IP}$ *Steingrímur gefið bókina*]]]?
        whom   has   Steingrimur   given   the-book

     d.*[$_{CP}$ *Hverjum* [$_{\pi P}$ *bókina hefur* [$_{IP}$ *Steingrímur gefið*]]]?
         whom   the-book   has   Steingrimur   given

In main clauses, extraction apparently cannot take place if Spec πP is occupied by another nominal. One way of describing this would be to say that extraction must take place via Spec πP; that is, to say that only elements in Spec πP can be extracted. The famous Tagalog ban on extraction of non-topics, of course, looks very similar:

(9) a. [$_{CP}$ *Sino ang* [$_{\pi P}$ *hinalikan* [$_{IP}$ *ni Maria*]]]?
       who   DT-kissed   A   Mary
       'Who did Mary kiss?'

     b.*[$_{CP}$ *Sino ang* [$_{\pi P}$ e *humalik* [$_{IP}$ *si Maria*]]]?

       who   AT-kissed   T   Mary
       'Who did Mary kiss?'

On this story, (9a) is parallel to (7a), and (9b) to (7b); if the subject is made into the topic, extraction cannot take place, while if the extracted object is made the topic the structure is well-formed.

Icelandic, like many V2 languages, has a process of "topic-drop" (Sigurðsson 1993, 254-255):

(10) a. (*Ég*) *þekki það ekki*
        (I)   recognize   that   not

     b. (*það*) *þekki ég ekki*
        that   recognize   I   not
       'I don't recognize that'

     c. *Núna þekki* *(*ég*) *(*það*) *ekki*
        now   recognize   I   that   not
       'Now I don't recognize that'

That is, Icelandic allows certain elements to drop just in case they are topicalized. Tagalog seems to behave similarly. For example, (11a), but not (11b), is an appropriate response to 'Why is Juan sick?' (Marlon Abayan, p.c.):

(11) a. *Baka kumain* (*siya*) *ng tambakol*
        maybe   AT-ate   T-he   G   mackerel
       'Maybe he ate mackerel'

     b. *Baka kinain* *(*niya*) *ang tambakol*
        maybe   GT-ate   A-he   T   mackerel
       'Maybe he ate the mackerel'

Rizzi (1992) notes a peculiar fact about topic-drop in German which he attributes to Cardinaletti (1990). Apparently dropped topicalized subjects can be of any person, while non-subject topics must be 3rd person (from Rizzi 1992):

(12) a. (*Ich*) *habe es gestern gekauft*
        (I)   have   it   yesterday   bought
       'I bought it yesterday'

     b. (*Das*) *habe ich gestern gekauft*
        (that)   have   I   yesterday   bought
       'I bought that yesterday'

c.*(*Dich*) *habe ich   nicht   gesehen*
  (you) have I   not   seen
'I haven't seen you'

At least some Tagalog speakers apparently have judgments which are strongly reminiscent of the German facts. The conversation in (13a), involving drop of a 3rd person goal-topic, is more felicitous than that in (13b), which contains a dropped 2nd person goal-topic (Marlon Abayan, p.c.):

(13) a. A: *Ano ang ginawa ng   kapatid mo kagabi?*
        what T   TT-did A   sibling your last-night
        'What did your sister do last night?'

    B: *Pumunta siya   sa   Makati*
        AT-went T-she L   Makati
        'She went to Makati'

    A: *A,   oo, nakita ko yata   doon*
        oh   yes TT-saw A-I it-seems there
        'Oh yes, I think I saw (her) there'

   b. A: *Ano ang ginawa ninyo   kagabi?*
        what T   TT-did A-you-pl last-night
        'What did you do last night?'

    B: *Pumunta kami   sa   Makati*
        AT-went T-we-excl L   Makati
        'We went to Makati'

  ?? A: *A,   oo, nakita ko yata   doon*
        oh   yes TT-saw A-I it-seems there
        'Oh yes, I think I saw (you) there'

On the other hand, (14) is perfectly acceptable (say, in place of B's response in (13a)):

(14)   *Pumunta (kami)   sa   Makati*
      AT-went T-we-excl L   Makati
      '(We) went to Makati'

Thus, in both Tagalog and German, dropped actor-topics, but not dropped goal-topics, can be 1st or 2nd person. On a theory in which German and Tagalog topics are syntactically similar, this is unsurprising; Tagalog thus appears to be behaving, again, like a V2 language.

It has often been noted that Tagalog topics must be specific, or definite, or "old information", or something of that kind. In fact, this has been one of the arguments offered against equating the Tagalog topic with subjects in other languages, on the grounds that such semantic restrictions seldom apply to subjects (Bowen 1965, Schachter and Otanes 1972). Interestingly, the Icelandic topic appears to also be associated with a definite reading. Topicalization of indefinites is considerably more awkward than topicalization of definites (Höskuldur Thráinsson, p.c.):

(15) a. *Jón keypti bókina*
      Jon bought the-book

   b. *Jón keypti bók*
      Jon bought a-book

   c. *Bókina   keypti Jón*
      the-book bought Jon

   d.??*Bók   keypti Jón*
      a-book bought Jon

Thus, in Icelandic, as in Tagalog, Spec πP appears to be associated with definiteness in some way.

The claim that *ang* marks absolutive case seems to encounter difficulties with phenomena involving weak crossover. Tagalog topicalization has effects on weak crossover that are unexpected if *ang* is a marker of absolutive case:

(16) a.*Nagmamahal ang kanyang$_i$ ama   ng bawat anak$_i$*
     AT-loves   T his   father G every child
     'His father loves every child'

b.?*Minamahal   ng   kanyang$_i$ ama   ang   bawat   anak$_i$
 GT-loves   A   his   father T   every   child
 'His father loves every child'

The quantifier in (16a) is unable to bind the pronoun as a variable. This effect is standardly attributed to the ban on weak crossover. QR of the quantifier to a position c-commanding the pronoun will create a weak crossover configuration, since no A-positions associated with the quantifier will c-command the pronoun. Interestingly, topicalization of the quantifier remedies this weak crossover effect, as we can see in (16b). This is not what we expect if *anak* in (16b) is absolutive and *ama* is ergative (Basque and Nisgha, from Bobaljik 1993, 60):

(17) a.*Nor   maite   du   bere   amak?
      who-ABS   love   AUX.3sA/3sE his   mother-ERG
      'Who$_i$ does his$_i$ mother love?'

    b.*Næ-gat   ɬ   ti-sipən-s   nɔx^w-t
      who-one   ND   FOC-love-DM   mother-3s
      'Who$_i$ does his$_i$ mother love?'

Here we see that making a quantificational element absolutive does not typically allow it to bind a variable in an ergative nominal. Icelandic topicalization, however, does remedy weak crossover, just like Tagalog topicalization (Höskuldur Thráinsson and Eirikur Rögnvaldsson, p.c.):

(18) a.*Foreldrar hans kenna sérhverjum   strák   að   keyra
       parents   his   teach   every   boy   to   drive
       'His$_i$ parents teach every boy$_i$ how to drive'

    b.?Sérhverjum strák kenna foreldrar hans   að   keyra
       every   boy   teach   parents   his   to   drive
       'Every boy$_i$, his$_i$ parents teach how to drive'

Once again, *ang*-marked nominals behave like Icelandic topics, and not like bearers of a particular case.

We have now seen arguments from a number of domains suggesting that Tagalog topicalization be treated as syntactically similar to topicalization in Icelandic. Several arguments have been advanced against understanding nominals marked with *ang* as

topics; let us move on to consider these. Kroeger (1991) offers one such argument. Following Bresnan and Mchombo (1987), he claims that topics are "what is under discussion, whether previously mentioned or assumed in discourse" (Bresnan and Mchombo 1987, 746). One prediction of this definition of topics, Kroeger says, is that Tagalog sentences like (19) should not be possible answers to a question like "What did you buy?". As he notes, this is false (Kroeger 1991, 79):

(19) Binili   ko   itong   damit
      GT-bought   A-I   T-this   dress
      'I bought this dress'

Here the topic *itong damit* 'this dress' must crucially be new information; it cannot represent information already presupposed in the discourse, or the answer would be uninformative. This is incompatible, however, with the notion of topic as Kroeger has defined it; topics must be old information, previously mentioned in the discourse. Kroeger concludes that *ang/si*-marked nominals cannot be topics, in his sense.

Here we run into terminological difficulties (probably mine, not Kroeger's). There may well be good reasons to reserve the term 'topic' for a class of nominals fitting Kroeger's definition. In any event, Kroeger's observations do not damage the case being made here for equating the Tagalog and Icelandic phenomena. Icelandic sentences which, on this account, are structurally equivalent to (19) are also well-formed responses to questions like "what did you buy?"

(20) βessa   bók   keypti   ég
      this   book bought   I
      'I bought this book'

Thus, Kroeger's arguments serve to show that the class of nominals referred to as "topics" in the literature probably do not form a natural class; at the very least, a distinction should be drawn between Icelandic and Tagalog topics, on the one hand, and topics in the sense Kroeger defines, on the other.[5] A full-scale investigation of the semantics of topicalization is beyond the scope of this paper, unfortunately, but it seems clear that with regard to the diagnostic discussed by Kroeger, Tagalog and Icelandic continue to behave similarly.

---

[5]   We might also want to distinguish the cases of 'topicalization' studied here from 'topicalization' in non-V2 languages, as in *Fish, I'm very fond of*.

Another argument against understanding *ang*-marked nominals as topics has been based on quantifier float. For some Tagalog speakers, as Schachter (1976) observes, the quantifier *lahat* 'all' may float to a position immediately following the verb. Such floated quantifiers are always construed as modifying the topic (Schachter 1976, 501):

(21) a. *Sumusulat   lahat   ang   mga bata ng   mga liham*
        AT-writing   all   T   pl. child G   pl. letter
        'All the children are writing letters'

    b. *Sinusulat   lahat   ng   mga bata ang   mga liham*
        GT-writing   all   A   pl. child T   pl. letter
        'The children are writing all the letters'

If quantifier float is a diagnostic of argumenthood, this is unexpected on the theory of Tagalog topicalization developed here.

Suppose we consider the properties of quantifier float in some other languages. Comparison of Tagalog with Icelandic is impossible in this case, as floated quantifiers in Icelandic are not subject-oriented (adapted from Bobaljik 1995, 249):

(22) a. *Strákarnir máluðu húsið* **allir** *rautt*
        boys-the   painted house-the all   red
        'All the boys painted the house red'

   b. *það   borðuðu   margir strákar bjúgun   ekki* **öll**
        there ate   many   boys   sausages-the not all
        'Many boys didn't eat all of the sausages'

Icelandic floated quantifiers may be associated either with the subject, as in (22a), or with the object, as in (22b). Note that this might make us uneasy about using floated quantifiers as a diagnostic for structure; floated quantifiers clearly have different properties in different languages, and we have no theory that tells us what those properties might be, or what the differences show.

Ignoring this problem for a moment, let us consider quantifier-float in English. The English equivalent of (22b) is not well-formed ("Many boys didn't eat the sausages all" is not a good English sentence), so perhaps English floated quantifiers are subject-oriented in the relevant sense, and a meaningful comparison with Tagalog can be made. Now we need to go on to investigate the notion of "subject-orientation" at work here. According to Branigan's (1992) theory of clause structure, in ordinary

transitive sentences the subject is base-generated in an internal subject position (Spec VP), moves to a position associated with nominative Case (Spec AgrSP), and then moves to Spec πP:

(23)
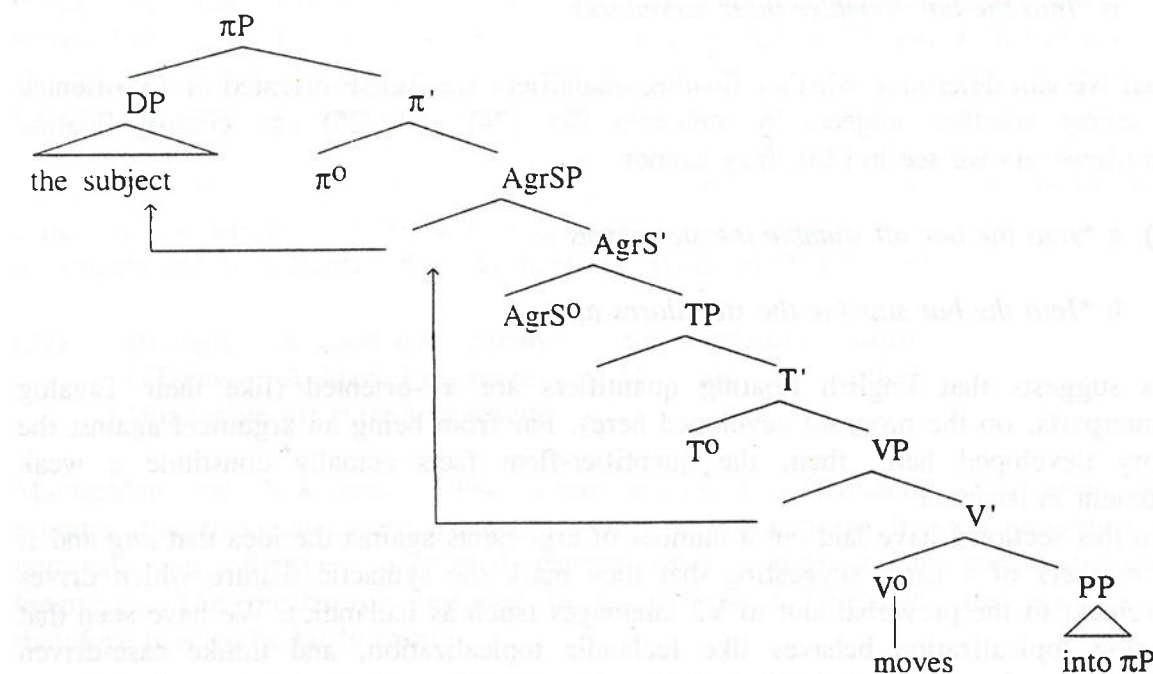
Thus, 'subject-orientation' might consist of Spec AgrSP-orientation, or of Spec πP-orientation, on this theory, since both of these positions are typically occupied by the subject. To determine which of these approaches to subject-orientation is the right one, we should consider some sentences in which elements other than the subject move to Spec πP. According to Branigan (1992), the sentences in (24) involve movements of this kind:

(24) a. *Into the bar stumbled three drunken stevedores*

    b. *"Tagalog is really Icelandic!" yelled the deranged linguist*

In (24a, b), the preposed elements *into the bar* and *"Tagalog is really Icelandic!"* are supposed to have moved into Spec πP. Note that the subject is still associated in some way with AgrSP, since it controls subject agreement:

(25) a. *Into the bar stumble three stevedores*

     b.*Into the bar stumbles three stevedores*

Thus, we can determine whether floating quantifiers are AgrSP-oriented or πP-oriented by seeing whether subjects in sentences like (24) and (25) can control floating quantifiers. As we see in (26), they cannot:

(26) a.*Into the bar **all** stumble the stevedores*

     b.*Into the bar stumble the stevedores **all***

This suggests that English floating quantifiers are πP-oriented (like their Tagalog counterparts, on the proposal developed here). Far from being an argument against the theory developed here, then, the quantifier-float facts actually constitute a weak argument in its favor.

In this section I have laid out a number of arguments against the idea that *ang* and *si* are markers of a case, suggesting that they mark the syntactic feature which drives movement to the preverbal slot in V2 languages (such as Icelandic). We have seen that Tagalog topicalization behaves like Icelandic topicalization, and unlike case-driven movement, with regard to binding theory, wh-extraction, "topic-drop", a definiteness effect, and weak crossover. I then went on to consider two arguments against understanding *ang* and *si* as topic markers. We saw that Tagalog and Icelandic behave alike with regard to Kroeger's semantic diagnostic for topichood. The properties of quantifier-float were then investigated and argued not to militate convincingly against the analysis proposed here.

If we take these parallels between Tagalog and Icelandic topicalization seriously, we are inclined toward an analysis that posits the same structure for both; that is, we are inclined to assume that topicalization in Tagalog, as in Icelandic, involves movement to an A' specifier c-commanding the position in which the subject gets case. In the next section we will see how this assumption simplifies our understanding of the structure of Tagalog.

3. Tagalog case

The preceding section reviewed a number of phenomena which are unexpected under a theory in which Tagalog topicalization involves movement to the position in which the subject gets case in English. By taking advantage of recent refinements of the Internal Subject Hypothesis, according to which the subject is associated with several external subject positions with different properties, we can account for these phenomena straightforwardly. Furthermore, by not assigning the topic either nominative or absolutive case, we avoid problems involving case-assignment to the other nominals in the sentence; this can now take place in the usual way.

According to the accusative view of Tagalog, the topic is in the nominative case, and direct objects which are not topics bear accusative case. Such a view must invent some ad hoc mechanism for assigning case to non-topic actors. Actors are clearly arguments and not adjuncts; they can bind reflexives, as shown in (27):

(27) *Ibinigay   ni Juan ang premyo   sa   kanyang   sarili*
     GT-gave   A Juan T   prize   D   his   self
     'Juan gave the prize to himself'

Maclachlan and Nakamura (1994) argue against the nominative-accusative view, pointing out that the case-marking on actors is identical to that on possessors of nominals, an otherwise unattested pattern (they claim) in nominative-accusative languages. The mechanism invented to assign case to the non-topic actor would therefore have to be fairly exotic:

(28) a. *lapis   ko*
     pencil   A-I
     'my pencil'

     b. *Binili   ko   ang lapis*
     GT-bought   A-I   T   pencil
     'I bought the pencil'

On the account developed here, on the other hand, no difficulties arise, since topicalization does not involve movement to a case position. The subject can receive case in the usual way. If we decide to view Tagalog as an accusative language, then Tagalog actors and possessors both bear nominative case. This <u>is</u> an attested pattern in nominative-accusative languages; Tagalog possessors look very much like their

Hungarian counterparts ((30) is Hungarian, from Abney 1987, 44 and 46):

(29) a. *lapis   ko*
        pencil   A-I

     b. *aking   lapis*
        D-I   pencil
        'my pencil'

(30) a. *a   te   vendeg -e   -d*
        the   you-NOM guest   POSSESSED 2s
        'your guest'

     b. *Peter -nek   a   kalapja*
        Peter   DAT   the   hat
        'Peter's hat'

Hungarian possessors can be assigned either nominative or dative case, as can be seen in (30). On the account developed here, *ni* could be seen as a marker of nominative case, and Tagalog could be taken to have both dative and nominative possessors as well.

According to the ergative view of Tagalog structure, sentences like (31) must be antipassives, since the subject is in the absolutive case; that is, the object must not receive case from the verb:

(31)   *Kumain   ang   lalaki   ng   tambakol*
        AT-ate   T   man   G   mackerel

In the story developed here, on the other hand, the marker *ang* has nothing to do with the case assigned to *lalaki*; (31) may very well be an active sentence. Kroeger 1993 has argued convincingly against the view that (31) is an antipassive, since the direct object behaves syntactically like an argument of the verb. Kroeger outlines several tests for argumenthood, only one of which I will describe here. PRO in adjunct clauses cannot be controlled by non-arguments in Tagalog. This can be seen in (32) (adapted from Kroeger 1993, 43); here, the marker *sa* on *hari* 'king' effectively marks it as an adjunct:

(32)   *Bumisita   si Juan_i sa   hari_j   [nang   nagiisa PRO_{i/*j}]*
        AT-visited T Juan D   king   Adv   AT-one
        'Juan visited the king alone'

(32) can only be understood as meaning that Juan was alone, not that the king was alone. (33a) and (33b), by contrast, are both ambiguous (adapted from Kroeger 1993, 47):

(33) a. *Hinuli ng polis_i   ang magnanakaw_j   [nang pumapasok PRO_{i/j} sa bangko]*
        GT-caught A police   T   thief   Adv AT-enter   D bank
        '(The) police caught the thief entering the bank'

     b. *Nanghuli ng magnanakaw_i ang polis_j   [nang   pumapasok PRO_{i/j} sa bangko]*
        AT-caught G thief   T   police   Adv   AT-enter   D bank
        'The police caught a thief entering the bank'

Note that in (33b), PRO can be controlled by the direct object even when the verb is in the actor-topic form; that is, even in the putative antipassive. This seems to suggest that non-topic direct objects are in fact arguments of the verb, contrary to the ergative view. Accounts of Tagalog that describe it as ergative must therefore invent a new mechanism to license the direct object in "antipassives", just as the nominative-accusative account must invent a new way for non-topic subjects to acquire case. The theory developed here, on the other hand, does not need to complicate existing accounts of case, since topicalization is not taken to be case-driven movement.

4. Conclusion

At the beginning of this paper I noted that a controversy exists on the question of whether Tagalog is a nominative-accusative or ergative-absolutive language. This controversy has been predicated on the assumption that topicalization is case-driven movement; that is, that markers like *ang* and *si* mark nominals as bearing either nominative or absolutive case. I have tried to show here that this view is incorrect, and that movement to the topic position in Tagalog is more like movement to the pre-verbal slot in languages like Icelandic than it is like passive in English. If this conclusion is right, the question of Tagalog's case system will have to be completely re-thought. There are, of course, interesting differences between Icelandic and Tagalog topicalization which may shed light on this question (for example, the unmarked

Tagalog topic appears to be the direct object, which is not the case in Icelandic), but these are matters beyond the scope of this paper.

According to what I have said here, the major difference between Tagalog and Icelandic is simply that Tagalog topicalization is associated with rich verbal and nominal morphology, while Icelandic topicalization is signalled by changes in word order. In the domain of case, pairs of this kind are well-documented; we typically think of languages like Latin as expressing case by means of rich morphology, while languages like Chinese use changes in word order. If the theory sketched here is right, Tagalog and Icelandic are another such pair with respect to a licensing system which has nothing to do with case.

REFERENCES

ABNEY, S. 1987. The English noun phrase in its sentential aspect. Ph.D. dissertation, MIT.

BOBALJIK, J. 1993. Ergativity and ergative unergatives. In: Colin Phillips (ed). *Papers on Case and Agreement II*, [MIT working papers in linguistics], 45-88.

BOBALJIK, J. 1995. *The Syntax of Verbal Inflection*. Ph.D. dissertation, MIT.

BOWEN, J. D. 1965. *Beginning Tagalog*. Berkeley and Los Angeles: University of California Press.

BRANIGAN, P. 1992. *Subjects and complementizers*. Ph.D. dissertation, MIT.

BRESNAN, J. AND S. MCHOMBO. 1987. 'Topic, pronoun and agreement in Chichewa.' *Language* 63.4:741-782.

CARDINALETTI, A. AND I. ROBERTS. 1991. 'Clause structure and X-second.' ms., University of Geneva.

DE GUZMAN, V. 1988. 'Ergative analysis for Philippine languages: an analysis.' In: McGinn (ed), 323-345.

GERDTS, D. 1988. 'Antipassives and causatives in Ilokano: evidence for an ergative analysis.' In: McGinn (ed), 323-345.

GUILFOYLE, E., H. HUNG, AND L. TRAVIS. 1992. 'Spec of IP and Spec of VP: two subjects in Austronesian languages.' *Natural Language and Linguistic Theory* 10. 375-414.

HARLEY, H. 1995. 'Abstracting away from abstract Case.' To appear in Proceedings of NELS 26.

JONAS, D. 1992. 'Checking theory and nominative Case in Icelandic.' In: Höskuldur Thráinsson and Susumu Kuno (eds), *Harvard working papers in Linguistics* 1, 175-195.

KROEGER, P. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI.

MACLACHLAN, A. AND M. NAKAMURA. 1994. 'Case checking and specificity in Tagalog.' ms., McGill University.

MCGINN, R. (ed). 1988. *Studies in Austronesian Linguistics*. Athens, OH: Ohio University Center for International Studies.

PAYNE, T. 1982. 'Role and reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog.' *Studies in Language* 6. 75-106.

RICHARDS, N. 1990. 'Tagalog subjects.' Unpublished ms., Cornell University, Ithaca.

RICHARDS, N. 1993. 'Tagalog and the typology of scrambling.' Unpublished honors thesis, Cornell University, Ithaca.

RÖGNVALDSSON, E., AND H. THRÁINSSON. 1990. 'On Icelandic word order once more.' In J. Maling and A. Zaenen (eds), *Modern Icelandic syntax* [Syntax and Semantics 24], 3-38. San Diego: Academic Press, Inc.

RIZZI, L. 1992. 'Early Null Subjects and Root Null Subjects.' ms., Université de Genève.

SCHACHTER, P. 1976. 'The subject in Philippine languages: topic, actor, actor-topic, or none of the above?' In Ch. Li (ed) *Subject and topic*, 491-518. New York: Academic Press.

SCHACHTER, P., AND F. OTANES. 1972. *A Tagalog reference grammar*. Berkeley and Los Angeles: University of California Press.

SHIBATANI, M. 1988. 'Voice in Philippine languages.' In M. Shibatani (ed) *Passive and voice*, 85-142. Amsterdam and Philadelphia: John Benjamins.

SIGURASSON, H. 1993. 'Argument-drop in Old Icelandic.' *Lingua* 93:247-280.

VOSKUIL, J. 1993. 'Verbal Affixation in Tagalog (and Malay).' In *McGill Working Papers in Linguistics* 8:75-106. Montréal: Department of Linguistics, McGill University.

ZAENEN, A., J. MALING, AND H. THRÁINSSON. 1990. 'Case and Grammatical Functions: The Icelandic Passive.' In J. Maling and A. Zaenen (eds), *Modern Icelandic syntax* [Syntax and Semantics 24], 95-106. San Diego: Academic Press, Inc.

Building 20D
Massachussetts Institute of Technology
Cambridge, MA 02139
United States
[norvin@mit.edu]

ASPECTS OF VERBAL MORPHOLOGY IN TAGALOG

Jan E. Voskuil

Leiden University/HIL

This paper presents an argument concerning the status of certain verbal affixes in Tagalog. Rather than promoting specific arguments to subject (or topic, focus, etc.), they change the verb's argument structure. The resultant verb is passive by default. Evidence will be drawn from comparison with affixes in other languages.

1. Overview

In this paper I present an argument concerning the interaction of voice and verb-deriving morphology in Tagalog. The standard analysis of, for instance, the affix *-an* is that it attaches to transitive verbs; the function of this affix is to promote a designated argument of the verb (the "location" argument) to subject. In this analysis, Tagalog has many kinds of passive voices: a "plain" passive with *-in*, which passivizes the patient, a conveyance passive with *i-*, which passivizes the theme, a "locational" passive with *-an*, which passivizes the location, etc. etc. Thus, a simple verb like 'plant', as in *Celia planted flowers in the garden*, has two passives. One (with *i-*) promotes *flowers* to subject, the other (with *-an*) promotes *the garden* to subject. This analysis can be called the MULTIPLE PASSIVE ANALYSIS.

I will argue for a different interpretation. The affixes form particular types of verbs — that, and nothing else, is their function. The verbs thus formed are passive by default. Consequently, Tagalog has only *one* passive, although it has many verb types. The suffix *-an* can be compared to *be-* in *besprinkle*; it specifies a semantic relation between the verb and its deep-object. The voice system then promotes the deep-object to subject. This analysis can be called the TWO-STAGE ANALYSIS.

The argument is based on a comparison with affixes in other languages, notably Dutch and Indonesian. These affixes appear to have the same functions as *-an*, except that they are neutral as to the voice of the verb. Similar syncretisms exist in the case of *i-*. Therefore, it must be possible to charactarize the functions of these affixes independently of the voice system.

The paper is structured as follows. I section 2 I will present the basic data, that is,

the morphological syncretisms aluded to above. In section 3 I will briefly discuss some basic notions involved in these syncretisms, and make explicit some assumptions about how they can arise. In section 4, I demonstrate that the Dutch and Indonesian affixes are independent of voice. Also, I will briefly sketch the Tagalog voice system. Section 5 presents the main argument, based on the above preliminaries.

2. Remarkable correspondences

From a diachronic perspective, nothing changes as fast as the function range of an affix. It is surprising that three languages — Tagalog, Indonesian, and Dutch — share a very specific affix with complex behavior. In Dutch, it comes out as *be-*, in Tagalog as *-an*, in Indonesian as *-i*. To get a feel for the affix *-an/-i/be-*, consider words like *belated*, *bespectacled*, *beloved*, *to besmearch*, and *to belabour*. Some of them only occur as adjective, not as verb. But the affix is, historically, the same as the Dutch variant of *-an/-i/be-*. The three affixes derive verbs that belong to a broad and varied spectrum of verb classes. Most often, it is possible to predict from the nature of the stem what the derived verb will be. Adjectives like 'pregnant' derive verbs like 'to make pregnant'. Nouns like 'weapon' derive verbs like 'to provide with weapons'. Verbs derived from 'to plant' and 'to spray' are roughly synonymous with their morphologically simplex base, but the affixation goes along with the appearance of a preposition meaning 'with'. Table 1 is a schematic overview of similarities — which is not to say that there are no differences. A similarly remarkable syncretism is that between Tagalog *i-* and Indonesian *-kan*. As shown in table 2, they have roughly the same range of functions.

			Dutch *be-*	**Indon.** *-i*	**Tagalog** *-an*
(I)	A A + aff	pregnant to make pregnant	*wapen* *bewapenen*	*senjata* *senjatai*	*sandata* *sandatahan*
(II)	A A + aff	weapon to give weapons	*zwanger* *bezwangeren*	*bunting* *buntingi*	*buntis* *buntisan*
(III)	V V + aff	to plant *x* in *y* to plant *y* with *x*	*planten* *beplanten*	*tanam* *tanami*	*tanim* *tamnan*
(IV)	V V + aff	to send *x* to *y* to send *y x*		*kirim* *kirimi*	*padala* *padalhan*

(V)	V V + aff	to look at to scrutinize	*kijken* *bekijken*	*lihat* *lihati*	*masid* *masdan*

Table 1

			Indonesian *-kan*	**Tagalog** *i-*
(I)	A A + aff	free to set free	*bebas* *bebaskan*	*bulós* *ibulós*
(II)	N N + aff	pocket to pocket	*kantong* *kantongkan*	*bulsá* *ibulsá*
(III)	V V + aff	to plant to plant (smt) in (smt)	*tanam* *tanamkan*	*taním* *itaním*
	V V + aff	to go inside (of) to bring inside (of)	*masuk* *masukkan*	*pasók* *ipasók*
(IV)	V V + aff	to buy (smt) to buy (smo) (smt)	*beli* *belikan*	*bilí* *ibilí*
(V)	V V + aff	to ask for to ask earnestly for	*mohon* *mohonkan*	*luhog (-um-)* *iluhog*
(VI)	V V + aff	to buy, go shopping to spend (smt)	*belanja* *belanjakan*	*bilí* *ibilí*
(VII)	V V + aff	to sit to seat (smo)	*duduk* *dudukkan*	*upô (-um-)* *iupô*

Table 2

In short, the same morphological patterns consistently recur cross-linguistically — patterns labyrinthine enough to rule out sheer coincidence. Which properties of universal grammar are responsible for this recurrence? There is little point in framing

the question in a Chomskian manner, and ask how speakers of these languages "know" or "cognize" — whatever that means — the morphological regularities at hand. Speakers of Tagalog and Indonesian know that "be-pregnant" means 'to make pregnant' and can be used literally, while speakers of Dutch know that the same verb in their language means the same but can only be used non-literally in certain types of archaic collocations, as in *De lucht was bezwangerd van een zoete geur* "The air was filled (lit. made-pregnant) with a sweet fragrance." In other words, they know what the conventions are that govern the use of these verbs. The question, then, is how it can be that conventions that arise in these different speech communities resemble each other so closely, without being identical.

Baker and Hacker (1984: 374f) compare the system of a language to the network of trunk roads in Brittain. "One is predisposed to view the road system ... as the more or less haphazard product of a host of piecemeal improvements (and natural disasters!) spread over two millenia." This picture applies especially aptly to morphological systems of the type at hand. Philippinos, Indonesians, and Dutchmen have, through the ages, haphazardly constructed a morphological trunk-road system. However, the landscape appears to be virtually the same in all three cases: an abyss here, an important well there, a place particularly suited for shelter a few miles in that direction. It is not surprizing to find that the road-networks that actually evolved closely resemble each other, without being identical. The landscape, then, is a biologically determined part of the human mind. I will sharpen these ideas in the next section.

It will not be possible, however, to go very deeply into these matters; the purpose of this paper is merely to argue for a particular interpretation of the status of the affixes *-an* and *i-* in Tagalog. It is essential for the argument to be sound (as opposed to valid) that the syncretisms be investigated and analysed in detail, and that the conceptual issues involved in explaining these syncretisms are made explicit and coherent. I hope to come back to these issues at another occasion. In the present article I will content myself with the basic data presented above — putting aside many important details — and some sketchy remarks about the nature of the syncretisms in the next section, proceeding as quickly as possible to the main argument. The background assumptions into which the following discussion is embedded are — very roughly speaking — those of Reichling (1935) and Uhlenbeck (1978). Affixes are interpreted as vectors that locate the affixed word within a taxonomical space, usually relative to a stem. In this framework, it does not make sense to say that affixes have meanings, or that they can be polysemous.

3. How many functions can go in an affix?

The function of an affix in a word is exhaustively determined by the properties of the stem and those of the word. Each occurrence of an affix in the lexicon determines a unique function; an affix has at least as many functions as there are words containing that affix. Although it is tempting to think that the five functions of, say, Tagalog *-an* in table 1 are, in some sense, "really" only one abstract function, this is not true at any fundamental theoretical level. What matters is that the five functions, as defined by the five *-an*-words plus their stems, are to some extent *similar*. One can still entertain an informal, pretheoretical concept of identity. For instance, one can sensibly say that Tagalog *-an* in the five functions is "one" affix, but that the Tagalog affix *-an* that, say, forms *lungtián* 'greenish' from *lungtî* 'green' constitutes a "different" affix. To put it more precisely, one would have to say that the former five functions of the affix are relatively more similar to each other than any of them is similar to its function in the Tagalog word for 'greenish'. Think of the affixes in the table as constituting a species. Darwinism has shown that the concept of species is an arbitrary convention, handy for biologists, but devoid of theoretical status (Dennet 1995). It does not matter whether we want to view the lesser black-backed gull and the herring gull as one species, or, for that matter, coyotes, wolves and dogs. What matters is what their location is relative to each other in the design space of organisms.

The specific function of an affix in a word is atomic. We can define taxonomically higher groupings as structured conglomerates of atomic functions. These conglomerates, in as far as they are homogeneous relative to some measure of similarity, correspond to the traditional, intuitive idea of the function of an affix.

We can define similarity between affix functions as a function of similarity between the derived words. To see this, consider the word *besprinkle*. The stem *-sprinkle* indicates that the word is similar to the verb *sprinkle*, the prefix *be-* indicates similarity with *besmear*, *bestride*, *beset*, *betray*, *bewitch*, but also with *below*, *before*, and *beyond*. Inspection of the data shows that *besprinkle* is more akin to *besmear* than to *beyond*, and that *sprinkle* is more akin to *smear* than to *yon*. This must be reflected in the classification of the functions that *be-* has in all these examples.

Explicit reference to the stem of a derived word is not even necessary if it is true — as I will assume — that stem properties can constitute a secondary taxonomic criterion in the lexicon of derived verbs. Thus, similarity can now be characterized as follows: the more similar two words derived by an affix, the more similar the functions of the affix in those two instances. The function range of an affix constitutes a contiguous area to the extent that the range of derived words is a contiguous area in the design

space of words — the virtual lexicon.

All this does not yet suffice to explain the recurrence of syncretisms. Why would affixes strive to accumulate a contiguous range of functions? This follows from the purposes that language in general, and morphology in particular, are designed to serve: to facilitate communication. Uhlenbeck (1978) defines morphology as the means by which the lexicon can be *systematically* expanded. Consider what would happen if an affix would expand the lexicon in unsystematic ways. An individual language user can do this freely, idiosyncratically forming new words *ad lib* with his own personal affix. But other language users would be unable to tell the meaning of those words, unless by explicit instruction, and presumably they also would have difficulty remembering them. This puts selectional presure on the evolution of affixes: an affix is more highly valued if its function range constitutes a contiguous area. It now follows that a morphological process of the type $x + affix \rightarrow y$ has a greater chance to arise if the range of y is more systematically constrained.

The tendency to systematicity is a selectional bias, not a principle of language. It leaves open the possibility to find affixes with a function range that consists of rather haphazard blotches with lots of empty space in between. For instance, there would be nothing surprising about a language that has an affix, say, -*prüt*, that forms causatives from, say, many or most class VIII verbs — whatever that may be — and also diminutives from a small but remarkable group of feminine nouns. On the (reasonable) assumption that causatives and diminutives are located in very different areas in the virtual lexicon, it would be surprizing to find in some other, historically unrelated, language an affix that would have almost exactly the same range of atomic functions. The chances that such an affix evolves on two independent occasions are vanishingly small, because no selectional considerations would favor a development in that direction. These considerations make comparative morphology a valuable tool for investigating the innate properties of the lexicon.

In sum, the only way to make sense of the syncretisms of the type found in table 1.1 is to assume that the lexicon is structured. This structure provides a measure of similarity, which, in turn, underlies the selectional pressure to which the evolution of affixes is subject: affixes strive to have homgeneous conglomerates of functions. In this sense, affixes can be said to partition the lexicon. The data discussed in section 2 can now be interpreted as showing that each of the three affixes in table 1 partitions the lexicon in (roughly) the same way; the same goes for the two affixes in table 2. This proposal will play a key role in the argument to be developed below.

4. Voice

4.1. Overview. The central argument in this paper concerns the status of the Tagalog affixes -*an* and *i*- within the Tagalog voice system. The two affixes have close counterparts in Indonesian and Dutch, as shown earlier. In the next two subsections, I will demonstrate that the Indonesian and Dutch affixes are neutral as to voice. Section 4.3 will deal briefly with some general questions concerning the Tagalog voice system.

4.2. Voice in Dutch. The voice system of Dutch resembles that of English so closely that it hardly deserves discussion here. It involves a morphologically derived participial form of the verb and a copula, *worden* 'be (passive)' (the verb also has other uses, including 'become'). The participle is formed from the stem by either suffixing it with -*d* or -*t*, or by ablaut, -*en*-suffixation, suppletion etc. etc. in the case of strong stems (*built, drunk, taken, come*). In addition, however, a prefix, *ge*-, is added, unless the verb stem already contains a prefix. Thus, *slaan* 'to hit' yields *geslagen*, but *beplanten* 'to plant (st) with (st)', which has prefix *be*-, yields *beplant*, not **gebeplant*. In this limited sense, there is interaction between affixation with *be*- and the voice system. Like in English, the agent is demoted into an optional 'by'-phrase. Sample sentences are given in (1). Note that affixation of *planten* 'to plant' with *be*- yields a verb that only occurs in the 'spray-with' construction. Like in English, it is the direct object that passivizes.

Given that in this pertinent construction, the "location" of planting is the direct object, it will become subject. This is, strictly speaking, not interaction of the two morphological systems. The one takes as input the output of the other.

(1)   *Gerard beplant de tuin met bloemen*
       Gerard be.plant the garden with flowers
       'Gerard plants the garden with flowers'

       De tuin werd beplant met bloemen (door Gerard)
       the garden was be.plant with flowers (by Gerard)
       'The garden was planted with flowers (by Gerard)'

4.3. Voice in Indonesian. Like most Austronesian languages in the area, Indonesian has a passive voice form characterized by two important, presumably correlated properties that clearly set it apart from the passive in English c.s. In the first place, the agent is a bare, non-demoted noun phrase; secondly, passive sentences do not have the

peculiar discourse-functional properties of passive sentences in English. That is, the proper translation of a passive sentence in Indonesian is very often a plain active sentence in English (Myhill 1992). The non-demoted agent is verb-adjacent. I assume that it is marked with genitive case. Cases are not morphologically marked in Indonesian; however, the typology of Austronesian languages suggests that this is so, as I have argued elsewhere (Voskuil 1993).

The Indonesian transitive verb has four voice forms: morphological active and passive, and bare active and passive.[1] The active/passive alternation is characterized by a difference in case patterns: nominative agent, accusative patient (active); and nominative patient, genitive agent (passive). The morphological/bare voice alternation is characterized by the form of the verb (presence vs. absence of *meN-* (active) and *di-* (passive)), and constituent order (agent-verb in the bare passive, verb-agent in the morphological passive). In the bare passive, the agent must be referential, preferably a proper name or referential pronoun. In the morphological passive, a genitive agent must be third person. The morphological passive can also take an agent in a *by*-phrase, in which case there is no person restriction on the agent; or it can take no agent at all. In those cases, the sentence has all the discourse-functional properties of a passive sentence in English (Myhill 1992). There are many tests for subjecthood; these show that in those sentence types analysed as passive, it is the deep-object (the ''patient'') that has the subject function. For discussion of these tests, I refer to Chung (1976).

(2)   *Dia   membaca   buku   itu*   (morphological active)
        3s   meN.read book   that
        'She reads the book'

(3)   *Buku   itu   dibaca   -nya*   (morphological passive)
        book   that   di.read 3s
        'She reads the book (or: the book is read by her)'

---

[1] There is some controversy about the bare active; prescriptive and descriptive sources alike often claim that the construction is unacceptable or, at the very least, ''sub-standard''. I submit that the construction is typical of *colloquial* Indonesian, but corpus research and research with informants — on which I hope to report on another occasion — lead me to the conviction that the bare active is perfectly ordinary, even ''standard'' Indonesian.

(4)   *Buku   itu   dibaca   oleh -nya*   (morph. passive + *by*)
        book   that di.read   by   3s
        'The book is read by her'

(5)   *Dia   telah   baca   buku   itu*   (bare active)
        3s   PERF   read   book   that
        'She has read the book'

(6)   *Buku   itu   telah dia baca*   (bare passive)
        book   that   PERF   3s   read
        'She has read the book'

In one sense, lexical morphology of the type discussed in section 2 does not interact with voice at all — broadly speaking. The affixes in question, *-i* and *-kan*, derive verbs from a variety of bases. The derived verb simply behaves as underived transitive verbs of the type *baca* 'to read' do: it has four voice forms, etc. etc.

In another sense, it does interact, namely, in as far as it alters the subcategorization properties of the base; thus, *tanami* 'to plant (st) with (st)' stands in a different semantic relation to the direct object than *tanam* 'to plant (st)'. But the direct object, independent of its semantic interpretation, is syntactically just a direct object. The following examples show this for the morphological voices; the bare voices work analogously.

(7)   *Mayang   menanam   bunga mawar   di   kebun*   (active)
        Mayang   meN.plant roses   in   garden
        'Mayang planted roses in the garden'

(8)   *Bunga mawar   ditanam   Mayang di kebun*   (passive)
        roses   di.plant   Mayang in garden
        'Roses were planted in the garden by Mayang'

(9)   *Mayang   menanami   kebun   dengan   bunga mawar* (active)
        Mayang   meN.plant.i garden with   roses
        'Mayang planted the garden with roses'

(10)   *Kebun ditanami   Mayang dengan bunga mawar*   (passive)
       garden di.plant.i   Mayang with   roses
       'The garden was planted with roses by Mayang'

Thus, there are two systems of verbal morphology, a verb-forming one, and one concerned with voice. They can be thought of as independent algorithms. Following certain rules, the one derives a transitive verb; the other, following other rules, derives a voice form — from a transitive verb.

In the details, there appear to be "real" interactions — cases where unexpected distributional properties in both systems correlate. For instance, in Classical Malay, verbs with *-i* are overwhelmingly passive, in more than 95% of the cases (Cumming 1994). This is an important fact, because, as we will see, in Tagalog a similar situation is found: the same verbs, derived with *-an* or *i-*, are 100% passive — they do not even show explicit markers for being passive.

4.4. Voice in Tagalog. Bloomfield's (1916) definition of a *passive verb* is a transitive verb in a construction where something else than the agent is the subject. Is what many linguists, including Bloomfield (1917), consider the subject in Tagalog indeed the subject? If not, then the Tagalog passive is not a passive. I will argue that the question is only indirectly important to the issue of Tagalog verbal morphology, because the pertinent affixes — *-an* and *i-* — are not part of the passive (or non-passive) system at all. This is a controversial standpoint, since the morphology is standardly taken to be the very engine that drives the "voice" machinery. It is therefore important to see clearly what the issues are that surround the notions subject, passive, and the like.

Let me first describe the Tagalog voice system in essentially Bloomfield's terms, and then explain the problems that are often thought to suggest a different analysis. Consider the sentences in (11) and (12), which are built around the verb *basa* (-*um*-, -*in*) 'to read'.

(11)   *Bumasa   siyá   ng   libro*
       PERF.um.read   3s.NOM   ACC   book
       'He read a book (not: He read the book)'

(12)   *Binasa   niyá   ang   libro*
       PERF.read.in   3s.GEN   NOM   book
       'He read the book'

The parenthesized affixes denote the affix class the verb belongs to: it forms its active with *-um-*, and its passive with *-in*. This particular passive affix is dropped when the verb carries the perfective or, in combination with reduplication, imperfective marker for passive verbs, the infix *-in-*, so that it does not show in (12). In my terms, the case marker *ng* (pronounced "nang") in (11) marks accusative case, while it marks genitive case in (12) — it does not show up, since pronouns have suppletive forms instead of case markers, but cf. *Binasa ng lalaki ang libro* "The man (GEN) read the book (NOM)," with *ng lalaki* "man" for the 3s genitive pronoun *niyá* 'his'. The subject is preceded by the nominative marker *ang*. Note that the passive sentence does not have the discourse functional properties of passive sentences in English; it normally translates into an active sentence in English, as do Indonesian passive sentences (Myhill 1992).

The distinction between accusative and genitive case is not always made because they look identical (cf. *ang libro niyá* "NOM book 3s.GEN, i.e., his book" and *ang libro ng lalaki* "NOM book GEN man, i.e., the man's book"). But the two functions express systematically different semantic relations — the genitive agent function expresses the agent-relation, the accusative direct object function expresses a broad range of relations, as in English. Moreover, there is a constraint on definite direct objects, which does not hold for the genitive agent. As indicated, (11) cannot be interpreted as if the book were a definite book (Adams and Manaster-Ramer 1988). Although the precise extent of the condition is unclear, and although there exist specific classes of counterexamples, the broad outlines are beyond doubt. Pronouns, which are clearly definite, do not occur in the object function; a sentence like "He hit me" comes out as passive (*Pinukól niyá akó*, with nominative 1s pronoun *akó*); the active, with nominative 3s pronoun *siyá* and genitive 1s *ko*, is completely out (**Pumukól ko siyá*).[2] Thus, there is a distinction to make. Cross-linguistic evidence leads to associating the passive agent with genitive case, and the active direct object with accusative (Voskuil 1993).

The question that concerns us at this point is whether the alternation in (11)–(12) is a passive alternation at all. There are many facts that would suggest that it does, but problems exist, as first noted by Schachter (1975). The most fundamental of these

---

[2] There is an alternative strategy, which consists in marking the definite object with oblique case (*Pumukól siyá sa akin*). Although I will frequently make use of this strategy in examples to be discussed, the precise nature of this construction is not clear. Examples in the literature typically involve verbs that undergo the conative alternation in English (*He hit (at) me*); but it is unclear whether *all* verbs systematically allow for this.

concern binding and control: the "genitive agent" relates to the *ang*-phrase as if the former were a subject, and the latter a direct object that has undergone topicalization of sorts. If true, that would entail *ang* should be glossed as a topicalization marker — objects can, as it happens, be topicalized —, and the "genitive agent" in (12) is really the nominative subject. For further discussion, I refer to Richards (this volume) and references cited there.

A good argument in favor of "*ang*-marks-subject" and against "*ang*-marks-topic" is that only noun phrases that are selected by the verb can be marked with *ang* — below, I will further sharpen this to the claim that only the agent and the deep-object of the verb can be thus marked. In other words: only noun phrases with a specific, structurally defined deep grammatical function can be marked *ang*. The range of thematic relations that a direct object can bear vis-à-vis the verb is somewhat broader than in English (on account of its richer morphology). However, the range of elements eligible for *ang*-marking is incomparably smaller than the range of elements that can be topicalized in English, and especially so in German, Dutch, and Icelandic. In these verb-second languages, *any* constituent of the sentence can move to the preverbal topic position: prepositional phrases, adjunct noun-phrases like *Sunday* in *Sunday I will arrive*, adverbials of all denominations, certain types of floated quantifiers, etc. etc. Verb-second languages are especially relevant here, since Richards (this volume) analyses *ang*-promotion as parallel to topicalization in Icelandic.

A typology of landing sites that distinguishes between noun phrase positions and XP-positions — positions available to anything — clearly suggests that promotion to the function marked by *ang* involves landing in a noun phrase position: A-movement (Chomsky 1981). This seems to tip the balance in favour of the "*ang* marks nominative" analysis. In any case, however, the question is orthogonal to the central claim of the present paper, which is that the verbal affixes *-an* and *i-* are not directly involved in the voice system — or topicalization system. For concreteness I will continue to refer to *ang* as a nominative marker, and designate verbs with *-an* or *i-* as passive. If we would have to conlude that this is wrong, the argument still stands.

5. Voice and lexical morphology in Tagalog

Tradition has it that the Tagalog voice system is rich in distinct types of passivization rules; I call this the multiple passive analysis. A verb like *taním* (*mag-, i-, -an*) 'to plant' has, besides an active with *mag-*, two passive forms, each of which promotes a different argument of the verb to subject, as shown in (14) and (15): *itaním* has as

subject the things planted (flowers), *tamnán* has as subject the "location" of the planting (the garden). The affixes pick out a labelled argument of the verb — agent, patient, location, etc. etc. — and promote it to nominative.

(13) *Nagtaním   siyá   ng   bulaklák   sa   hardin*
      PERF.mag.plant 3s.NOM   ACC   flowers   OBL   garden
      'She planted flowers in the garden'

(14) *Itinaním   niyá   ang   bulaklák   sa   hardin*
      PERF.i.plant   3s.GEN   NOM   flowers   OBL   garden
      'She planted the flowers in the garden'

(15) *Tinamnán   niyá   ng   bulaklák   ang   hardín*
      PERF.plant.an GEN.3s   ACC   flowers   NOM   garden
      'She planted the garden with flowers'

Surprisingly, perhaps, the same verb has the same properties and more in English. It has a passive form that promotes the thing planted (*The flowers were planted in the garden*) and it has one that promotes the "location" argument (*The garden was planted with flowers*). That English fails to make a morphological difference between the two passives is a result of its poor morphology — Indonesian and Dutch do make such a difference. Unlike Tagalog, however, English, Dutch, and Indonesian have an active construction that corresponds to the "locational passive": *Celia planted the garden with flowers*. The Tagalog active verb *magtaním* realizes the 'plant-in' alternant, as in (13), but not the 'plant-with' alternant, as shown in (16). Similarly, *tamnán* in (15) realizes the 'plant-with' alternant in the passive, but there is no active counterpart, with either *mag-* or *-um-* (the two active voice markers); there is no *magtamnán* or *tumamnán*.

(16) * *Nagtaním   siyá   ng   hardín   ng   bulaklák*
       PERF.mag.plant NOM.3s   ACC   garden   ACC   flowers
       'He planted the garden with flowers'

In this respect, *-an* differs from Indonesian *-i* and Dutch *be-*: although it marks the 'plant-with' alternant in the 'plant-in/plant-with' alternation, like its Indonesian and Dutch counterparts, the derived verb exclusively occurs in the passive voice.

Note, furthermore, that Tagalog verbs derived with *i-* or *-an* not only require the

passive voice, but also fail to carry overt, independent markers of the passive voice. Thus, whereas the Indonesian verb *tanami* 'plant-with' carries the usual markers of passive voice when it occurs in that voice (for instance, the passive marker *di-* in the morphological passive, see ex. (10) above), the Tagalog verb *tamnán* 'to plant-with' has no voice morphology at all. Specifically, it does not carry the passive marker *-in*, exemplified in (12), over and above the suffix *-an*; there is no such thing as *tamnanín*.[3] Tagalog verbs derived with *i-* or *-an* are passive by default.[4]

The multiple passive analysis holds that the Tagalog "voice" affixes directly encode the semantic relation that the subject bears to the verb. Ultimately, the present view fully endorses this: it cannot be denied that the pertinent relation covaries with the affix on the verb. However, the present view fundamentally differs from the traditional one in the way in which this encoding is established. My analysis posits two levels: a "lexical" level, and a "voice" level. The affixes reflect the semantic relation between the verb and its deep-object; the voice system promotes the deep object to subject. Thus, what holds true of the interaction between lexical verb morphology and voice in Indonesian, also holds true of Tagalog: the two systems can be thought of as independent algorithms. Following certain rules, the one derives a transitive verb; the other, following other rules, derives a voice form — from a transitive verb.

Schematically, the argument in favor of the present view — which I call the two-stage analysis of the Tagalog voice system — is as follows.

1. The function of certain affixes in Indonesian and Dutch is to derive specific types of verbs from specific types of stems.

2. The partitionings in the verbal lexicon established by these affixes are identical (at the correct level of abstraction, thus allowing for reasonable margins).

---

[3] However, passive verbs have a different aspectual paradigm. The perfective and imperfective aspects are formed with an infix, *-in-*, which does not occur on active verbs. The issue is rather complicated, because active verbs with *mag-* have *nag-* in the pertinent aspects. Historically, the *mag-* ~ *nag-* alternation can be shown to involve the infix *-in-*: *nag-* can be traced back historically to **minag-* (Wolff 1971).

[4] Recall that Dutch verbs with *be-* refuse the passive marker *ge-* (see 3.2), a similarity that I will not explore.

---

3. The functioning of the pertinent affixes in both languages is independent of the functioning of the voice system.

4. The Tagalog affixes *i-* and *-an* establish the same partitionings in the verbal lexicon as those referred to in 2.

5. Therefore, there must be level at which the functioning of the Tagalog affixes is independent of the functioning of the voice system.

The proper level is the interface between word-formation and voicing. A central notion is DEEP-OBJECT, which I define as the noun phrase that is accusative in the active and nominative in the passive (it has nothing to do with "deep-structure"). Thus, the Tagalog affixes reflect the semantic relation between the deep-object and the verb; next, the voice system passivizes the verb, promoting the deep-object to nominative. The passive rule has nothing to do with (is blind to) the semantic relation between the verb and the deep object. If this is correct, Tagalog only has one single passive alternation.

The two-stage analysis is not merely terminologically different from the traditional view. Let me briefly discuss an important empirical difference. Consider once again the sentence in (15) above, *Tinamnán niyá ng bulaklák ang hardín* "She planted the garden with flowers." The two-stage analysis leads to the following conclusion: *ang hardín* "the garden" is nominative, hence, it is the deep-object. This entails that *ng bulaklák* "flowers" is not the deep-object. Rather, it must have a status comparable to *flowers* in *to plant the garden with flowers*: it is a prepositional complement of sorts. Thus, although the noun phrase looks exactly like the direct object in (13) (*Nagtaním niyá ng bulaklák sa hardin* "She planted flowers in the garden"), the two-stage analysis forces us to posit a fundamental difference between the two.

The multiple passive analysis, on the other hand, is forced to analyse both instances as accusative direct objects: they are structurally identical. The affix on the verb in (15), *-an*, promotes the location argument of the verb to subject; the verbal affix in (13), *mag-*, promotes the agent. By definition, neither of the affixes interacts with the theme argument (flowers); this argument is a constant factor.

Strikingly, there exist important syntactic differences between *ng bulaklák* "flowers" in (13) and *ng bulaklák* "with flowers" in (15), which I have discussed elsewhere (Voskuil 1994) — there, I argue in addition that these differences suggest that *ng bulaklák* in (15) is indeed a prepositional complement, with an invisible preposition. The main difference is that *ng bulaklák* in (13) behaves as an ordinary

direct object, that is, it cannot be definite, and undergoes alternations typical of direct objects. *Ng bulaklák* in (15) has none of these properties. It does not behave as a direct object.

In the above discussion I have consistently referred to *ang* as a nominative marker. An alternative theory exists: *ang* is a topic marker. The outcome of that debate will have consequences for the analysis of the Tagalog voice system. But the issue of the multiple passive vs. the two-stage analysis will remain largely unaffected by whatever the outcome is. This can be seen from the fact that both sides in the subject-or-topic debate assume the traditional multiple passive analysis. What the one glosses as ''location-topic marker,'' the other will gloss as ''location-subject marker'' (compare, for instance, Kroeger (1993) and Carrier-Duncan (1985)). The two-stage analysis is especially immune for developments on that score, since it disconnects the functioning of the affixes from the functioning of the ''voice'' system — or whatever it is.

REFERENCES

BAKER, G.P. AND P.M.S. HACKER. 1984. *Language, Sense and Nonsense*. Blackwell, Oxford.

BLOOMFIELD, L. 1916. 'Subject and Predicate.' Paper presented at the 48th Annual Meeting of the APA. Reprinted in C.F. Hockett (ed.), *A Leonard Bloomfield Anthology* (abridged edition), University of Chicago Press, Chicago, 1987, pp. 47–54.

BLOOMFIELD, L. 1917. *Tagalog Texts with Grammatical Analysis*. University of Illinois Press, Urbana.

CARRIER-DUNCAN, J. 1985. 'Linking of Thematic Roles in Derivational Word Formation.' *Linguistic Inquiry* 16, 1–34.

CHOMSKY, N. 1981. *Lectures on Government and Binding*. Foris, Dordrecht.

CHUNG, S. 1976. 'On the Subject of Two Passives in Indonesian.' In C.N. Li (ed.) *Subject and Topic*. Academic Press, New York, pp. 57–98.

CUMMING, S. 1994. 'Semanticization and the Malay/Indonesian applicatives.' Paper presented at the 7th International Conference on Austronesian Linguistics (ICAL 7), Leiden University, August 1994.

DENNET, D. 1995. *Darwin's Dangerous Idea*. Allan Lane/The Penguin Press, London.

KROEGER, P. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. CSLI Publications, Stanford, California.

MYHILL, J. 1992. *Typological Discourse Analysis. Quantative Approaches to the Study of Linguistic Function*. Blackwell, Oxford.

REICHLING, A. 1935. *Het Woord*. Tjeenk-Willink, Zwolle.

RICHARDS, N. (this volume). 'Another Look at Tagalog Subjects'

SCHACHTER, P. 1976. 'The Subject in Philippine Languages: Topic, Actor, Actor-Topic, or

None of the Above.' In P. Li (ed.), *Subject and Topic*. Academic Press, New York.

UHLENBECK, E.M. 1978. *Studies in Javanese Morphology*. Martinus Nijhoff, The Hague.

VOSKUIL, J.E. 1993. 'Abstract Case and Malagasy.' Unpublished ms.

VOSKUIL, J.E. 1994. 'The Tagalog Absolute Accusative.' Paper presented at the 7th International Conference on Austronesian Linguistics (ICAL 7), Noordwijkerhout, The Netherlands, August 1994.

WOLFF, J.U. 1971. 'Verbal Inflection in Proto-Austronesian.' In A.B. Gonzales, *Parangal kay Cecilio Lopez*, pp. 71–91. Linguistic Society of the Philippines, Quezon City.

Dept. of Linguistics
Leiden University
P.o. Box. 9515
2300 RX Leiden
The Netherlands
[voskuil@rullet.leidenuniv.nl]

NUSA
LINGUISTIC STUDIES OF INDONESIAN
AND OTHER LANGUAGES IN INDONESIA

NUSA volume 1
John W.M. Verhaar (ed.), *Miscellaneous Studies in Indonesian and Languages in Indonesia,* Part I, Jakarta 1975.

Rp 6.500,00/US$ 5.00

NUSA volume 2
Dale F. Walker, *A Grammar of the Lampung Language,* Jakarta 1976.

Rp 6.500,00/US$ 5.00

NUSA volume 3
Amran Halim (ed.), *Miscellaneous Studies in Indonesian and Languages in Indonesia,* Part II, Jakarta 1977.

Rp 7.000,00/US$ 5.00

NUSA volume 4
Soepomo Poedjosoedarmo (ed.), *Miscellaneous Studies in Indonesian and Languages in Indonesia,* Part III, Jakarta 1977.

Rp 7.000,00/US$ 5.00

NUSA volume 5
Ignatius Suharno (ed.), *Misellaneous Studies in Indonesian and Languages in Indonesia,* Part IV, Jakarta 1977

Rp 7.000,00/US$ 5.00

NUSA volume 21/22
Keith Michael McCune, *The Internal Structure of Indonesian Roots,* Part I, Jakarta 1985.
Rp 14.500,00/US$ 11.50

NUSA volume 23
Keith Michael McCune, *The Internal Structure of Indonesian Roots,* Part II, Jakarta 1985
Rp 12.500,00/US$ 11.50

NUSA volume 24
Gloria Risser Poedjosoedarmo, *Role Structure in Javanese,* Jakarta 1986.
Rp 7.500,00/US$ 8.00

NUSA volume 25
Soenjono Dardjowidjojo (ed.), *Miscellaneous Studies of Indonesian and Other Languages in Indonesia,* Part VIII, Jakarta 1986.
Contents: Gloria Poedjosoedarmo (Subject Selection and Subject Shifting in Indonesia), Mark Durie (The Use of Passive-Like Strategies in Acehnese), Gloria Poedjosoedarmo (The Symbolic Significance of Pharyngeal Configuration in Javanese Speech; Some Preliminary Notes), Joan M. Rosen (Phonemes, Verb Classes and Personal Endings in Maumere), Author Index NUSA volumes 18-23, 1984-1985.

Rp 5.750,00/US$ 7.50

NUSA volume 26
Janet Bateman, *Iau Verb Morphology,* Jakarta 1986.

Rp 6.500,00/US$ 7.50

NUSA volume 27
Soenjono Dardjowidjojo (ed.), *Miscellaneous Studies of Indonesian and Other Languages in Indonesia,* Part IX. Jakarta 1987.

Rp 8.500,00/US$ 9.50

NUSA volume 30
Bambang Kaswanti Purwo (ed.), *Towards a Description of Contemporary Indonesian: Prelimanary Studies* III. 1988.

Rp 9.500,00/US$ 9.50

NUSA
LINGUISTIC STUDIES OF INDONESIAN
AND OTHER LANGUAGES IN INDONESIA

NUSA volume 6

John W.M. Verhaar (ed.), *Miscellaneous Studies in Indonesian and Languages in Indonesia,* Part V, Jakarta 1978.
Contents: J.V. Dreyfuss (*meN, di-, ber-* in Indonesian), Michael R. Thomas (Unmarked Verbs in Indonesian), J.W.M. Verhaar (Passive in Indonesian), Stephanus Djawanai (Nga'da Discourse), L.L. Kamengmai and W.A.L. Stokhof (Woisika Text).

Rp 1.600,00/US$ 5.00

NUSA volume 7

Amran Halim (ed.) *Miscellaneous Studies in Indonesian and Languages in Indonesia,* Part VI, Jakarta 1979.
Contents: J.V. Dreyfuss (Definition of 'Nouniness' in Indonesia), Sofjan Abdurrahman and Colin Yallop (Komering Phonology), R.H. Barnes (An Austronesian Relationship Name), Don A.L. Flassy and W.A.L. Stokhof (A Note on Tehit).

Rp 2.250,00/US$ 5.00

NUSA volume 8

Bintaro, *Javanese Transitive Verbs: A Tagmemic Grammar,* Jakarta 1980.

Rp 1.500,00/US$ 4.00

NUSA volume 9

Kay Ikranegara, *Melayu Betawi Grammar,* Jakarta 1980.

Rp 4.175,00/US$ 4.00

NUSA valume 10

Robert A. Blust (ed.), *Historical Linguistics in Indonesia,* Part I, Jakarta 1981.

Rp 2.450,00/US$ 5.50

NUSA volume 11

Muhadjir, *Morphology of Jakarta Dialect, Affixation and Reduplication,* Jakarta 1981.

Rp 3.350,00/US$ 6.50

NUSA volume 12

Ellen Rafferty, *Discourse Structures of the Chinese Indonesian of Malang,* Jakarta 1982.

Rp 2.100,00/US$ 5.50

NUSA volume 13

Alan T. Walker, *A Grammar of Sawu,* Jakarta 1982.

Rp 2.300,00/US$ 5.50

NUSA volume 14

Richard McGinn, *Outline of Rejang Syntax,* Jakarta 1982.

Rp 2.750,00/US$ 6.50

NUSA volume 15

Soenjono Dardjowidjojo and John W.M. Verhaar (ed.) *Miscellaneous Studies of Indonesian and Other Languages in Indonesia,* Part VII, Jakarta 1983.
Contents: Edmund A. Anderson (The Meaning of Variation in Indonesian), Soenjono Dardjowidjojo (A Classifier, *itu, -nya,* or None of the Above: The Way the Indonesian Mind Operates), Rachel Fleming (Cohesion in Wandamen Narrative), Joan M. Rosen (Rembong and Wangka: A Brief Comparison), Ellen van Zanten and Vincent J. van Heuven (A Phonetic Analysis of the Indonesian Vowel System: A Preliminary Acoustic Study).

Rp 2.750,00/US$ 6.00

NUSA volume 16

James T. Collins (ed.), *Studies in Malay Dialects,* Part I, Jakarta 1983.
Contents: Ellen Rafferty (Contrasts in the Syntax of a Malay and an Indonesian Oral Narrative), C.D. Grijns (Language-Use in Three Neighbourhoods of Tanah Abang: The Speaker's View), Jeff Dreyfuss (The Backwards Language of Jakarta Youth (JYBL), a Bird of Many Language Feathers), K.A. Adelaar (Malay Consonant-Harmony: An Internal Reconstruction), Author Index NUSA volumes 1-14, 1975-1982.

Rp 2.100,00/US$ 5.50

NUSA volume 17

James T. Collins (ed.), *Studies in Malay Dialects,* Part II, Jakarta 1983.
Contents: C.L. Voorhoeve (Some Observations on North-Moluccan Malay), Paul Michael Taylor (North Moluccan Malay: Notes on a 'Substandard' Dialect of Indonesian), James T. Collins (Syntactic in Ambonese Malay: The Possessive Construction), Hein Steinhauer (Notes on the Malay of Kupang (Timor)), W. Seiler (The Lost Malay Language of Papua New Guinea), Author Index NUSA volumes 15-17, 1983.

Rp 2.750,00/US$ 6.00

NUSA volume 18

Bambang Kaswanti Purwo (ed.), *Towards a Description of Contemporary Indonesian: Preliminary Studies,* Part I, Jakarta 1984.
Contents: John W,M. Verhaar (Affixation in Contemporary Indonesian), John W.M. Verhaar (The Categorial System in Contemporary Indonesian: Verbs).

Rp 2.100.00/US$ 5.50

NUSA volume 19
John W.M. Verhaar (ed.), *Towards a Description of Contemporary Indonesian: Preliminary Studies,* Part II, Jakarta 1984.
Contents: Russell Jones (Loan-Words in Contemporary Indonesian), J.W. de Vries (Adaptation of Polymorphemic European Loanwords in Contemporary Indonesian), Bambang Kaswanti Purwo (The Categorial System in Contemporary Indonesian: Pronouns).
Rp 2.750,00/US$ 6.00

NUSA volume 20
Jilis A.J. Verheijen, *Plant Names in Austronesian Linguistics,* Jakarta 1984.
Rp 3.000,00/US$ 6.50

NUSA volume 28/29
Olaf H. Smedal, *Lom-Indonesian-English & English-Lom Wordlists Accompanied by Four Lom Texts.* Jakarta 1987.
Rp 8.000,00/US$ 11.00

NUSA volume 31
James N. Sneddon (ed.). *Studies in Sulawesi Linguistics,* Part 1, Jakarta 1989.
Content: Timothy Friberg and Thomas V. Laskowske (South Sulawesi Languages), Michael P. Martens (The Badaic Languages of Central Sulawesi), Ulo Sirk (On the Evidential Basis for the South Sulawesi Language Group), J.N. Sneddon (The North Sulawesi Microgroups: in Search of Higher Level Connections).
Rp 8.500,00/US$ 11.50

NUSA volume 32
Bambang Kaswanti Purwo (ed.), *Miscellaneous Studies of Indonesian and Other Languages in Indonesia,* Part X, Jakarta 1990.
Contents: Janet Bateman (Pragmatic Functions of the Tone Morphemes on Illocutionary Force Particles in Iau), Janet Bateman (Iau Segmental and Tone Phonology), William U. Brown (Mai Bratt Nominal Phrases), E.O. van Reijn (Texts from Kumun (Kerinci)).
Rp 8.200,00/US$ 11.00

NUSA volume 33
J.N. Sneddon (ed.), *Studies in Sulawesi Linguistics,* Part II, Jakarta 1991.
Contents: René van den Berg (Muna Historical Phonology), Robert L, Busenitz and Marilyn J. Busenitz (Balantak Phonology and Morphophonemics), Nikolaus P. Himmelmann (Tomini-Tolitoli Sound Structures), Timothy Friberg and Barbara Friberg (Notes on Konjo Phonology).
Rp 10.250,00/US$ 13.50

NUSA volume 34
Donals A. Burquest and Wyn D. Laidig (eds.), *Descriptive Studies in Languages of Maluku,* Jakarta 1992.
Contents: Cindy Blood (Subject-verb Agreement in Meher), Craig Marshall and Sarah Marshall (Reduplication in Fordata), Masahiro Takata and Yuko Takata (Kola Phonology), Yuko Takata (Word Structure and Reduplication in Kola), Sandra G. Wimbish (Pagu Phonology).
Rp 9.500,00/US$ 12.50

NUSA volume 35
Jilis A.J. Verheijen, *Glossary of Taxonomic - Indonesian Plant Names.* Jakarta 1993.
Rp 8.500,00/US$ 11.50

NUSA volume 36
René Van den Berg (ed.), *Studies in Sulawesi Linguistics,* Part III, Jakarta 1994.
Contents: Robert L. Busenitz (Marking Focus in Belantak), Marjo Kahunen (The Noun Phrase in Padoe), Kathryn B. Laskowske (Negation in Seko Padang), David F. Matti (Mamasa Pronoun Sets), Kari K. Stromme (Person Marking in the Mamuju Language).
Rp 13.500,00/US$ 15.00

NUSA volume 37
René Van den Berg (ed.), *Studies in Sulawesi Linguistics,* Part IV, Jakarta 1995.
Contents: Susanne Valkama (Notes on Duri Transitivity), Kari Valkama (Person Marking in Duri), Paula Vuorinen (Person Marking in Padoe).
Rp 15.500,00/US$ 18.50

NUSA volume 38
Wyn D. Laidig (ed.), *Descriptive Studies in Languages of Maluku*, Part II, Jakarta 1995.
Contents: Edward A. Kotynski (Glottal Stop and the Nasal Prefix in Tabaru and Other North Halmaheran Languages), Wyn D. Laidig and Carol J. Laidig (A Synopsis of Larike Phonology and Syntax), Joost J.J. Pikkert and Cheryl M. pikkert (A First Look at Tidore Phonology), John A. Severn (The Phonemic Syllable in Sahu: A Computer-Aided Phonological Analysis), Kathleen B. Taber and Mark H. Taber (On Being Partially Pregnant: Transitivity in Luang).
Rp 16.000,00/US$ 19.00

NUSA volume 39
Marian Klamer (ed.), *Voice in Austronesian*, Jakarta 1996.
Contents: Robert Blust (Notes on the Semantics of Proto-Austronesian *-an 'locative'), Marian Klamer (Kambera Has No Passive), Norvin Richards (Subjects in Tagalog and Icelandic), Jan E. Voskuil (Aspects of Verbal Morphology in Tagalog).
Rp 15.000,00/US$ 18.00