Obligatory Incorporation and ‘Have’ in Tukang Besi

MARK DONOHOE
Monash University

1. The paraphrasability of Noun Incorporation structures

Mithun (1984), in an excellent cross-linguistic survey of Noun Incorporation (NI) phenomena, asserts that NI as a morphosyntactic device is only found in a language when there is an alternative, non-incorporating structure in the same language. She remarks that

Interestingly, all languages which exhibit such morphological structures (noun incorporation – MD) also have syntactic paraphrases. If we know that, in Koryak, one can say tiqoyanmâtekin ‘I-reindeer-slaughter’, then we can correctly predict the existence of a sentence like Tinnètkin qoyâwg’e ‘I-slaughter reindeer.’ (1984:847-848)

This is not quite the same generalisation as that captured by Baker’s (1988) Uniformity of Theta-role Assignment Hypothesis (UTAH). Mithun qualifies her initial comment; while ‘productive morphological constructions of this type never exist in a language without syntactic analogs’ (1984:848), she also claims that ‘[a] wide range of productivity associated with individual N and V stems is a standard feature of incorporating languages’ (1984:889). While Mithun claims that if a language allows incorporation it will also allow at least some predicates to appear without incorporation, but does not stipulate that all verbs must allow a non-incorporated coding option, Baker’s claim is that if an argument shows alternation in its surface structure, then that reflects the same thematic role assigning conventions at deep structure. In Baker’s view, NI is taken to be a case of an object receiving thematic roles through normal transitive verb procedures, and then NI obscuring that pattern. Crucially, under the UTAH any NI structure implies the existence of a non-incorporating structure.

Depending on the definition of ‘incorporation’, there are some apparent counter-examples to this claim, counter-examples which are presented by Mithun and Baker, among others. Mithun (1984:897), quoting from Bogoras’ description of Chukehee (1922), (in her Section 6, ‘Decay’, discussing the fate of previously productive incorporation structures as languages develop) notes that in some languages there are

small sets of derivational suffixes which, when added to N’s, function much like incorporating V’s. They supply meanings such as ‘to fetch’, ‘to take off (clothes)’, ... Suffixed to N’s, they derive V stems denoting unitary activities, as in Koryak ...

pčai-tivdi
boot-take.off
‘He took off his boots.’
It is likely that these suffixes are simply V roots which, in the modern language, never occur without an [incorporated][Noun].

In the same section Mithun notes that 'relics of older NI processes have developed into productive systems of affixation.' (p. 891), and that 'a number of languages have affixes which correspond to incorporating V's in other languages.' (p. 887). These phenomena, while resembling noun incorporation, do not exhibit the same patterns that Mithun describes as being associated with NI. Massam's (2001) writings on Niuean 'pseudo noun incorporation' come closer to the mark in describing the construction that we find in Tukang Besi, an Austronesian language from central Indonesia (though see also Asudeh and Ball 2005 for an alternative, and perhaps more accurate, analysis that is more compatible with the Tukang Besi facts). While most bivalent verbs in Tukang Besi allow for both incorporated and non-incorporated structures, I shall show that there are some instances for which NI may not be paraphrased (short of lexical substitution) with a non-incorporation structure. Specifically, the verb hato- 'have', however, appears only in noun-incorporating structures, and has no corresponding non-incorporating paraphrase. As a prelude to this presentation I present a sketch of some of the relevant morphosyntax of Tukang Besi in Section 2, a discussion of more conventional NI in Section 3, and then the problematic cases in Section 4. A discussion of some possible reasons for this exceptional treatment are discussed in Section 5, along with a discussion of the place of this extreme form of (constructional) polysynthesis in a typology of incorporation types, as well as a quick review of other apparently obligatory incorporating verbs in other languages. A possible explanation of the unusual obligatory incorporation phenomenon is provided referring to the argument structure of the verb, specifically the unusual status that hato- has as a stative, yet bivalent, verb, a condition that is vanishingly rare in Tukang Besi.

2. Tukang Besi

Tukang Besi is an Austronesian language spoken natively on the Tukang Besi archipelago of Southeast Sulawesi, in central Indonesia, and in various trading communities between Singapore and New Guinea (Donohue 1999, 2004a).

It is a verb-initial, subject-final Philippine-type language syntactically. Morphologically it shows obligatory agreement for the S,A by verbal prefixes, and optional agreement for the P by enclitic. Nominal case marking follows a Philippine-style pattern (Schachter 1976, 1977 and many others both before and since), and I follow Bell (1976, 1983) and Kroeger (1993) in using the label NOM for 'nominative' to gloss the case that has variously been called the 'pivot', 'subject', 'focus' or 'topic' in these languages, and in Tukang Besi is marked with na. The other nominal cases are the genitive ma, the non-nominative core case te, here glossed CORE, and the oblique i / di, the former used in future/irrealis contexts, the latter in other cases. The agreement and case marking systems interact as shown in table 1, where s- and a- represent the obligatory prefixal agreement markers, and -p represents the optional enclitic agreement marker.

20 It could be argued that Baker's claim was that incorporation structures exist in contrast to unincorporated structures in a given language, but not necessarily for every verb in a language. This is not explicit in his writing (though it is a distinction that Mithun makes), but is a possible interpretation. It is certainly the interpretation that most closely fits the facts in Tukang Besi, as will be demonstrated in the sections following.
<table>
<thead>
<tr>
<th>Clause type</th>
<th>variant</th>
<th>Agreement</th>
<th>word order</th>
<th>case marking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivalent</td>
<td>P indexed on V</td>
<td>a-V=(p)</td>
<td>V A P ((\neg VPA))</td>
<td>te</td>
</tr>
<tr>
<td>P not indexed on V</td>
<td>a-V</td>
<td>V P A</td>
<td>na</td>
<td>te</td>
</tr>
<tr>
<td>Monovalent</td>
<td>s-V</td>
<td>V S</td>
<td></td>
<td>na</td>
</tr>
</tbody>
</table>

Table 1: Agreement, word order and case marking in Tukang Besi

The alternation between nominative and non-nominative Ps is shown in (1) and (2), along with the passive in (3), and a monovalent clause in (4) (the verb in (4) is unaccusative, but clauses with unergative verbs show the same morphosyntactic patterns). I present a phonetic transcription of the sentences as well as the usual three-line glosses in order to indicate stress placement (which is regularly trochaic), information that becomes important in the discussion of incorporation.\(^{21}\)

(1) *No-kili=mo* te lante (\(na \ ia\)).
\(3R\)-clean=PF CORE floor NOM 3SG
'She cleaned (a/the/some of the) floor.'
\(\text{[noki}li\text{imo te la}nte (\text{na \ ia})]\)

(2) *No-kili=\(e\)=mo* na lante (\(te \ ia\)).
\(3R\)-clean=3\(P\)=PF NOM floor CORE 3SG
'She cleaned the floor.'
\(\text{[noki}l\text{ii}t\text{emo na la}nte (\text{na \ ia})]\)

There are additional differences between the sentences in (1) and (2). The sentence in (2) shows a higher degree of prominence of the P (more specific, more definite, more individuated), and in most respects more 'transitivity' (in the sense employed by Hopper and Thompson 1980).

(3) *No-to-kili=mo* na lante (*te / di [etc.]\(ia\)).
\(3R\)-PASS-clean=PF NOM floor CORE OBL:PST 3SG
'The floor was/has been cleaned (*by her).'
\(\text{[notoki}lio na }\text{lante]}\)

(4) *No-moturu=mo* (\(na \ ia\)).
\(3R\)-sleep=PF NOM 3SG
'She fell asleep.'
\(\text{[nomotu}r\text{u}mo (\text{na \ ia})]\)

(the free pronouns *na ia* and *te ia* in (1), (2) and (4) are not strictly necessary, since enough information about the subject is encoded on the verb (Tukang Besi could be described as a 'Pro-drop' language), and their presence implies

\(^{21}\) The following abbreviations have been used in glossing examples: A, S and P follow Comrie (1978), designating syntactic roles; 1, 2 and 3 indicate first, second and third person, respectively, and SG and PL are used for singular and plural, along with R for reals (S,A prefixes appear in two sets, reals and irreals). The other abbreviations are: ANTICAUSS: anticausative, APPL: applicative, CAUS: causative, COM: comitative, CORE: core non-nominative, DAT: dative, FUT: future, GEN: genitive, INSTR: instrumental, NOM: nominative, OBL: oblique, PASS: passive, PF: perfective, PP: p-prefix, PST: past, R: realis, REC: reciprocal, S/A: inflix,
a particular pragmatic focus on the identity of the subject. See Donohue (1999) for discussion of the status of free and bound pronouns.

It could be argued that there is really only one basic bivalent clause type, the form shown in (2), and that (1), which shows no agreement for the P, is an ‘antipassive’ derivation from it. Indeed, similar analyses have been proposed for other Philippine-type languages. Problems with applying this analysis to Tukang Besi arise when we consider that we have to assume:

- that there is no explicit antipassive morphology on the verb, whereas there is explicit voice morphology in the language: when passive a verb shows overt marking, with the prefix to- (see example (3));
- that the ‘antipassive’, not the active, clause is used as input to the passive derivation, which cannot appear with P enclitics;
- that the by-phrase in the ‘antipassive’ is still a core argument of the verb, not an oblique (demonstrable by time-adverb placement restrictions, and case-marking choices);
- and that the derived ‘subject’ of an antipassive clause (the cleaner, te ia, in (2)) behaves the same with respect to relativisation as does the A of non-‘antipassive’ bivalent verb (see the discussion following (4)).

For these reasons I have adopted the analysis of Tukang Besi as showing a voice system which I describe as being of the ‘Philippine-type’, not adequately characterisable as either nominative-accusative with passives or ergative-absolutive with antipassives, but having two basic bivalent clause types (Donohue 1999:160-166, or Donohue 2004a; numerous other authors make similar claims for other western Austronesian languages, most famously Tagalog).

Two points must be noted about the morphosyntax of (1)-(4). The first is that the nominals (and pronominals) appear in a phrase that is preceded by a case marker (all nominal elements of a clause must appear with an overt case marker, whether referential or predicative). Case marking is obligatory on all nominals in Tukang Besi. The second point is that the aspectual clitic =mo ‘perfective’ appears at the end of the verb, following agreement marking but preceding the (case-marked) NPs. These are two simple but reliable tests for the absence of N1 in a clause. The behaviour of causes in causative constructions (of the Lai-Mandati dialect) is a reliable test for transitivity. Causativised versions of (1) and (4) are given as (5) and (6). Notice the oblique case marking on the cause in (5), which is built about a base-bivalent verb, but the use of the core case marker te on the causee in (6):

(5) Ku-pa-kili=mo te lante i ia.
    1SG-CAUS-clean=PF CORE floor OBL 3SG
    ‘I made her clean the floor.’

(6) Ku-pa-matur=mo te ia.
    1SG-CAUS-sleep=PF CORE 3SG
    ‘I put her to sleep.’

(this case-marking pattern in causatives formed with pa- is not found in the Rupu or Wanse dialects that I have reported on elsewhere (Donohue 1999 ).

In Rupu Tukang Besi the causee is expressed as the P of the causative verb. In this dialect, (5) appears as Kupakilimo te ta te lante, with identical case marking on both the causee and the causand (though with different syntactic behaviour). The requestive causative hape- appears with a different pattern again, in all dialects, with all causees appearing as obliques, as in: Kuhepekilimo te lante i ia, ‘I asked her to clean the floor (for me).’ and Kuhepetumoturumo i ia, ‘I asked her to go to sleep.’)
Even when they do not appear with an overt nominal or pronominal P, the causee of a putatively bivalent verb in a causative construction must still appear with an oblique case marker, and not the core one that would be expected if the verb was monovalent, as seen in (7) and (8).

(7) *Ku-pa-kili=mo ia.  
1SG-CAUS-clean=PF OBL 3SG  
'I made her clean it.'

(8) *Ku-pa-kili=mo te ia  
1SG-CAUS-clean=PF CORE 3SG

If the verb in (7) was treated morphosyntactically as monovalent, the case marking of the causee would be with te, as in (6). The sentence in (7) cannot be interpreted as 'I made her clean it', with a referential, but unexpressed P (which would be expressed by *Kupakili'emo i ia, with a third person P enclitic on the verb). Rather, it has at best a generic object, 'dirty things'. This is found with other verbs as well; manga 'eat', for instance, takes kaijawa 'cassava' as its assumed object when no other is given from its clause of from the immediate context. Nonetheless, it does not behave as a monovalent verb with a core case marker on the causee.

With this base on transitivity and aspect placement we can consider NI structures in general, and in Section 4 NI in clauses headed by the verb hoto- 'have'.

3. Noun Incorporation in Tukang Besi

Incorporation is a productive process in Tukang Besi, but is not as prevalent in texts and conversation as are the many valency-changing devices (applicatives, causatives, passives) in the language. Nonetheless, many verbs show an alternation of a fully cased noun appearing in a constituent apart from the verb, and a noun without overt case appearing in close constituency with the verb. Some examples of this construction are given in (9)-(11).22 Note particularly that the V and N form a single phonological unit: stress assignment in (9), for instance, indicates that there is a single prosodic phrase extending over nokilikamara, with regular trochaic stress on the penultimate syllable and secondary stress found at the beginning of each foot preceding it.

(9) *No-kili kamara na ia.  
3R-clean room NOM 3SG  
'She cleans rooms.'

[nokili ka'rama na 'ia], *[nokili ka'rama na 'ia]]

(10) To-ala kaluku di koranga.  
1PLR-fetch coconut OBL:PST garden  
'We fetch coconuts from the gardens.'

[toqala kaluku di ko'rama]

---

22 The kind of incorporation described here appears to involve the V and the N compounding in a V', and not the V0, level. See Mohanan (1995, 1997) for some recent treatments of similar constructions in Hindi. Note that the nominal unit that participates in incorporation in Tukang Besi is an N', not an NP as described by Massam (2001) for Niue.
(11) *Salala 'u-wini simbuku=mo ara i kolo ana.
always 2SG.R-pull.in.on.line octopus=PF if OBL bay this
'You would always pull in octopuses when we went fishing on this bay."
[sa'ala 'u-wini simbuku=kumo]are kolo=ana]

As can be seen in (11), aspect clitics may follow the verb and noun combination. The
incorporated noun is also notably lacking a case marker, and may not appear in a preverbal
topicised position. These ungrammatical possibilities are shown in the following sentences,
based on (10), along with the possibilities for aspect clitic placement. Note the difference in
interpretation of aspectual scope in (12)a and (12)b.

Aspect clitics: attach to V base, or to V+N unit

(12) a. To-ala kaluku=mo.
IPLR-fetch coconut=PF
'We would fetch coconuts.'
[toala kaluku]

b. To-ala=mo kaluku.
IPLLr-fetch=PF coconut
'We came to be fetching coconuts.'
[toala=mo kaluku]

No case marker on incorporated Ns

(13) a. To-ala kaluku.
IPLR-fetch coconut
'We fetch coconuts.'
[toala kaluku]

b. *To-ala te kaluku
IPLR-fetch CORE coconut
*[toala te kaluku]

Case marker required on non-incorporated NPs

c. To-ala te kaluku.
IPLR-fetch CORE coconut
'We fetch the/some coconuts.'
[toala te kaluku]

b. *To-ala kaluku
IPLR-fetch coconut
*[toala kaluku]

No topicalisation of bare Ns, acceptable with non-incorporated NPs

(14) a. *Kaluku to-ala.
coconut IPLR-fetch
'Coconuts, we would fetch.'
[kaluku toala]

b. Te kaluku, to-ala.
CORE coconut IPLR-fetch
'The coconuts, we fetched.'
[te kaluku toala]

Note that the variability in aspect clitic placement is not available with other suffixal
material on the verb. If we were to use P enclitics (see example (2)) rather than an
incorporating N in a structure similar to (12) we would find that the only possibility for clitic
placement is following the entire verb.

(15) a. To-ala=1e=mo.
IPLR-fetch=3P=PF
'We fetched them.'
[toala=1e=mo]

b. *To-ala=mo=1e.
IPLR-fetch=PF=3P
*[toala=mo=1e]

---

29 Frequently a core case marker marking a P following a verb (that is, VP-internal) will be
'phonologically incorporated' into the prosodic phrase that includes the verb. An alternative
realisation of (13) could thus be [toala=te kaluku].

184
This sort of incorporation exemplified in (9)-(11) is restricted to Ps of bivalent verbs; the S of an unaccusative verb may not be incorporated into the verb, as can be seen in (17), which is an ungrammatical NI structure based on (16).

(16) *No-kengku na uwe.
    3R-cold NOM water
    ‘The water is cold.’

(17) a. *No-kengku uwe
    3R-cold water

b. *No-uwe kengku
    3R-water cold

c. *Kengku uwe
    cold water

d. *Uwe kengku
    water cold

Similarly, from a trivalent verb such as hu’u ‘give’ incorporation is not possible, with either the recipient or the theme.

(18) No-hu’u te ana te ‘onitimu.
    3R-give CORE child CORE watermelon
    ‘They gave the child a watermelon.’
    [nohu’u te ana te ‘onitimu]

(19) *No-hu’u ana te ‘onitimu
    3R-give child CORE watermelon
    ‘They gave children watermelons.’

(20) *No-hu’u ‘onitimu te ana
    3R-give watermelon CORE child
    ‘They would give watermelons to children.’

Examining the transitivity test found in the case marking options for causes in causative constructions, we find that the clause with NI appears as a monovalent one, with the causee marked not with the oblique i, but with the core te:24

(21) Ku-pa-kili lante=mo te ia.
    1SG-CAUS-clean floor=PP CORE 3SG
    ‘I made her clean floors.’
    [ku-pakili lante=mo te ia]

(22) *Ku-pa-kili lante=mo i ia
    1SG-CAUS-clean floor=PP CORE 3SG

Another difference between verbs with NI and the verbs in the non-incorporating clauses shown in Section 1 is the ability of the verb to appear with P agreement clitics, regardless of the relative order of the P clitic and the incorporated N with respect to the verb. Compare the grammatical (2) from Section 1, with the ungrammatical attempts to maintain agreement with the incorporated N shown in (23) and (24).

24 Note that the sentence in (21) is minimally different from that in (5). The sentence in (5) shows the case-marking found in the Lia-Mandari dialect of Tukang Besi, while that of (21) shows the more widespread case marking pattern.
 Parenthetically, we should note this is a syntactic, not a morphological, test. It is possible for an P clitic to appear on a verb with NI, but only when that verb has already been causativised, and so appears with a causative prefix, and is thus a multiple-predicate trivalent construction, as seen in (25).

(25) *No-kili \ lante=`e=mo.
   3R-clean=3P floor=3P=PF
   'She cleaned (the) floor(s).'
   [ku:pa\kili\lante\=e\=mo]

In this example the P clitic indexes the causee ("her"), not "floor". With a second person causee, second person P clitics would be used, as in (26), making it clear that the clitic refers to a different argument, and is not indexing the incorporated N.

(26) Ku-pa-kili \ lante=ko=mo.
   1SG-CAUS-clean floor=2SG.P=PF
   'I made you clean floors.'
   [ku:pa\kili\lante\=ko\=mo]

Another property can be used to test the difference in structure between the clauses with NI and the ones without it is the ability of adverbs to appear between a verb and its case-marked P, as in (27), but not between a verb and a bare P (though it may appear following the P, as in (29), an option (not illustrated here) also available for non-NI clauses).

(27) No-kili \ marasai \ te \ 3R-clean difficult CORE floor NOM 3SG
   lante (na ia).
   'She cleaned the floor with difficulty.'

(28) *No-kili \ marasai \ lante \ (na ia).
   3R-clean difficult floor NOM 3SG
   'She cleaned the floor with difficulty.'

(29) No-kili \ lante \ marasai \ (na ia).
   3R-clean floor difficult NOM 3SG
   'She cleaned the floor with difficulty.'

We can bring together the differences between non-NI and NI clauses as revealed in this different tests with the summary in (30).
Having established this battery of tests for NI, we shall examine a construction involving the verb *hoto- ‘have’* and its lack of alternation with a non-NI form.

4. Noun Incorporation and *hoto- ‘have’*

The ownership of alienable items, and of distant (defined as not consanguinal) kin terms can be expressed with the verb *hoto- ‘have/possess/own’.\(^{29}\) When *hoto-* is not used to indicate possession, either because of semantic restrictions as just mentioned, or speaker preference, a simple existential clause with possessive marking is used. The contrast between the two constructions can be seen in (31)-(34), showing alienable and inalienable nouns in each of the two constructions.

Alienable noun

(31) **Ku-hoto kabali.**
1SG-own machete
‘I have a machete.’
[kuho:to:ka?bal]  

(32) **Ane ke kabali=su.**
exist COM machete=1SG.GEN
‘I have a machete’
(Glossing literally, ‘My machete exists.’)

The inability of the *hoto-* construction to be used to mark the possession of inalienable kin terms is shown in (33).

Inalienable noun

(33) * **Ku-hoto iai (=su)**
1SG-own younger.sibling=1SG.GEN
‘I have a little sister.’

(34) **Ane ke iai=su.**
exist COM younger.sibling=1SG.GEN
‘I have a little sister.’
(Glossing literally, ‘My little sister exists.’)

Although (31) and (32) serve the same pragmatic purpose, they are significantly different in entailment. Uttering (31) implies the current existence of a machete in the speaker’s possession, whereas the utterance of (32) merely states that a machete has (ever) existed that is in the speaker’s possession, but not necessarily at the time of the speech act. (32) is thus not

---

\(^{29}\) Despite its translation as ‘have’, *hoto-* does not function as an auxiliary verb in Tukang Besi, not occurring with tense-aspect-modality meaning in conjunction with other verbs.
a paraphrase of (31), since it does not entail the same current existence of the possessum that is part of the meaning of (31).

Applying the other tests for incorporating structures that were developed in Section 2 and summarised in (30), we can see that the construction in (31) is indeed an NI structure. The test for causees cannot be applied to the verb, though this is not a feature of the verb hoto- specifically, but of it being a bivalent verb with a non-agentive subject argument (other bivalent verbs with non-agentive subjects, such as 'awa 'get', tarima 'receive', 'ita 'see (by chance), rodongo 'hear (by chance)' are similarly constrained against appearing in causative constructions). 26 Returning to hoto-, we can see that the verb can appear with an incorporated P. In these instances the construction behaves morphosyntactically just as does any other clause with an NI structure with respect to case marking, with no morphological case being assigned to the P.

Allows a non-case marked P?
(35) No-hoto kolikoli.
3R-own canoe
'He has a canoe.'
[no,hoto,kolikoli]

When we examine the behaviour of the hoto- incorporations with regard to the other tests that were developed to distinguish NI from non-NI structures in (30), we find significant differences:

Allows a case marked P?
(36) *No-hoto te kolikoli
3R-have CORE canoe
'He has a canoe.'

Aspect may follow P?
(37) a. No-hoto jonso=mo.
3R-have outboard.motor=PF
'They've got an outboard motor.'

26 The verbs 'ita and rodongo may appear in causative constructions, but only with the senses 'look at (intentionally)' and 'listen to (intentionally)', respectively, not 'see (by chance)' and 'hear (by chance)'. Compare the grammatical and ungrammatical readings of (35).

(i) Ku-pa-'ita te ika -wini=su 1 1 ama=su.
1SG-CAUS-see CORE fish PP-pull.in.on.line=1SG.GEN OBL father=1SG.GEN
'I showed the fish I'd caught to my father.'

Morphosyntactic tests proving the validity of the two readings of these verbs have been presented in Donohue (1996a). These may serialise with the verb sala 'do by accident/chance', but may not then form a causative or applicative construction, as in (ii).

(ii) * Ku-pa-sala-'ita te ika -wini=su 1 1 ama=su.
1SG-CAUS-accident-see CORE fish PP-pull.in.by.line=1SG.GEN OBL father=1SG.GEN
*I made my father happen to see the fish that I'd caught.'
* 'I accidentally showed the fish I'd caught to my father.'
* 'I caused my father to happen to catch sight of the fish I'd caught.'

The verbs 'awa and tarima (and others), which are lexically specified as taking non-agent As, may never appear in a morphological causative construction: * nopa'awa, * nopatarima.
b. *No-hoto=mo jonso.
   3R-have=PF outboard.motor
   ‘They’ve now got an outboard motor.’
   (eg., they no longer borrow or lease, but not own)

P may appear with coreferential P agreement?

(38) *No-hoto=e na wunua
   3R-have=3p NOM house
   ‘They’ve got a house.’

We might think of the ungrammaticality of (38) as being purely morphological: the verb hoto-
does not permit the use of P enclitics. But we should note that it is possible for just the P
enclitics to appear with the verb, as long as there is no separate nominal, either incorporated
or in its own NP.

(39) a. No-hoto=e
   3R-have=3p
   ‘They’ve got (it/them/one).’

b. *No-hoto wunua=e
   3R-have house=3p
   ‘They’ve got a house.’

   [noho?o?e]

   Adverb precedes P?

(40) *No-hoto molengo wunua
   3R-have long.time house
   ‘They’ve had a house for ages.’

Compare the performance of the incorporation found with hoto- constructions with that
observed for other verbs:

<table>
<thead>
<tr>
<th>(41)</th>
<th>General NI</th>
<th>hoto- NI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allows a case marked P?</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Aspect may follow P?</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Causee is oblique?</td>
<td>–</td>
<td>n/a</td>
</tr>
<tr>
<td>May have coreferential P clitic?</td>
<td>–</td>
<td>– (+)</td>
</tr>
<tr>
<td>Adverbs may precede P?</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Clearly the construction found with the verb hoto- is more similar to that described in (30) for
most other bivalent verbs (essentially, all those with theme or patient Ps) than it is to non-
incorporating structures. Despite this similarity, a syntactic paraphrase of the NI structures is
not available: the P of hoto- may not appear as a regular argument of the verb, in a fully cased
NP, but rather must appear in an NI structure. We thus appear to have an example of a
morphological noun incorporation structure which does NOT have a syntactic paraphrase;
when we know that we can say Kuhotokabali ‘I-machete-have.’ in Tukang Besi, as in (31),
we cannot correctly predict the existence of either * Kuhoto te kabali or * Kuhoto’e na kabali,
‘I-have a/the machete.’, as seen in the analogous (36) and (38). This is not the case for all NI
structures, but is true for those based on the verb hoto-. There are, however, mitigating
factors, and these will be addressed in the following section.
5. Possible models

The argument structure of hoto- might be relevant to this pattern. If we accept hoto- as a bivalent verb, since monovalent verbs in Tukang Besi do not allow NI and do not allow P enclitics, then it is the only bivalent verb that denotes a state (or inceptive state), rather than an activity (i.e., an unbounded / atelic predicate) (Donohue 1999:91).

Verbs that might be expected, from a comparative perspective, to be bivalent and stative, such as 'love' or 'hate', are coded as monovalent reciprocal verbs in Tukang Besi, not bivalent ones—love, for instance, is expressed with the verb root ilu 'lust after, be sexually attracted to'. This verb can be derived with the reciprocal prefix: po-ilu, in which both parties are preferentially coded as a single NP representing the coordinated S, as in (42), or else with an accompaniment applicative to introduce the P, as in (43) a and b, showing forms without, and with, pronominal enclitics for P.28

(42) No-po-ilu na [no La Kasi kene ta].
   3R-REC-lust NOM La Kasi COM 3SG
   'She and La Kasi are in love.'

(43) a. No-po-ilu-ngkene te la na La Kasi.
   3R-REC-lust-APPL CORE 3SG NOM La Kasi
   'La Kasi loves her.'

b. No-po-ilu-ngkene=’e te La Kasi.
   3R-REC-lust-APPL=3P CORE La Kasi
   'La Kasi loves her.'

Some speakers allow forms such as (44), in which poilu appears to function as a normal morphologically simple bivalent verb, showing two arguments without the use of the applicative. Clause such as (45), however, in which the P is shown as an overtly distinct separate NP, are rarely accepted, suggesting that the process of reanalysis of poilu has not gone to completion, since this predicate does not allow all the options of normal bivalent verbs (as shown earlier in (1) and (2)).

(44) No-po-ilu=’e te La Kasi.
   3R-REC-lust=3P CORE La Kasi
   'La Kasi loves her.'

(45) *Ha? No-po-ilu te ia na La Kasi
   3R-REC-lust CORE 3SG NOM La Kasi
   'La Kasi loves her.'

Similarly, expressions of hating will involve a verbal+nominal complex predicate, which is monovalent. The hated person can be expressed either in a dative construction, with the

27 The one exception being motion verbs that are serialised with numeral verbs, as discussed in Donohue (1999:196-199, 2004b:235-236).
28 For a point of comparison, note (i), which uses the bivalent verb root in its undervived form, allowing agreement with P enclitics without the necessity of a further applicative derivation.

(i) No-ilu=’e te La Kasi.
   3R-lust=3P CORE La Kasi
   'La Kasi lusts after her.'
polyfunctional ako, or with the locative applicative on the verb. Attempts to treat ja’olaro as an unanalysable bivalent verb are even less welcome that with pohlh, shown in (48).

(46) **Ku-ja’o laro ako te La Kasi.**
    1SG-bad inside DAT CORE La Kasi
‘I hate La Kasi.’

(47) a. **Ku-ja’o laro-mi te La Kasi.**
    1SG-bad inside-APPL CORE La Kasi
    ‘I hate La Kasi.’

    b. **Ku-ja’o laro-mi=’e na La Kasi.**
    1SG-bad inside-APPL=3P NOM La Kasi
    ‘I hate La Kasi.’

(48) a. *Ku-ja’o laro=’e.
    1SG-bad inside =3P
    ‘I hate him.’

    b. * Ku-ja’o laro te La Kasi
    1SG-bad inside CORE La Kasi
    ‘I hate La Kasi.’

There is a general lexical constraint against bivalent verbs expressing stative predicates (see also Donohue and Donohue 2004). The only bivalent, stative verb in the language is hoto-. The unusual behaviour with respect to incorporation structures is then, perhaps, not so surprising. Additionally, we might note that, in contrast to clauses such as (38), the sort exemplified in (49) is fully grammatical (earlier presented as (39)a).

(49) **No-hoto=’e.**
    3R-have=3P
    ‘They’ve got it/one.’

We can infer however, that in contrast to other verbs in Tukang Besi, there is a requirement with hoto- that the P be expressed directly on the verb, either as an incorporated nominal or as a pronominal P agreement marker; a free nominal representing the possessor is not an acceptable alternative, as seen in the ungrammaticality of both (36) and (38), regardless of whether there is a pronominal clitic on the verb or not. This shows that, while statements such as Mithun’s regarding NI and its paraphrasability may well prove to be true when applied to entire languages, there are constructions within languages, such as the verb hoto- in Tukang Besi, that function without regard for these generalisations.

With reference to the earlier discussion of apparent ‘obligatory incorporation’ in Eskimo, the Tukang Besi data cannot be treated simply as derivational because they can appear with only pronominal inflection following, and no full nominal. That is, while the Chukchee, Eskimo and Niue data show constructions that must appear with an independent lexical item (or phrase, in the Niue case), the Tukang Besi hoto- allows for a construction with no independent lexeme, as in (49). Hence the verb appears with only inflectional marking, while the morphemes in the other languages mentioned here are not morphologically ‘complete’ without a lexical item: they are specified as attaching to a word/phrase of a particular category.
The question remains as to why *huto-* would require incorporation. As mentioned at the start of this section, *huto-* is the only stative bivalent verb. It might be that there is a constraint on the realisation of argument structure such that only non-stative verbs allow a two-place predicate to be formed in syntax. Support for this hypothesis is the fact that most bivalent verbs cannot be used with non-agentive subjects. Translations of such English sentences as ‘The knife cut the meat.’ or ‘The falling rock broke the bottles.’ involve monovalent main verbs, with instrumental causers, as seen in (50) and (51). We can see that either the anticausative *mo-* is used to derive a stative, monovalent verb from the normally bivalent *hugu,* or else an ambitransitive verb such as *plisa* is used in its monovalent function, to mean ‘be broken, (happen to) break’. There are very serious constraints on the realization of non-sentient participants as an A or a clause, with a verb that requires an agent.

(50) No-mo-hugu=mo na ne'i (ako) te poda.
   3R-ANTICAUS-chop=pF NOM meat INSTR CORE knife
   ‘The meat is chopped (by the knife).’

(51) No-pisa=mo na botoru (ako) te watu b[um]uti.
   3R-break=pF NOM bottle INSTR CORE knife fall SI
   ‘The bottle broke (because of the falling stone).’

The realisation of two arguments can thus only be accomplished lexically, through NI. This would appear to be an instance of morphological restrictions applying to a construction in a language based purely on the semantic roles of the (one of) the arguments of a verb (see Donohue 1996b for documentation of semantic role restrictions in Tukang Besi relative clause formation). (It is probably relevant that the ‘object’ of *hoto-* is, as has been seen, indefinite (though not necessarily nonspecific). It is possible that this also plays a part in the incorporating behaviour of this verb, but the fact that not all indefinite and nonspecific objects must be incorporated weakens this hypothesis.)

Given the lexical nature of the NI advocated here, we might think that Rosen’s (1989) theories on compound versus classifier NI might be relevant. Rosen claims that there is a cluster of properties that can be associated with NI structures, and which can result in morphologically monovalent predicates. Among these properties are the existence of null proforms: pro-drop or null NP heads are allowed in what Rosen labels ‘classifier NI’ languages. This parameter is seen in the acceptability of phrases like ‘the big [one]’ with no pronoun (which is, of course, ungrammatical in English). In Tukang Besi we find the morphological intransitivity that Rosen predicts to be associated with compound NI, but, unlike the other examples of languages with this kind of NI pattern discussed by Rosen, Tukang Besi allows both types of null pro-forms, which should indicate classifier NI, rather than compound NI. This combination of grammatical features in the one language does not negate Rosen’s conclusions, but does break the mould in that Tukang Besi NI constructions do not allow stranding of modifiers outside the verb, the archetypical property of classifier NI. Even if the Tukang Besi *hoto-* construction is not a syntactic NI structure, it still is an unpredicted example of obligatory compounding with a non-bound lexical item, and poses problems for analysis.

6. The size of the incorporated element

I present here a short note on the size of the incorporated element. I have referred to the process of nouns appearing in the same syntactic and prosodic word as verbs as being ‘noun
incorporation', but the existence of clauses such as (52) as well as those of the form seen in (53) is suspicious. Although *kuhotokabali* and *mohama* form separate phonological words for the purposes of stress placement, they are not separate words syntactically. In (54) we can see that aspectual clitics can only appear preceding, or following *kabali mohama*, but not intruding into this unit (compare with the examples in (37), in which we have seen that aspectual clitics may intrude between a verb and its incorporated nominal). This, along with the lack of case marking, suggests that *mohama* is not simply a stranded modifier, such as has been reported in many other incorporating languages, but is in fact a part of the incorporation structure.

(52) *Ku-hoto kabali mohama.*
    1SG-own machete sharp
    'I have a sharp machete.'
    [*kuhoʔtokaʔbalimoʔhama*]

(53) *Ku-hoto kabali.*
    1SG-own machete
    'I have a machete.'
    [*kuhoʔtokaʔbalai*]

(54) a. *Ku-hoto=mo kabali mohama.*
    1SG-have=PF machete sharp

b. *Ku-hoto kabali mohama=mo.*
    1SG-have machete sharp=PF

c. *Ku-hoto kabali=mo mohama*
    1SG-have machete=PF sharp

Further support for the idea that a noun and an adjective form a special constituent can be found when we examine the structure of normal cased NPs (nominative NPs behave somewhat differently in this respect; see Donohue 1999:304).

Genitive clitic and adjective modifiers
(55) a. *Ku-ala te kabali mohama=su.*
    1SG-fetch CORE machete sharp=1SG.GEN
    'I fetched my sharp machete.'

b. *Ku-ala te kabali=su mohama.*
    1SG-fetch CORE machete=1SG.GEN sharp
    'I fetched my sharp machete.'

Genitive clitic and demonstrative modifiers
(56) a. *Ku-ala te kabali=su iso.*
    1SG-fetch CORE machete=1SG.GEN yon
    'I fetched that machete of mine.'

b. *Ku-ala te kabali iso=su*
Genitive clitic and relative clause modifiers

(57) a. Ku-alá te kabali=su i-hoko-hama.
1SG-fetch CORE machete=1SG GEN PP-FACT-sharp
‘I fetched my machete which I’d sharpened.’

b. *Ku-alá te kabali i-hokohama=su

Genitive clitic and numeral modifiers

(58) a. Ku-alá te kabali=su dua-mata.
1SG-fetch CORE machete=1SG GEN two-CLF
‘I fetched my two machetes.’

b. *Ku-alá te kabali dua-mata=su

The structure of the ‘core’ part of the NP can be represented as shown in (59) (from Donohue 1999:305).

(59)

NP

| N=POSS Rel. Cl. Demonstrative

| NUM-CLF

| (Adj)¹

The important feature of the NP is that possession is not marked on the noun, but on an N’ unit that consists of the head noun and any one adjective modifying it. Importantly this is an adjective, not an adjective phrase; if the adjective is expanded to a phrasal unit, it must appear in a relative clause, and outside the N’ that is the domain of genitive marking.

Plain adjective precedes genitive marking

(60) a. te ‘obu to’oge=su
CORE dog big=1SG GEN
‘my big dog’

b. *te ‘obu saori to’oge=su
CORE dog very big=1SG GEN
‘my very big dog’

Extended adjective follows genitive marking

c. *te ‘obu [rc s[um]aori to’oge]=su
CORE dog very SI big=1SG GEN
‘my very big dog’

d. te ‘obu=su [rc s[um]aori to’oge]
CORE dog=1SG GEN very SI big
‘my very big dog’

In short, while the unit that can show incorporation is not a full NP, it is larger than a simple lexical head. The incorporatable unit is an intermediate phrasal category, an N’ that maximally allows for a noun, a single adjective, and a genitive clitic. Note the following ungrammatical attempts to incorporate N+ units that are larger than an N’.

194
7. Other obligatory incorporation

The other type of so-called ‘obligatory incorporation’ that has been widely reported in the literature, the appearance of what Baker (1988:142) calls ‘N-V postbases’ which ‘have the morphophonological properties of bound forms, and are always listed as derivational affixes’ in Eskimo languages, is described by Sadock (1980) as being ‘not at all the process [nomin incorporation] – Sadock] described by Sapir (1911). The Eskimo phenomenon under consideration is a derivational process, as Sapir pointed out.’, and Mithun (1986:32), discussing these same affixes, notes that ‘In incorporating languages, a verb minus its incorporated Noun is still a well-formed verb; but in Greenlandic, a denominal verb minus its noun stem would be no word at all.’

This same label of ‘obligatory incorporation’ might be thought to apply to the Niuean fai ‘existential’ construction. However, Seiter (1980:77-78) notes that ‘although the fai possessive/existential construction apparently involves an incorporated nominal, the two facts just mentioned [obligatoriness and the possibility of modification by a relative clause – MD] suggest that the fai-construction is not derived through the rule of Noun Incorporation posited earlier.’ This would appear to rule out an NI analysis involving obligatory incorporation, but more recent work on Niuean (Massam 2001) shows that, rather than incorporating NPs in a problematic fashion, Niue does in fact exhibits full phrasal NP-‘incorporation’. The fai construction described by Seiter for Niuean looks like the following clause (Seiter 1980:76):

(64) To fai falaoa apogipogi.
   FUT exist bread tomorrow
   ‘There will be bread tomorrow.’

It is likely that the kind of obligatoriness that is described here for Tukang Besi is similar to the construction in Niue, though there are some differences between them (Tukang Besi shows N incorporation, and cannot incorporate full NPs, while this is reportedly possible in Niue; Asudeh and Ball present an analysis that is closer to the Tukang Besi facts than is Massam’s account of Niue).

There are other examples of what is described as ‘obligatory incorporation’ in Baker (1988:130ff), but these are dealing with affixes that are never found without a lexical root (such as the antipassive affix, analysed in Baker as an instance of noun incorporation, in which the ‘noun’ is obligatorily incorporated, but does not correspond to any lexical nominal).
References


