Voice and valency alternations in Karo Batak

CLODAGH NORWOOD

1 Introduction

Karo Batak, the second largest of five distinct Batak languages,[1] is a western Malayo-Polynesian (WMP) language spoken in an area of the north-western Sumatran highlands extending south-east from the city of Medan to the large volcanic Lake Toba. Partly due to the relatively inaccessible highland location, the language and cultural traditions of the Karo have been maintained to a remarkable extent. A large number of Karo people adhere to a Christian organisation which actively promotes the use of the Karo language, while skill in song and speech-making during traditional ceremonies is highly regarded even by younger people.

Of the twenty-four groups that the WMP languages are currently, though tentatively, divided into (Ross 1995), the Batak languages as a whole fall into the North-West Sumatra/Barrier Islands group that comprises Gayo, Batak, Nias, Mentawai and Enggano. Nothofer (1991) put forward the hypothesis that this group of languages may be more closely related to an older Paleo-Hesperonesian family that once covered the whole WMP area, than to the more centrally located Hesperonesian languages in this area. Superceded by the later Hesperonesian group, the geographically peripheral languages of the area, such as those of the North-West Sumatra/Barrier Islands group, and northern Sulawesi and southern Philippines languages, are purported to share certain features by virtue of their outlying location in the western Austronesian speaking area.

A detailed descriptive grammar of Karo Batak is available (Woollams 1996). The current paper is part of work in progress which aims to describe the forms and functions of two Karo verbal affixes in relation to the Philippines-type 'focus' system, and ultimately to place Karo Batak in a typological categorisation of WMP languages.

A semantically transitive verb in Karo can occur with either no prefix (indicated by *Ø-* in the gloss for expository purposes) as in (1), or with a homorganic nasal prefix, *N-*, as in

1 The estimated number of speakers varies according to whether the relatively large numbers of speakers who live in other parts of Indonesia are included. The population of Taneh Karo itself is about 600,000.

Fay Wouk and Malcolm Ross, eds, *The history and typology of western Austronesian voice systems*, 181-208.
Canberra: Pacific Linguistics, 2002.

example (2). In focus terminology the *Ø*-prefix would be an undergoer focus affix and the *N*- prefix an actor focus prefix, but as Karo does not exhibit a fully functioning focus system, these terms are not appropriate, whereas the less restrictive term *voice* is (see Fox & Hopper 1994). Other grammatical terms used in this paper are described in §3.2.[2]

(1) 352[3]
*Ku-**jumpai** ləbe duana*
I-*Ø*-meet first both
'I'll meet them both first.'

(2) 309
*adi aku səkali **n-jumpai** ia*
if I once *N*-meet him
'if I ever meet him'

This paper examines the distribution and function of the actor voice (*N*- verb form) constructions, and that of their undergoer arguments. It demonstrates that these constructions are not simply syntactically transitive alternatives to the much more common undergoer voice *Ø*- forms. Rather the distribution of the *N*- constructions is quite circumscribed. They are found most frequently (75% of all *N*- constructions) in subordinate clauses where the deleted pivot is the agent. Those found in main clauses fall into a number of discrete groups with clearly identifiable functions, most of which relate to detransitivisation. Other main clause functions, notably topicalisation of the agent, can be explained as relics of an earlier focus system.

This study is based on the examination of the text of a recorded conversation consisting of over 750 conversational turns.[4] Transcribed and translated with assistance from a native speaker, it will henceforth be referred to as the text. Five hundred of these turns comprise spontaneous conversation between three participants about the raising of money through a proposed land sale and contain a number of semantically transitive verbs which are repeated throughout the conversation. Four other data sets were also checked in order to verify the conclusions drawn from analysis of the text, three written stories, and one short journalistic piece that included quoted conversation. Examples have been taken from the text wherever possible (text examples are preceded by a turn number).

2 Basic morphosyntax

2.1 Nominal marking

Table 1 shows the forms of the pronominal system.

---

2 DAT dative; DEF definite; DEM demonstrative; EMPH emphatic marker; EXCL exclusive; EXIST existential; FAM familiar; INC inclusive; INTER interrogative; PERF perfective marker; PL plural; PTL particle; PURP purposive; REL relative marker; SG singular; TRS transitiviser.

3 Intonation breaks should be assumed at the beginning and end of all examples but are shown by double slashes (//) when they occur in the same line of an example.

4 I wish to thank Ramli Ginting and his family for assistance with data gathering, in particular his mother who has since died and to whom I here wish to record my respect. Personal names have been changed in the text. I also thank Barry Blake and Nikolaus Himmelmann for critical comments, and Andyda Meliala and Edimon Ginting. In particular I want to thank Fay Wouk for her generosity and conscientious editing throughout. Any errors in interpretation or analysis are entirely my own.

Table 1: Pronominal paradigms[5]

	Free pronouns (FF)	Actor affixes/clitics of *Ø-* forms	Genitive affixes/clitics
1SG	*aku*	*ku-*	*-(ng)ku*
2SG (formal)	*kam*	*-ndu*	*-ndu*
2SG (familiar)	*(əng)ko mu*	*-ko*	*-m*
3SG	*ia*	*-na*	*-na*
1PL.INCL	*kita*	*si-*	*-(n)ta*
1PLEXCL	*kami*	*kami*	*kami*
2PL	*kena*	*kena*	*kena*
3PL	*kalak*	*-na/kalak*	*-na/kalak*

As Table 1 shows many of the A(ctor)[6] suffixes of the *Ø-* forms are identical to the corresponding genitive forms. However first person singular and plural inclusive As of *Ø-* forms are bound prefixes, not genitive suffixes.[7] Apart from the affixed As of the *Ø-* verbs, the remaining pronominal arguments of both verb forms (i.e. elsewhere) and the S(ingle) argument of intransitives are represented by independent pronouns, as shown in (3) for the P argument of a *Ø-* verb, and (2) above for both arguments of a *N-* verb.

(3) 635

Ku-idah ia rusur ku rumah ənda.
I-*Ø*-see her often at house this
'I see her often at this house'

Full NP arguments are not morphologically marked, but each voice form marks one argument by virtue of its constrained position vis-a-vis the verb. This marking follows the pronominal patterns in that the A of the *Ø-* verb forms must follow and be immediately adjacent to the verb, as must the P of the *N-* forms, except in those instances where the *N-* form is suffixed with *-sa* (§2.4 and §5.5).

Proper names in any function in Karo take the prepositional particle *si*, although it is mostly restricted to younger people's names as older people are addressed by relationship terms, which may include their clan name.

---

5 Third person pronouns, including possessives, are glossed in the examples according to the English meaning intended, not as 3SG or 3PL, and to be consistent, first and second person pronouns are also glossed in English.

6 The Dixon/Comrie terminology is used where A and P represent the arguments for Actor and Patient (or undergoer) respectively.

7 Himmelmann (1996:126) notes that in some South Sulawesi languages the occasional presence of a proclitic for first person in a paradigm that is otherwise comprised of enclitics is suggestive of incipient IRREALIS mode marking, while van den Berg (1996:91ff.) in the same volume takes the opposite view that the paradigm of proclitics represents reconstructable Goal Focus IRREALIS marking in the Proto Celebic focus system. The patterns in Karo and related languages need further investigation, the apparent lack of *i-/ni-*forms with first person *ku-/si-* being noteworthy, as are two prominent exceptions to first person proclitic forms. These are two commonly used lexical items which have a basically verbal function, but show nominal morphology by retaining the first person GEN suffixes (enclitics). They are *ate* 'wish' and *ni* 'say' (quotative). They will be glossed as nominals.

Dative marking with the preposition *man* (*+ba-* for pronominals) 'for/to', is used in ditransitive constructions. It is used to mark recipients, benefactives as in (6) below, and indirect objects of saying/telling etc., as well as the reflexive anaphor.

2.2 Verb morphology

In addition to the two voice prefixes under discussion, the other main verbal prefixes are *ər-* and *tər-*, *i-/ni-*, and a homorganic *N-* prefix which is used to derive most intransitive verbs. In this paper this intransitive *N-* prefix will be assumed to be merely homophonous with the *N-* voice prefix under discussion.

Intransitive verbs usually occur in one of three main lexically determined forms: they may be bare stem constructions, or marked with either the intransitive *N-* prefix as exemplified in (4) and (5), (which come from a written text), or with the *ər-* prefix. As noted above, the pronominal argument of intransitives is always an independent pronoun from Table 1. The preferred order is predicate initial, as in (4).

(4) *E maka* ***tedis*** *mə si Naktaki,* ***ng-andung*** *ia,* ***bərkat*** *ia.*
then stand *mə si* Naktaki *N*-cry he depart he
'Then Naktaki stood up, crying, and departed.'

However, pragmatic and syntactic factors may cause inversion of the V(X)S pattern, as with *mulih* in (5).

(5) ***Ng-andung*** *duana, sabab la i-sangka nande-na ia* ***m-ulih.***
N-cry both because not *i*-expect mother-her she *N*-return
'They both cried because her mother had not expected her to return.'

The very productive *ər-* prefix derives intransitive verbs from nominals and, with one exception, occurs with only one core argument, as in (6).

(6) 38
Si ənggo ər-dahin kena
REL PERF *ər*-work you

Sangap kena ibas dahin kena.
content you in work your
'Those of you who are already working, be content in your work.'

The exceptional form with this prefix, which is possibly a relic from an erstwhile transitive paradigm, is derived from the verb *b(ah)an* 'do/make'. This form takes two arguments and frequently occurs with the *-sa* suffix as in (7). Further investigation is required to clarify what this might imply about the historical development of the *ər-* prefix.

(7) 144
Uga ər-bahan-sa kita m-uat si məhuli?
how *ər*-make-*sa* we.INCL *N*-get REL good
'How do we make things good for ourselves?'

The *tər-* forms are basically passive verbs, as in (8). Accidental passives are included in this group.

(8) 215

Pərən ku-idah kərina juma ah // lanai ***tər-dahi***
grass I-Ø-see all garden DEM never *tər*-tend
'I noticed grass everywhere in that garden...not being tended.'

Second and third (but not first) person agents can be expressed with the *tər-* forms, second person using the same pronominal forms as *Ø-* verb forms, as in (9), and the third person using the enclitic *-sa*. However, two argument constructions which include the A argument are rare with *tər-*.

(9) 141

aku lah gia ***tər-usih*** *kena nak-ku*
I EMPH EMPH *tər*-follow you son-my
'if you follow me (my advice), my son'

The intransitive nasal verbs (including a class that glosses as adjectives) and other prefixed forms such as the *ər-* and *tər-* forms, will not be considered further here. Where they occur in examples from henceforth their derivational prefixes are not included in interlinear glossing.

The morphosyntactic structure of the *i-* forms, illustrated in (10), is basically the same as that of the *Ø-* forms, with one difference. Either there is no *i-* prefix on first person forms, or first person *i-* forms have been conflated with the *Ø-* forms.

(10) 209

O nande kai gundari ***i-suan-ndu*** *i juma-ta Pindan ah?*
oh mother what now *i*-plant-you in garden-our name DEM
'Oh Mum, what have you currently got planted in our Pindan garden?'

Speakers appear to regard *i-* as a freely variable allomorph of *Ø-*.[8] It is particularly prevalent in written genres, where *Ø-* forms are seldom used. But while there is a large degree of overlap, there is also evidence of a distinction between the two in the spoken varieties of the language.

Two criteria suggest that the *i-* forms, along with the now infrequently used *ni-* constructions, typically function as passives in spoken language. First, the *i-* forms have a low discourse frequency and second, the *Ø-* constructions occur with an A argument about 60% of the time, while the *i-* forms are much less frequently accompanied by an A argument. If there was once a modal distinction between the two forms, this no longer appears to be the case, although the prevalence of *i-* forms in written narratives is worth further investigation.

The *ni-* forms, which are said to be used only by older speakers, now occur rarely even in written genres; a comparison of passages of Neumann's translation of the Bible with the same passages from the modern version over fifty years later, for example, shows that where Neumann used *ni-* forms, the modern version uses *i-* forms or alternative phrasing.

There are also three valency increasing affixes: the causative prefix *pe-* and two applicative suffixes *-i,* and *-kən.* Verbs derived with *-kən* in particular have often become lexicalised, for example, *bəre* is often best glossed as 'allow', but *bərekən* means 'give'. One or two other verbal suffixes, including *-ən,* have limited productivity or very restricted distribution.

---

8 This observation is supported by Woollams (1996).

2.3 Emphatic particles

A possibly unusual feature in Karo is the presence of three frequent particles, *ngə*, *mə* and *pe*, which may be described as emphatic markers, although they are not optional in every context. Only one speaker uses them with very high frequency, mainly in contexts where their use is optional. Although they usually mark nominals, their specific distribution and function is the subject of ongoing research and they are presumed to be undergoing a functional change, moving perhaps from a vestigial form of case marking to discourse particles.[9] In (11), from a written text, as in many other cases, *ngə* appears to be associated with the undergoer of the *Ø-/i-* forms, while *mə* appears to be associated with the agent of these forms and *pe* with the agent of the *N-* forms, but these are not absolute correlations as many counterexamples testify.

(11) *tapi kərina ngə i-simbak-na*
but all *ngə* *i*-push.away-he
'but he pushed them all away'

2.4 The nasal verb suffix *-sa*

N- forms may take a suffix of the form *-sa* (or a phonologically determined allomorph *-ca*). Woollams (1996:115) claims this suffix is a third person object (undergoer) form for the *N-* forms. Certainly in some contexts it appears this way, as in (12), but in many other contexts the situation is not this simple.

(12) *Maka ngadi mə ia ng-ajar-**sa**.*
then stop *mə* she *N*-learn-*sa*
'Then she stopped teaching (them).'

The suffix *-sa* cannot freely be used in all contexts with *N-* forms. As well as examples like (12), where it appears to substitute for the independent third person pronoun *ia*, it may also co-occur with this pronoun, as well as with first or second person undergoers in a few cases. This suffix is discussed in more detail in §5.5, where it will be argued that its function with the Karo *N-* forms is similar to that of the oblique case marker, *sa* in Tagalog which is used to mark specific, individuated undergoers of the actor focus forms.

3 The *Ø-* verb forms

3.1 Formal properties

The *Ø-* form clauses function as the basic active construction in Karo. They show an ergative morphological pattern whereby only the pronominal A(ctor) argument is a verbal affix. Table 2, which shows the forms of the affixes/clitics, is virtually a repeat of the second column of Table 1. In written forms of the language the prefix *i-* is usually used instead of *Ø*.

---

9 They are present in many of the examples where they are glossed with their own form to distinguish them from other particles.

Table 2: The *Ø-/i-* verb form affixes/clitics

	Ø-/i- form affixes/clitics = A
1SG	*ku-* (*i-* prefix not overt in first person form)
2SG	*-ndu/-ko*
3SG	*-na*
1PL.INCL	*si-* (*i-* prefix not overt in first person form)
1PL.EXCL	FF*
2PL	FF
3PL	*-na*/FF

* FF = free form pronouns as shown in Table 1.

Unlike the A argument of these verbs, their undergoer (P) argument is not a suffix or clitic. Pronominal Ps take the free forms shown in Table 1, while any lexical NPs are marked with *si* only if they are proper names. Examples (13) and (14) show pronominal A arguments in first and second person respectively, and example (15) a full NP A.

(13) 338
ənta ku-jumpai si Neli.
let I-Ø-meet PN Neli
'Allow me to meet Neli.'

(14) 351
*Jadi ate-ndu **jumpai**-ndu ləbe bibi nguda?*
become wish-your Ø-meet-you first aunt young
'Do you want to meet younger auntie first?'

(15) 412
***Uukur** bapa-na ah motor ngə kap.*
Ø-buy father-his DEM car ngə EMPH
'His father bought a car (for him).'

3.2 Grammatical relations

If one assumes the existence in WMP languages of a coherent and underlying system of grammatical relations (but see Himmelmann 1996, especially p.121, who diverges from this point of view) the undergoer or patient argument (P) is the subject of the *Ø-* verb constructions.

Subject is defined here as the most privileged grammatical relation, sometimes called pivot (or trigger in Indonesian languages). It is not restricted to semantic agents, the A (or S) relations. Thus the P argument of the ergatively patterning *Ø-* forms, as well as the A argument of the accusatively patterning *N-* forms, and S, the single argument of intransitives, are all subjects. The other two arguments, the A of the *Ø-* forms, and the P of the *N-* forms will be called non-subjects. In one section (5.5), where an analysis of the suffix *-sa* is proposed, the non-subject P of the *N-* forms is referred to as a potential object.

These syntactic terms, subject and non-subject, are only used in discussion of syntactic functions. Otherwise as long as reference to the appropriate verb form is unambiguous, the semantically based terms actor and undergoer are used, often abbreviated as A and P respectively.

Returning now to the description of the *Ø-* forms, both the P and the A arguments of these constructions demonstrate a number of core properties, the P relation exhibiting the larger number of them. There is a strong tendency for this argument to be definite and specific but this is not an absolute constraint. In the following description of core properties some examples are elicited as no illustrative examples were available from the corpus.

3.2.1 Relativisation

The basic gap strategy used for relativisation in Karo is exemplified first of all with an intransitive S in (18). In (19) the relativised and gapped argument is a *Ø-* construction P. Square brackets, [ ], arbitrarily placed after the relativiser, represent the relativised argument.

(18) 333
Lit *dənga* *ka* *bibi* *si* [ ] *nggəluh.*
EXIST still PTL auntie REL live
'There is auntie who is still living.'

(19) 773
Lit *dənga* *ngə* *sada* *kaset* *nari* *si* [ ] *ənggo* *ku-isi.*
EXIST still *ngə* one tape from REL PERF I-*Ø*-fill
'There is another tape I have already filled.'

The A argument of the *Ø-/i-* constructions can also be relativised,[10] as in the elicited example (20). With these arguments the strategy is one of pronoun retention. Woollams (1996:293) gives examples whereby possessor and dative arguments can also be relativised by this same strategy.

(20) *Guru* *ng-ərgai* *murid* *si* *ingət-[na]* *pəlajarən.*
teacher *N*-scold pupil REL forget-[she] lesson
'The teacher was scolding the child who forgot her lessons.'

3.2.2 Quantifier float

A quantifier, such as *krina* 'all' in the following examples, will always be understood to belong with the P argument of a *Ø-/i-* construction. Thus in (21) whatever the arrangement of the constituents, *kərina* modifies the undergoer, not the actor argument *-na* of *ihamati* 'respect'.

---

10 Woollams (1996:293) claims this only occurs if there is not an equivalent *N-* verb to satisfy the requirements of an actor subject. However the example he gives uses a verb that does have an *N-* equivalent.

(21) 773

jəlma kərina i-hamati-na (*/ i-hamati-na kərina jəlma/ jəlma i-hamati-na kərina*)
people all *i*-respect-he
'He (the king) respected everyone.'
* 'Everyone respected him.'
(but for this gloss see example (34) below, the *N-* form equivalent)

3.2.3 Imperatives and reflexivisation

The *Ø-* construction is the basic form used for both imperatives and reflexivisation. Examples (22) (an elicited example) and (23) are reflexive constructions, in which the actor is always the antecedent of the reflexive anaphor. In (23) the anaphor is preceded by an optional dative pronoun. It should be noted, as Woollams (1996:205) has observed, that the reflexive anaphor agrees in person with its antecedent, unlike that of a *N-* form reflexive anaphor which has an invariant form (see §4.1).

(22) *Ku-pəkpək ba-ngku.*
I-*Ø*-hit self-my
'I hit myself.'

(23) *Si-tandai-lah rusur man ba-nta.*
we-*Ø*-know-EMPH often DAT self-our
'We often indeed know ourselves.'

In imperatives the addressee, (the actor/agent), is freely omitted, as in example (24), unlike the situation in Tagalog, for example. The omission of the A argument in imperatives, and its control of agreement in reflexives, indicate its status as a core argument in these constructions.

(24) 122

Pəsai ləbe dukut-na, agoi ləbe dukut-na.
Ø-clean first grass-DEF // *Ø*-remove first grass-DEF
'First clean up the grass, remove it...'

3.3 Ordering constraints

The basic predicate-initial order of Karo is reflected in the preferred **V A P** order of the *Ø-* constructions. Full nominal A arguments, like pronouns, must immediately follow the verb (with the exception of the first person prefixes). Most of the examples above represent the basic order, for example (15).

Unlike the cliticised A argument, the position of the subject, P, is extremely flexible. It frequently precedes the verb, as in (25). The use of the particle *pe* 'also' here is optional; its function is purely semantic. Nevertheless a particle, whether *pe* or another form of emphatic particle, marked an inverted subject in most of the examples I found of these constructions; that is, the pragmatic and semantic significance of the clause initial position often attracts extra marking. In (26), however, the pronominal undergoer preceding the verb is unmarked.

(25) 189
Sen *pe galang-galang bərekən-na man ba-ngku.*
money *pe* big-big *Ø*-give-he DAT self-my
'He also gave a lot of money to me.'

(26) *ənggo təlu kali* ***kam*** *ku-idah landək.*
PERF three times you I-*Ø*-see dance
'I've seen you dance three times.'

3.4 Obligatoriness of arguments

Neither argument of the *Ø*- constructions, whether in main or subordinate clauses, is obligatory in spoken discourse, and once either A or P is an established entity it may be omitted in contexts where its reference will be understood. In (27b), which was uttered almost immediately after (27a), the undergoer subject has been omitted.

(27)a. 635
Ku-idah ia rusur ku rumah ənda.
I-*Ø*-see her often at house this
'I see her often at this house.'

b. 637
Ku-idah rusur // aku usur lewat arənda.
I-*Ø*-see often // I often pass here
'I often see [her], I often pass here.'

In (28) and (29) the non-subject (A) argument of the *Ø*- construction main clauses has been omitted. Example (29) could mean 'we'll sell the land', the most likely interpretation in the context, or it could be an imperative 'sell the land'.

(28) 735
Tukur ndai bəras-na // tukur ndai ikan-na.
Ø-buy PTL rice-their // *Ø*-buy PTL fish-their
'(I) just bought (them) their rice and fish.'

(29) 20
Emaka dayakən juma e muat ənəm puluh juta.
so *Ø*-sell land DEM *N*-get six ten million
'So (I'll/we'll) sell the land to get sixty million.'

Apart from the zero anaphora of these examples, various particles may serve as anaphors.

Omission of arguments from subordinate clauses is discussed in §5.3.2.

3.5 Summary of *Ø*- verb forms

The *Ø*- forms, which have a morphologically ergative pattern of pronominal affixation, are the most frequent type in main clauses. Overall, the conclusion is that both A and P are core arguments of *Ø*- constructions, with P, the subject, ranking higher than A. The P of the

Ø- constructions (and the single argument of intransitives, S) can be relativised through a gap strategy, the primary strategy of relativisation in Karo. The non-subject A argument can also be relativised, if required, through a secondary strategy of pronominal retention in the relative clause. The A argument of the *Ø-* constructions also controls the reflexive anaphor, which agrees in person and number with it. In straightforward two-argument transitive clauses with a quantifier, the quantifier will be understood to modify the subject, the P argument, regardless of its position in the clause. Both A and P can be deleted, and frequently are in conversational discourse, and both can function as controllers or controllees in subordination with appropriate verbs. Finally, the *Ø-* constructions are the canonical form of imperatives.

4 The *N-* verb forms

4.1 Formal properties

The pronominal forms that the *N-* forms take are the free form (FF) pronouns seen in Table 1. Full NPs are unmarked (apart from the marking with *si* for proper names). P arguments, both pronouns and full NPs, usually follow the verb, although there are exceptions, including those constructions where the suffix *-sa* is present on the verb. An example of an *N-* construction with two pronominal arguments is (2) above, repeated here as (30). It is worth noting that, due to the distribution and syntactic conditions under which the *N-* form occurs, this example is one of the very few from the whole corpus that illustrate an *N-* form construction with two pronominal arguments. (31) shows both pronominal and full NP arguments and (32) full NP arguments only.

(30) 309
adi aku səkali n-jumpai ia
if I once *N*-meet him
'if I ever meet him.'

(31) 762
pərbahan aku muat uang sekolah
...because I *N*-get money school
'...because I have to earn money for school.'

(32) 446
səbab pərtambar-ku pe ənggo n-jutai ənda utang
because medicine-my *pe* PERF *N*-throw DEM money
'because my medicine has already wasted the money'

4.2 Grammatical relations

The grammatical subject of the *N*-form constructions is the A(ctor). It passes various tests for subjecthood as discussed below. On the other hand the undergoer, P, of these constructions passes none of the following tests for core status, except that under certain circumstances it may be omitted, and that it is able to control deletion in a lower clause, as shown in §5.3.1.

4.2.1 Relativisation

The actor argument of the *N-* verbs is readily relativised using the normal gap strategy.

(33) 296

Lang	*aku*	*la*	*kah*	*min*	*aku*	*si*	[ ]	***nukur e***	*ya.*
not	I	not	EMPH	PTL	I	REL	[ ]	*N*-buy	DEM EMPH

'I am not the one who is buying it.'

But there is no strategy, including that of pronominal retention used for various other arguments, whereby the P of the *N-* form clauses can be relativised.

4.2.2 Quantifier float

With the use of an *N-* form, the actor, *jəlma* 'person', of (21) above, (*jəlma kərina i-hamati-na* ('he respected them all')), can be made the actor subject of the clause and can be quantified in this role. The quantifier *kərina* may appear in several positions but it always pertains to the subject, *jəlma* 'people'.

(34)

Kərina	*jəlma*	*ngə-hamati*	*ia.*	(OR: *jəlma ngəhamati ia kərina* OR: *jəlma kerina ngəhamati ia*)
all	people	*N*-respect	him	

'They all respected him.'

Note that **jəlma ngəhamati kərina ia* is not possible because P, *ia*, no longer directly follows the verb.

4.2.3 Imperatives and reflexives

Both imperatives and reflexives can be constructed from *N-* forms, but they are not canonical. The *N-* form in example (35), *nungkun*, is one of the very few that appear to occur in imperative mood and, like the more common Ø-form imperatives, can omit the actor/agent as in (35).

(35) 353

Nungkun	*ləbe*	*man*	*pərmərintah*	*ənda.*
N-ask	first	DAT	government official	DEM

'You'd better ask the government first.'

Reflexive anaphors of *N-* constructions have an invariant third person form showing no person agreement with their antecedents. Example (36) is from Woollams (1996:205); there were no examples in the corpus.

(36)

Kam	*muji*	*ba-na*	*usur.*
you	*N*-praise	self-3	always

'You are always praising yourself.'

4.3 Ordering constraints

Examples (34) and (36) above illustrate the basic **A V P** word order of these constructions. The position of the actor subject is more flexible than that of the non-subject undergoer which in most instances directly follows the verb, unless the suffix *-sa* intervenes.

As example (37) shows, while the actor subject generally precedes the verb it need not come immediately before it. Here the optional particle *pe* follows the subject.

(37) 275

Kalak *e* *pe* *la* ***nggit*** *nukur* *bage saja-ng(ə).*
people DEM *pe* not *N*-want *N*-buy thus only-*ngə*
'People may not want to buy (it), that's all.'

Generally, undergoer NPs do not precede the verb. However inversion has been observed, as in (38), where the undergoer is itself a complement clause. Here the complement clause, containing the verb *nukur* 'buy', precedes the main verb. In this instance the particle *pe* is not optional; it serves to disambiguate the main clause subject NP *kalak* 'people' from the unexpressed undergoer of *nukur*.

(38) 276

Nukur ***pe*** *kalak* *e* *la* *nggit.*
N-buy PTL people DEM not *N*-want
'Indeed others won't want to buy (it).'

Overall the order of the two alternative voice forms can be compared to those of Classical Malay (Cumming 1991:107 and151ff.) and Toba Batak (Schachter 1984). They are:

undergoer voice *Ø-/i-* verb clauses: **V**(erb)~**A** **[P]**
actor voice *N-* verb clauses: **[A]** **V**(erb)~**P**

where the tilde (~) indicates, for the argument that immediately follows it, a fixed and uninterruptible position with respect to the verb and, conversely, the square brackets ([ ]) indicate relative freedom of position for the argument they enclose. These two orders give an impression of a mirror image, typical of many WMP languages. This does not hold true for Karo however. In the case of the *Ø-* forms (undergoer voice), the pronominal A arguments are bound forms while both actor and undergoer pronominal forms for the *N-* forms are independent pronouns. Only when the suffix *-sa* is present on the *N-* forms, and no concurrent undergoer is present, is there a real mirror image. However as *-sa* is not an argument, as will be shown in §5, the impression of formal equivalence is false for Karo.

4.4 Obligatoriness of arguments

Although it is generally present in main clauses, an overt actor (the subject of the *N-* forms) is optional, as the following example shows.

(39) 33

N-darami *nakan* *i* *Jakarta* *ni-na.*
N-seek rice in Jakarta word-his
'(He is) looking for food in Jakarta he said.'

The undergoer may also be omitted. In many such cases the undergoer is not recoverable and the verb appears to be intransitive. But when the suffix *-sa* is present on the verb the undergoer is always recoverable, often from an immediately adjacent clause, sometimes from earlier in the discourse episode. For example, in one chain of utterances in the text with the verb *nukur* 'buy', the undergoer (the garden that was for sale, the main topic of discussion) was omitted in several clauses in succession, such as examples (37) and (38). Preceding this

series of utterances however was a solitary one *adi banta juma e, payo katamu ena?* 'this is about us buying the garden, am I correct?' which clearly established the identity of the undergoer. According to a native speaker these examples would be more grammatically correct if *-sa* had been suffixed to the verb. In Norwood (2000) I have demonstrated that undergoers of the *-sa* suffixed verbs, whether occurring in the same clause, or recoverable from a previous clause, are salient discourse entities, being both identifiable and specific.

Nevertheless, as indicated, a number of *N-* forms had no recoverable undergoer, nor was the addition of *-sa* judged possible in these circumstances. The syntactic status of these undergoers is discussed further in §5.5.

Omission of arguments in subordinate clauses is discussed in more detail in §5.3

4.5 Comparison of *N-* verb forms and *Ø-* verb forms

Table 3 sums up some of the main features of the two types of constructions described above. In addition to the features outlined in the table, various tests show that while the two arguments of the *Ø-* constructions participate in a number of syntactic processes, only the A argument of the *N-* forms satisfies various tests for core status. The non-subject A argument of the *Ø-* constructions may, for example, be relativised or be the omitted target of control, but the non-subject P argument of the *N-* forms cannot. The function of the *N-* form suffix *-sa*, which is of relevance to the status of the *N-* form undergoers, is discussed below in §5.5.

5 Functional analysis of nasal verb forms

The analysis was designed to establish the morphosyntactic status of the *N-* forms, their distribution and their function. It has been shown above that only one of their arguments, the agent subject, fulfils many of the criteria for core status. Compared to the *Ø-* forms these *N-* forms show a number of features of lowered transitivity.

Table 3: The morphosyntactic differences between the two voices
(based on the most common type of main and subordinate clauses for each voice)

	Ø- verb form		*N-* verb form	
	Main clause	Sub. clause	Main clause	Sub. clause
Unmarked order:	VAP	VAP	AVP	[ ] VP
Subjects:				
Semantic role:	P	P	A	A
Obligatory	no	no	(yes)	omitted
Pronominal form:	FF	FF	FF	–
Full NP:	unmarked and free position	unmarked	unmarked	–
Non-subjects:				
Semantic role:	A	A	P	P
Obligatory	no	no	?*	?*
Pronominal form:	affix	affix	FF enclitic	FF enclitic
Full NP:	verb enclitic	verb enclitic	variable	variable

* omitted under certain syntactic or discourse conditions – see §2.2.4 above.

5.1 Distribution of the *N-* prefix

The most striking feature of the *N-* forms is their distribution. Out of 109 examples of *N-* forms analysed from the text two stems occurred in this form only (that is, there were no *Ø-* equivalents). These were two complement taking predicates (CTPs) with 30 tokens between them. Out of the remaining 79 *N-* form clauses, less than 20 occurred in main clauses, or approximately one in four. There was a similar distribution for the written data set. This section examines the distribution of these verbs, the morphosyntactic features that are associated with them, and their functions in both main and subordinate clauses.

Table 4 shows the distribution of the main prefixes on semantically transitive verb stems from the text. Stems with more than five tokens, including those formed with valency increasing applicatives, are listed individually. There is a range of semantic types, including verbs of perception such as *idah* 'see', and of cognition such as *ətəh* 'know', and three common complement taking predicates (CTPs) (18, 21 and 22).

Two nominal forms noted in footnote 6, *ate* 'wish/like, *nin* 'word/say', and a third one *ukur* 'thought/think' were not included, both because of the nominal morphology of their first person forms and because they usually lack voice alternatives, although *N-* and *ər-* forms of *ate* and *ukur* are very occasionally found. If they had been listed with the *Ø-* forms, the figures would have been even more heavily weighted towards these constructions, especially given the high frequency of these three items.

The unequal distribution of the two voice prefixes clearly demonstrates that the *N-* forms do not function as basic transitive forms. This is particularly evident in looking at the first three stems, where the ratio of *Ø-* to *N-* forms is 100:3. Surprisingly, two verbs, *tukur* 'buy' and *buat* 'get', showed a greater frequency of *N-* forms. These are discussed below.

5.1.1 Semantic and syntactic parameters

A wide variety of factors were considered in interpreting the distribution given in Table 4. These included, among others, the specific parameters of transitivity from Hopper and Thompson (1980) shown to be relevant to other WMP languages, viz: (i) the number of participants, (ii) aspect, (iii) affirmation, (iv) mood and (v) individuation of O (P).

Only those parameters that proved significant are listed and discussed here.

1. Presence or absence of overt arguments in the clause.
2. The type and form of A and of P, whether a bound or FF clitic pronoun (as per Table 1), a demonstrative, or a definite, specific or indefinite NP
3. The presence in the containing clause of modals, tense/aspect adverbial particles, and complementisers and conjunctions.
4. Type of clause, whether (a) main clause — declarative, interrogative, negative or imperative — or (b) complement or other subordinate clause, including relative clauses. When a subordinate clause was recorded, the type and form of the main clause predicate was noted, for example desiderative, permissive, purposive, resultative or causative verbs, and whether the CTP was a verb, adjective or nominal.
5. The particular arguments involved in deletion, controller and deleted controllee.

Table 4: The distribution of prefixes on semantically transitive verbs in a conversational text

STEM	GLOSS	*Ø-*	*N-*	*tər-/ər-/i-*	TOTAL
1. *daya(kən)*	'sell'	34	1	0-2-4	41
2. *(ə)təh*	'know'	40	1	0-0-0	41[11]
3. *idah*	'see'	23	1	1-0-0	25
4. *katakən*	'tell'	14	8	0-0-2	24
5. *tukur*	'buy'	6	13	0-0-1	20
6. *sungkun*	'ask'	15	3	0-0-0	18
7. *jumpai*	'meet'	15	4	0-0-0	19
8. *suan*	'plant'	8	8	0-0-1	17
9. *buat*	'get'	5	9	1-0-0	15
10. *təkən*	'sign'	7	7	0-0-0	14
11. *dungi*	'finish'	6	3	2-0-0	11
12. *darami*	'seek'	5	2	0-0-0	7
13. *pindo*	'ask'	5	2	0-0-0	7
14. *sampati*	'help'	4	1	1-0-1	7
15. *suruh*	'ask for'	6	0	0-0-0	6
Sub-totals:		193	63	5-2-9	272
16. <5 tokens per stem*		72	16	not counted	88
Ø- FORMS ONLY:					
17. *ban*	'do'	80		0-2-2	84
18. *akap*	'feel/think'	29		0-0-0	29
19. *bərekən*	'give'	14		0-0-0	14
20. *dat*	'get'	8		0-0-0	8
N- FORMS ONLY:					
21. *nggit*	'like'		20	0-0-0	20
22. *ngasup*	'be able'		10	0-0-0	10
TOTALS:		393	109	5-4-11	525

* These verbs were not counted with computer assistance and the actual numbers may vary slightly.

5.2 Argument structure – parameters 1–2

In *Ø-* verb clauses P, the subject of these constructions, was found to be overwhelmingly identifiable, with definite marking. P was most often either a pronoun or a NP marked with a possessive clitic or demonstrative. The identifiablity or individuation of P accords with the findings of Hopper and Thompson regarding high discourse transitivity.

11 Eleven further tokens of *ətəh* as well as numerous forms of an allomorph [*dah*] occurred in second person forms – all variations of a discourse marker comparable to English 'you know'.

On the other hand the *N-* forms had almost equal numbers of identifiable and unidentifiable undergoers. This can be explained partly in terms of undergoer incorporation, which is discussed in §5.4.3.

The *Ø-* verb form clauses were five times more likely to contain two overt arguments than the *N*-form constructions. Given that the majority of *N-* forms were in reduced complement clauses (where deletion of the A subject coreferential with a main clause argument is obligatory), these findings are not surprising. Typical examples of each type with the verb stem *dayaken* 'sell' are (40) which has a *Ø-* form and two overt arguments (in both of its clauses), and (41), with the *N-* form which has a deleted subject argument coreferential with a main clause argument.

(40) 5

Juma e adi ənggo pagi dayakən-ndu,
garden DEM if PERF tomorrow Ø-sell-you

*kuja ngə ate-ndu **ban-ndu** **guna-na** **sen** ənda?*
where *ngə* wish-your Ø-make-you use-DEF money DEM
'If you sell the garden later on, what use will you make of the money?'

(41) 3

Mama // lit ate-ndu sura-sura-ndu
Uncle // exist wish-your wish-your

*guna n-dayakən **juma-ta** ənda?*
PURP *N*-sell land-our DEM
'Uncle, it is your intention to sell this piece of our land?'

Table 5 shows the results of parameters 1–2 above for a subset of the verbs listed in Table 4. The subset comprises only the first fifteen stems from Table 4. The results are divided to show the differences between clause types. Subordinate clauses themselves vary in type and deletion of subjects is not obligatory in all of them. Arguments that had zero expression are not included.

Table 5

	Ø- forms (No.=193)		*N-* forms (No.=63)	
	Main clause	Sub. clause	Main clause	Sub. clause
Overt A	106	29	9	9
Type of A:				
Identifiable NP	3	1	1	–
Non-identifiable NP	2	–	1	1
Pronominal	101	28	7	8
Overt P	93	24	6	28
Type of P:				
Identifiable NP	38	19	1	4
Non-identifiable NP	10	3	3	21
Pronominal	25	1	1	3
Clause	26	1	1	–
Enclitic *-sa*	–	–	14	11

5.3 Subordinate clauses

Analysis of parameters 3 and 4 showed that two in three *N-* form constructions in the text were subordinate clauses. A more casual analysis of semantically transitive *N-* forms in the written genres produced similar results. These findings tally with those of Woollams (1996:212), as do the relative proportions of *N-* versus *Ø-* forms in Table 2.

In a separate study, the *N-* form constructions were examined in the context of several inter-related syntactic and semantic parameters of subordination described by Lehmann (1988) and Hopper and Thompson (1980). These include the syntactic downgrading and desententialisation of the subordinate clause, as well as an increasing tendency to omit any explicit linking device as desententialisation increases. The majority of *N-* form clauses showed a high frequency of these features.

Some observations on the role of both the *Ø-* and the *N-* forms in subordination are recorded here, particularly in regard to control phenomena. Athough relative clauses were discussed above, the descriptions of control and controllee relations applies equally to them as to other kinds of subordinate clause.

5.3.1 N- *verb form control relations*

When the main clause is an *N*-construction, the actor subject may control deletion in a subordinate clause. Due to the scarcity of *N-* form main clauses in general this is uncommon, with the exception of the two *N-* form CTPs (nos. 21 and 22 in Table 4), one of which, *nggit* 'want', is exemplified in (42), the other being *ngasup* 'be able'. Square brackets represent the deleted argument.

(42) 295
Kalakai pe la ng-git [ ] *nukur adi la bəre Nini.*
people *pe* not *N*-want *N*-buy if not agree Nini
'People won't want to buy if Nini doesn't agree.'

Control of deletion by the undergoer of an *N-* form main clause also appears to be possible. There were no instances of this across a wide range of data however, and (43) is an elicited example.

(43) *Aku ng-idah ia sangana* [ ] *nuan-i galuh.*
I *N*-see him while *N*-plant-TRS banana
'I watched him while he planted bananas.'

With regard to the subordinate clause itself, when it is an *N*-form reduced complement clause, the actor subject must be deleted, as in the examples just shown and others scattered throughout the discussion. But in subordinate clauses which have a degree of independence, such as those introduced by a complementiser, as in (44), omission is not possible.

(44) 309
Adi aku səkali n-jumpai ia.
if I once *N*-meet him
'If I ever run into him.'

Except under the particular discourse conditions described in §4.4, or when the suffix *-sa* is present on the verb, the undergoer of an *N-* clause is not deleted. In general therefore, it may

be concluded that only the actor subjects of the *N-* constructions may be the target of controlled omission in subordinate clauses.

5.3.2 Ø- verb form control relations

Most often in cases of subordination, especially those involving *N*-construction reduced complements and deleted actor pivots,[12] the main clause is a *Ø-* construction. Normally its actor controller is overtly expressed. In (45) this is *ku-*.

(45) 252

***Ku**-dayakən nin-gku // sitik kəl kap e // [ ] muat dua juta.*
I-Ø-sell word-my // small very EMPH DEM // *N*-get two million
'I'll sell I tell you, just a small piece, to get two million.'

But example (29), repeated here as (46), demonstrates that, even when absent, the *Ø-* construction actor can control deletion from the subordinate clause.

(46) 20

Emaka dayakən juma e [ ] muat ənəm puluh juta.
so Ø-sell land DEM *N*-get six ten million
'So (we'll) sell the land to get sixty million.'

The *Ø-* construction subject, P, also controls deletion with a number of verbs, including verbs of speaking and cognition such as *suruh* 'tell/order' in (47) or *sangka* 'expect'.[13]

(47) 577

Ku-suruh ma-nguda ləbe [ ] ngərana ras Joko.
I-Ø-tell uncle-young first speak with Joko
'I told younger uncle to speak with Joko first.'

When the subordinate clause itself is a *Ø-* construction, as in the second clause of (48), its P subject may be deleted. In this sentence *ia* 'he' is both intransitive subject of the main clause and deleted undergoer of the *Ø-* construction complement.

(48) 485

Adi kari ia ng-git ndo ku-jumpai [ ] səkali nari ni-na.
if later he *N*-want allow I-meet-TRS once again word-his
'Later he would be willing if I met [him] another time, he said.'

Non-subject As in *Ø-* construction subordinate clauses are not normally deleted, but may be under some circumstances. Examples (49) (a) and (b) show a minimal pair using the same *Ø-* form CTP (*bere*) and the same embedded verb (*dayaken*) in the second clause. In the (a) example the A (*anak* 'child'/'he') is retained in the second clause; in the (b) example it has been deleted (as in fact has the understood undergoer in both utterances).

---

12 Dixon (1994:17) points out that ergative languages always have a syntactic operation which feeds the need for an underlying A to function as a pivot.

13 Woollams (1996:313-314) provides an excellent and detailed discussion on the possibilities of a raising analysis for these verbs.

(49)a. 415

*La pagi bəre-**na** dayakən **anak** // pərutangkən.*
not tomorrow Ø-allow-he Ø-sell child // mortgage
'If he doesn't agree to sell, mortgage (it).'

b. 320

*Adi bage (ə)nda la bəre-**na** dayakən [ ] // məsui kal ate-ku.*
if thus DEM not Ø-allow-he Ø-sell // sick very feeling-my
'If he doesn't agree to sell, I'll feel quite ill.'

The most common CTP in the text is the nominal *ate* 'desire/wish'. It is generally, though not invariably, followed by a Ø- form complement. When this CTP is used, omission of core arguments from the complement clause with either Ø- or *N*- verb forms did not occur.

5.3.3 Summary

This brief discussion of the phenomena of control and deletion of arguments in complex constructions, while not complete, serves to illustrate the fact that either argument of both the *N*- and the Ø- constructions may be controller of deletion. However only the Ø- forms allow either of its arguments to be deleted controllees. That is, Ø- construction subjects (P) and non-subjects (A) may be deleted controllees, and, as well, *N*- construction subjects (A). But the undergoer of the *N*- construction may not be deleted as the controllee. Furthermore the form of the CTP is a significant factor in determining the following clause type and the possibility of deletion.

Table 6 sums up the possibilities for controller and controllee relations.

Table 6

	Controller	Controllee
Ø-verb forms		
Subject	yes	yes
Non-subject	yes	yes
Dative	yes	yes
***N*- verb forms**		
Subject	yes	yes
Non-subject	yes	no *
Dative	?	?

* except when the suffix *-sa* is on the *N*- form in the subordinate clause.

5.4 The *N*- verb forms in main clauses

Out of 79 semantically transitive *N*- verb forms in the text (excluding numbers 21 and 22 in Table 4, i.e. the *N*- forms which were complement taking predicates) only twenty-two were in main clauses. This section describes the factors that condition the use of this relatively small number of *N*- form main clause constructions, i.e. irrealis mood (§5.4.1), topicalisation (§5.4.2), and undergoer incorporation (§5.4.3).

5.4.1 Irrealis mood

Anderson (1985:193-194) points out that, cross-linguistically, subordinate clauses often use the subjunctive form of the verb, this same 'subjunctive' form appearing in main clauses to mark functions such as irrealis mood. While the *N-* forms of Karo, so commonly found in subordinate constructions, are not the main form of marking for irrealis mood, in main clauses they are typically found in an irrealis context.

Realis and irrealis are defined, following Barr (cited in Himmelmann 1996:123)[14] as:

> REALIS: Past, completed action, a state or action already existing or occurring, a characteristic which is real, existing, fact, fully actualised.
>
> IRREALIS: Non-past action, hypothetical, not yet realised action or state, a characteristic not yet real, not fully actualised. In this respect irrealis shares some features of subjunctive.

A feature of those *N*-forms found in the main clause subjunctive/irrealis context, as well as in subordinate clauses, was a lack of associated temporal marking. While tense/aspect adverbial particles were commonly found in *Ø-* form clauses, there were few in any of the *N-* clauses in the text. Examples (50) and (51) exemplify irrealis constructions.

(50) 535

Ei *pensiun* *bapa* *e* *nekolahken* *adek.*
AFFIRM pension father DEM *N*-school-TR sister
'It's Dad's pension that is schooling (paying school fees for) sis.'

(51) 632

Aku ***n-dungi-sa*** *ras* *ia.*
I *N*-finish-*sa* with her
'I'll sort it out with her.' (* 'I sorted it out with her.')

Although main clauses with *N-* forms may be used for irrealis mood it does not follow that irrealis is generally expressed by the use of the *N-* form. Most irrealis clauses are either future, marked by the future tense particle *kari,* or conditional, marked by *adi* 'if'. While either particle can occur with an *N-* form, they most frequently occur with *Ø-* verb forms. Additionally, as noted above, transitive imperatives, another class of irrealis, are always *Ø-* forms.

It was noted at the beginning of this section (5) that two stems showed disproportionate numbers of *N-* forms. These were the verbs *tukur* 'buy' and *buat* 'get'. That these two verbs appear most often in this form makes sense in the context, since the most uncertain elements of the whole endeavour are finding a buyer and getting the money. This does not totally explain why *dayaken* 'sell' has only one *N-* form, as the selling is also uncertain. But there is no uncertainty in the intention to sell and I presume that this is where the difference lies.

In a focus system, such as the alternative voice forms of Karo are based on, Hopper and Thompson's 1980 transitivity hypothesis predict that the *N-* forms would be used for imperfective and irrealis semantics, compatible with lowered transitivity. Except for the fact that the future marker is used with the *Ø-* forms, Karo conforms to their predictions regarding the association of irrealis marking and the use of the actor voice (*N-* forms).

---

14 Barr and Himmelman make the distinction in relation to Sulawesi languages.

5.4.2 Topicalisation

Two examples definitely did not fall into the category of irrealis mood. In (52) the speaker has just previously interrupted another speaker to correct a description of an event in which only the interrupting speaker had personally participated. With a particularly emphatic tone this speaker starts a long and uninterrupted passage, repeating the first person pronoun before proceeding.

(52) 325

Aku // aku ras bapa-ndu mbarənda ***n-jumpai-sa*** *ku pəngadilən ah.*
I // I and father-your past *N*-meet-*sa* at court DEM
'It was ME, me and your late father who met (her) at the court.'

This is a topicalised construction; the emphasis on the first person pronoun subject is shown by its repetition after the intonation break. The use of the *N-* form *njumpai* allows the first person pronoun to be the subject of the sentence and to appear in sentence initial position.

Another example occurred further on in the same emphatic utterance:

(53) 327

Kami n-jumpai-sa dua-dua ras katua mbarənda.
we *N*-meet-*sa* together with older.sister late
'We met together with (our) late sister.'

5.4.3 Undergoer incorporation

In the text a very high proportion of the *N-* forms were associated with just three stems. One of these, *nuan* (stem = *suan* 'plant'), had four tokens in one sentence with three further tokens in quick succession. Native speakers appear to regard these forms, consisting of *N-* form +indefinite P, as a form of nominalisation, indicated by the gloss in (54).

(54) 105

Ku juma dahin-ku rusur ***nuan*** *wartel //*
in garden job-my often *N*-plant carrots //

nuan *kentang //* ***nuan*** *lacina əntah pe* ***nuan*** *markisah.*
N-plant potatoes // *N*-plant chilli or *pe N*-plant passion-fruit
'Frequently my job in the garden is carrot-planting, potato-planting, chilli-planting or passionfruit-planting.'

In describing these constructions Woollams (1996:179) points out that between the verb and its undergoer there can be no intervening material, not even a phonological pause; nor can the undergoer be modified in any way. This has the effect of creating an intransitive verb wherein the undergoer is incorporated into the verb. This is consistent with the proposal of Starosta, Pawley and Reid (1982) that the focus affixes were once nominalising affixes which became reanalysed as verbal focus affixes.

Table 7: Summarises the distribution of main clause functions of the *N-* forms

1. Irrealis mood	16
2. Topicalisation	2
3. Undergoer Incorporation	4 (plus 4 in subordinate clauses)
Total	22

5.5 The suffix *-sa*

As briefly discussed in §2.4 the *N-* forms may take a suffix of the form *-sa* (or *-ca*). Although this suffix cannot be used in all contexts with *N-* forms, in many instances where it is absent the judgement of a native speaker is that it could occur, and that its presence would make the construction grammatically more 'complete'. It has been claimed that this suffix is an allomorph of the independent form of the third person pronoun, *ia*, for the object[15] or undergoer of the *N-* forms (e.g. Woollams 1996:115). Example (12) above, repeated here as (55) is an example where an analysis of third person undergoer is apparently unproblematic.

(55) *Maka ngadi mə ia ng-ajar-sa.*
then stop *mə* she *N*-learn-*sa*
'Then she stopped teaching (?them).'

However, in addition to being found in constructions without an overt undergoer NP, such as (55), it may also co-occur with the usual free form third person pronominal undergoer *ia*, or with a simple third person NP in the same clause, as in example (56). In these cases it appears to cross-reference the undergoer NP. This creates problems for the third person analysis, but they are not serious ones. One could simply say that *-sa* is a third person object marker which can be used either to represent or to cross-reference the object.

(56) 638
Ia kang ng-gorəng-sa kuning-ndu?
she INTER *N*-cook-*sa* yellow-your
'Did she cook up your yellow (medicine)?'

However, *-sa* also occurs with non-third person pronouns, as in (57), an example from Woollams (1996:219; his example 6.195). This presents much more serious problems for an analysis of *-sa* as a third person object marker. To retain this analysis, one would have to say that it was an object marker that was not restricted as to person, and also that it could be used either to represent or to cross-reference the object.

(57) (Woollams 1996:219, e.g. 6.195)
Pitu wari nari rəh kami mərang-sa kam kərina.
seven day from come we *N*-attack-*sa* you all
'In seven days we shall come and attack you.'

Woollams interprets these examples of *-sa* as a perfective marker, an analysis that is incompatible with his own (and my) analysis of the *N-* form constructions in main clauses as imperfective. Additional problems for a third person object analysis arise from the fact that whether the undergoer is overt or absent, replacing *-sa* with a free form pronoun causes the construction in question to undergo semantic and syntactic changes, as (58a and b) show.

(58)a. 325
Dibata lah ng-idah-sa si bage e.
God EMPH *N*-see-*sa* REL thus DEM
'May God witness this, that he is thus (doing).'

---

15 The term object, a syntactic designation, is used with reference to the alternative theories being discussed. It is interchangeable with the semantic term undergoer.

b. *Dibata lah ng-idah ia ərbahan si bage.*
God EMPH *N*-see him do REL thus
'Indeed God is watching him do this.'

In (58a) *-sa* apparently refers to the event summed up in the following headless RC *si bage e.* If a free third person pronoun, *ia*, is used in this clause, as in (58b), this leads to changes in its syntactic structure and meaning such that the free form *ia* functions as an undergoer, followed by a complement clause; *ia* cannot cross-reference a following NP (such as the headless relative in (58a)).

Even if these syntactic differences in the use of the two forms are not taken into account, the presence of the suffix *-sa* with an undergoer which is either first or second person is sufficient evidence to refute the claim that *-sa* is an alternative form of third person undergoer, either as an argument in itself, or as a third person cross-reference form. The discourse properties of the undergoer that *-sa* represents or cross-references do, however, correlate with the presence or absence of this suffix. The undergoer of an *N*-verb suffixed with *-sa*, whether co-occurring in the same clause, or mentioned in a previous clause, is always an NP that is specific and identifiable; that is, individuated. It may be a personal pronoun, a lexical noun marked with a definite marker of some description, a proper name or a specific clausal undergoer. In every case it marks specific rather than general properties of the NP, in terms of the flow of information. It may not therefore be coincidental that the particular form of the suffix *-sa* is the same as the prepositional oblique case marker in certain Philippines languages, where its major function, apart from locative marking, is to mark specificity of the following NP (D. Zorc, pers. comm.). It seems plausible that it was once a prepositional marker of the undergoer NP that immediately follows the *N*- verb, and from this position was reanalysed as a verbal suffix. At the same time the overt mention of the NP itself immediately after the verb became redundant under certain conditions.

The suffix *-sa* also represents the third person pronominal agent of the stative passive *ter-* forms. While the contexts under which this occurs have not yet been investigated, it is possibly the case that these contexts are the same as those for its occurrence on the *N*- verbs, that is, it cross-references a specific and identifiable NP. This is a plausible explanation for the case of the passive agent in (59), who was identified by name in the previous clause.

(59) 226
La ter-dahi-sa ban-na ka tomat-na je sitik.
not *ter*-care for-*sa* *Ø*-do-he PTL tomato-his PTL little
'It can't be looked after by him, he does the tomatoes.'

The suffix *-sa* also occurs as an object of certain prepositions. Woollams (1996:116) gives examples of this function with *kempak* 'to', and *ras* (comitative) 'with' as in (60) (his example 4.75).

(60) *I-pindo-na gəlah banci ia tading ras-sa.*
i-ask-he PURP able he stay with-*sa*
'He asked if he might be able to stay with him.'

In both sets of circumstances, either as passive agent (or cross-reference to passive agent), and as the object of prepositions, *-sa* marks an oblique role. This gives support for the analysis of *-sa* as an erstwhile oblique preposition which, having lost its following NP under certain circumstances, has been reanalysed as a verbal or prepositional suffix which itself represents the NP.

5.6 Summary of section 5

The *N-* verb constructions function primarily to allow argument sharing between the main and subordinate clause where the actor is pivot. They are also lower in transitivity than their *Ø-* verb counterparts. The syntactic status of their undergoers reflects this lowered transitivity; they are generally unaffected, unindividuated undergoers, and not tracked in the discourse. However, certain constraints can motivate the use of an *N-* construction when the undergoer is specific and identifiable (as indicated by formal marking patterns, such as definite marking, pronominalisation or zero anaphora, or a proper name). In those cases, the suffix *-sa* may function as a cross reference form on the *N-* verb, though its presence is not obligatory. This use of *-sa* points to it being related to the oblique marking preposition found in Tagalog and other Philippines languages.

N- constructions are uncommon in main clauses, constituting only twenty percent of all uses of *N-* verbs. When they occur, they either indicate irrealis mode or are used for undergoer incorporating constructions or for topicalisation of the A argument.

It may be argued that as a widespread feature of the Austronesian family, the syntactic pivot function of the *N-* constructions, along with the vestiges of the focus system, are inherited features. The focus system as described by Starosta, Pawley and Reid (1982) wherein the verbal focus affixes originate from earlier Proto Austronesian (PAn) nominalisers provides information about the semantic role of the focused NP, and its pragmatic prominence.

If pragmatic prominence of the NP was the major motivation for the use of the *Ø-* construction (undergoer voice or patient focus) in Karo, then there would be no need for explanations. However, in Karo, this is not the case. There are a number of *Ø-* construction undergoers that are not topical or referential, and not marked with identifying markers. It seems clear therefore that while the *Ø-* constructions are higher in transitivity by a number of the parameters of transitivity described by Hopper and Thompson (1980), the *N-* constructions are lower, by these same parameters. This suggests that discourse transitivity has been a major factor in conditioning the use of either construction type (see also Wouk 1984).

6 Conclusion

This paper has shown that the Karo *N-* verb constructions cannot merely be regarded as alternative syntactically transitive constructions, as in Cirebon Javanese, for example (Ewing 1999), nor does it maintain the focus sytem as found in Philippine languages such as Tagalog, although its functions are clearly derived from an older focus system. Rather, Karo *N-* forms have several divergent functions. This, I suggest, is to a large extent a reflection of historical change.

The most common use of the *N-* forms is to manipulate the argument structure of subordinated constructions so that A may function as a syntactic pivot. This function is reconstructable for PAn as described by Starosta, Pawley and Reid (1982). According to Starosta et al. the focus affixes originate from earlier PAn nominalisers, and the focus system served to provide information about the semantic role of the pragmatically prominent NP in a clause. And indeed the coreferential pivot NP of the *N-* constructions in subordinate clauses still has pragmatic prominence.

Two infrequent functions, topicalisation and object incorporation, are also reconstructable for PAn. Topicalisation is a main clause function of the *N-* forms, and one does not necessarily entail detransitivisation. The topicalised A is frequently associated with *pe*, one of three particular emphatic particles (*pe*, *ngə*, and *mə*), which may be vestiges of an older noun phrase marking system associated with the verbal focus system. The nominalising function of undergoer incorporation, which is found in either main or subordinate clauses, does involve detransitivisation. Both these functions can be seen to reflect pragmatic prominence, as topicalised As are highly pragmatically prominent, and As of incorporated undergoers are necessarily more prominent than their undergoers.

However, in Karo, as in many other Western Austronesian languages, the focus system (or the remnants of it) no longer serves chiefly to mark pragmatic prominence. Rather, it is a marker of discourse transitivity as described by Hopper and Thompson (1980), with *N*-forms associated with low discourse transitivity and *Ø-* forms associated with high discourse transitivity (see also Wouk 1984). *N-* forms were found, with respect to several semantic and syntactic parameters, to be low in transitivity. In main clauses they are a marker of irrealis mood, although periphrastic adverbs in *Ø-* constructions are more commonly used for this purpose. A further aspect of the lowered transitivity manifested by the *N-* constructions is seen in the non-individuation of the undergoers of an unaffixed *N-* verb. The affixation of the suffix *-sa* on the verb marks their undergoer as an individuated entity. This suffix, analysed here as an earlier prepositional case marker, appears to be similar to the obliquely marked arguments of certain Philippines languages such as Tagalog.

The findings of this paper, then, can be seen to provide support for the interesting and challenging hypothesis put forward by Nothofer (1991), described at the beginning of this paper, that Karo and its near relatives may be part part of an older Palaeo-Hesperonesion group, related to other outlying western Austronesian languages.

References

Anderson, Stephen, 1985, Inflectional morphology. In T. Shopen, ed. *Language typology and syntactic description* 3: *Grammatical categories and the lexicon*, 150-201. Cambridge: Cambridge University Press.

Berg, René van den, 1996, The demise of focus and the spread of conjugated verbs in Sulawesi. In H. Steinhauer, ed., 1996:89-114.

Cumming, Susanna, 1991, *Functional change: the case of Malay constituent order*. Berlin: Walter de Gruyter.

Dixon, R.M.W., 1994, *Ergativity*. Cambridge: Cambridge University Press

Ewing, M., 1999, The clause in Cirebon Javanese conversation. PhD dissertation, University of California, Santa Barbara.

Fox, B. and P.J. Hopper, eds, 1994, *Voice: form and function*. Amsterdam/Philadephia: John Benjamins.

Himmelmann, N., 1996, Person marking and grammatical relations in Sulawesi. In H. Steinhauer, ed., 1996:115-136.

Hopper, P. and S.Thompson, 1980, Transitivity in grammar and discourse. *Language* 56:251-99.

Padia si Ndekah (*si Nipegendek*), The Old Testament (abbreviated version), 1953, tr. by J.H. Neumann. Djakarta: Lembaga Alkitab Indonesia.

Padia si Badia ras Ende-Enden (The Old and New Testaments and Hymns), 1977, Jakarta: Lembaga Alkitab Indonesia

Lehmann, C., 1988, Towards a typology in clause linkage. In John Haiman and Sandra Thompson, eds *Clause combining in grammar and discourse*, 181-225. Amsterdam: John Benjamins.

Norwood, C., 2000, The Kato Batak nasal verb suffix *-sa*: third person pronoun or transitiviser? *Melbourne Papers in Linguistics and Applied Linguistics,* Vol 0 Issue 1. 69-82.

Nothofer, Bernd, 1991, The relationship between the languages of the Barrier Islands and the Sulawesi–Philippines languages. In Tom Dutton and Darrell Tryon, eds, 1992, *Language Contact and Change in the Austronesian world*, 389-409. Berlin: Mouton de Gruyter

Ross, Malcolm, 1995, Some current issues in Austronesian linguistics. In D.T. Tryon, ed. *Comparative Austronesian dictionary* 1:45-120. Berlin: Mouton de Gruyter

Schachter, Paul, 1984, Semantic-role-based syntax in Toba Batak. In Paul Schachter, ed. *Studies in the Structure of Toba Batak,* 122-149. UCLA Occasional Papers in Linguistics, 5.

Starosta, Stanley, Andrew K. Pawley and Laurence Reid, 1982, The evolution of focus in Austronesian. In Amrin Halim, Lois Carrington and S.A.Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, 2:145-170. Canberra: Pacific Linguistics.

Steinhauer, Hein, ed., 1996, *Papers in Austronesian Linguistics No.3.* Canberra: Pacific Linguistics.

Woollams, G. 1996, *A grammar of Karo Batak*. Canberra: Pacific Linguistics.

Wouk, F., 1984, Scalar Transitivity and Trigger choice in Toba Batak. In Paul Schachter, ed. *Studies in the Structure of Toba Batak,* 195-219. UCLA Occasional papers in Linguistics, 5.

Yusmaniar, Lubis A. Hamid, S. Simorangkir and B. Surbakti, 1987, *Struktur Bahasa Karo.* Pusut Pembinaan dan Pengembangan Departemen Pendidikan dan Kebudayaan. Jakarta