In Tukang Besi the selection of subject is marked on the verb with the use of different agreement markers, while case marking is constant, thus leading to a system in which, depending on the diathesis employed, subject marking on the verb follows completely different paradigms. Some minor sentence types, and differently-marked subordinate sentence types, are also examined, and I explore the implications for the separation of argument structure and grammatical functional structure.

1. HOW SUBJECTS CAN BE DIFFERENT

Tukang Besi, an Austronesian language of Southeast Sulawesi in central Indonesia, is a language in which the grammatical subject of a bivalent clause may be marked in different ways independently of their referential properties and without any relation to the semantics of predicates. As with other more northern Austronesian languages, such as Tagalog, the determining factor is verbal diathesis. Case marking is (usually) an explicit indicator of subject status, while verbal agreement is not always indicative of the functional status of the argument it references.

Before attempting to show the ways in which subjects are uniformly or non-uniformly indicated in the morphology and syntax, I must clearly define what I mean by the term 'subject'. The notion of 'subject' has been used in several different ways by linguists. Without reviewing the extensive literature from a variety of different theoretical persuasions, we need to review the distinctions that were popularised by Schachter in two seminal papers (Schachter 1976, 1977) which detailed the difference between what can in modern terminology be described as argument structure positions and pragmatically/syntactically salient arguments.

In argument structure terms we can identify three 'positions' for monovalent and bivalent predicates, arranged in terms of the ordering of the different arguments of that predicate:

The arguments are ordered in terms of the thematic hierarchy; I assume the version of the hierarchy described in Bresnan and Kanerva (1989):

agent > beneficiary > goal/experiencer > instrument > theme/patient > locative

The details and labels of the hierarchy are not important for the purposes of this exposition, simply the relative position in the ordered hierarchy. Alternative ways to rank arguments without reference to

(H. de Hoop and P. de Swart (eds.), Differential Subject Marking, 247–279. © 2008 Springer.)
Following (approximately) Comrie (1978) and Andrews (1985:68) I shall use the abbreviations A, S and P to refer to the positions (loosely) described in (1), (2) and (3), respectively. An S is the single argument of a monovalent verb; it is the highest role in its subcategorisation frame, and can simultaneously be described as the lowest role in the frame. An A is the most agent-like argument of a polyvalent verb, the highest role, which is not simultaneously the lowest role. The label P refers to the non-A argument in a prototypical bivalent verb, and to the argument in a trivalent (or quadrivalent) predicate which shows the same morphosyntactic behaviour. It is the lowest (second highest; see footnote 2) role in its subcategorisation frame, but it is not simultaneously the highest. These are syntactic roles in the sense of relationships existing at argument structure that can frequently be shown to affect morphosyntactic categories in languages, and they are not descriptors of grammatical status, though the identity of grammatical functions such as subject and object may be defined, after examining the properties of appropriate constructions in the language, in terms of the appropriate groupings of these roles.

Using this terminology we can describe the syntactically privileged argument in most English syntactic constructions as being the S or A (as appropriate for the verb) of the clause; this is the category that Philippinists such as Schachter call 'actor', and which in case terms is the nominative grouping (S and A, as opposed to P, the accusative). 2 On the other hand the syntactically most privileged argument in many Mayan languages (such as Mam; see England 1983) or Inuit (see the summary in Manning 1996) is the S of a monovalent clause or the P of a bivalent clause; authors such as Gerdes (1988) describe this syntactic grouping as absolutive, the same label used to describe case or agreement referring only to this grouping of arguments, and in contrast to ergative, which refers to an A. Schachter shows that while some syntactic constructions are restricted to the S or A in the clause, other constructions refer to an argument which in a polyvalent clause might be the A, but might also be

semantic roles have been presented by Kiparsky (1997) and Wunderlich (1997). While these models do not refer to particular semantic roles and their ordering, they require a stipulatively ordered set of semantic explications to achieve a similar model.

The difference between specifying 'lowest role' and 'second role' emerges only when we examine three-place verbs. In some languages, including Tukang Besi, it is the second argument of the three core arguments, the recipient in an agent, recipient, theme frame, that is treated in the same way as the second and lowest role in the subcategorisation frame of a two-place verb (\(\langle_{-2},_{-3}\rangle\)), in some other languages (very few in number) it is the theme in such a subcategorisation frame that behaves as the second and lowest argument in a bivalent frame. For these languages we need to recognise a third 'position', one that is neither highest nor lowest in the subcategorisation frame.

Falk (2000a, 2000b) uses the label \(\text{GF}^\ast\) to refer to this (apparently universally relevant, though not universally privileged) grouping, to avoid implications of morphological marking that might be present in the label 'nominative'. \(\text{GF}^\ast\) should be read as 'grammatical function that is defined by being the highest role in argument structure.'

DIFERENT SUBJECTS, DIFFERENT MARKING

A different argument, such as the P. These latter properties are, following Kroeger (1993), taken to be tests for 'subject'. The constructions that can be described in terms of argument structure configurations alone, and which do not vary with the overtly marked verbal diathesis, are not taken to be evidence for a separate module of grammatical functions, and so cannot be interpreted as tests for subject status. Only those constructions which are sensitive to the verbal diathesis, and which are not easily explicated in terms of argument structure positions, are assumed to be subject tests. I shall take 'object' to refer to any direct argument of the verb that is not the subject; as a less-salient participant in the discourse, grammatical tests for object status are less common than those that test subject status. This asymmetry has been noted at least since Plank (1984).

Variation in subject marking is normally described as being found in terms of the semantic content of the predicate, the status of the predicate as a main or subordinate element of the sentence it is in, and the tense-aspect-mood of the clause. In Tukang Besi we find that all of these factors play a part in determining the morphological marking of subjects, as well as the diathesis system of the language, which is marked in a way that leads to a typologically unusual morphosyntactic system.

2. THE INFLECTIONAL MORPHOSYNTAX OF TUKANG BESI

Tukang Besi is 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, 2004).

It is a verb-initial, subject-final language. The syntax shows many similarities with the Philippine languages, but also a significant number of discrepancies. The verbs show obligatory agreement for the S or A by verbal prefixes, and optional agreement for the P by enclitic. 4 Case marking choices are limited to the nominative 'na', the genitive 'na', the oblique 'i / di' (the former used in future/irrealis contexts, the latter and underspecified form used in other cases), and the underspecified (core) case 'te'. 5 6 The agreement and case marking systems interact as shown in the following examples. While there is only one marking choice for monovalent clauses, described in 2.1, bivalent clauses have two variants, which are discussed in 2.2 and 2.3, as well as more detail in section 3.

4 These pieces of morphology cannot be simply considered instances of 'incorporated pronouns'; for arguments against this stance, see Donohue (1999:123-129).

5 Tukang Besi examples are transcribed in the phonemic transcription used in Donohue (1999). The conventions are broadly equivalent to IPA standards, with the following exceptions: b represents a voiced impled bilabial stop, d represents a voiced impled dental stop, t represents a glottal stop, g represents a velar nasal, and a represents a high back unrounded vowel.

6 The gloss 'nominative' is used here in the tradition found in, amongst others, Bell (1976, 1983) and Kroeger (1993), and does not refer to a grouping of S and A. Rather, it refers to a case that marks the grammatical subject of the clause. The use of this gloss pre-empts the analysis to be presented in section 3, but is maintained here for the sake of consistency.
2.1 Marking Ss

In (4) the subject of the monovalent clause is marked by the agreement prefix ku- on the verb and by the use of the nominative case marker na on the NP (with local pronominal arguments the use of free NPs is at best optional, and is strongly dispreferred).  

(4) No-waliaco=mo (na ia) i kampo
3R.S/A-return=PFV NOM 3SG OBL village
'S/he returned to the village.'

The same agreement and case marking patterns can be used with all monovalent verbs, regardless of their semantic content/event structure. Evidence of this can be seen in the following examples.  

agent S
(5) No-rau na ana
3R.S/A-yell NOM child
'The child is yelling.'

experiencer S
(6) No-malino na ana
3R.S/A-lonely NOM child
'The child is lonely.'

patient S
(7) No-tunu na kau
3R.S/A-burn NOM wood
'The wood is burning.'

The prefixal agreement paradigm varies in realis versus irrealis contexts, for all but first person singular Ss. The differences between the two prefix sets can be seen in the following paradigm for mo'aro 'be hungry'.

<table>
<thead>
<tr>
<th>Realis</th>
<th>Irrealis</th>
<th>'was/is hungry.'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>ku-mo'aro</td>
<td></td>
</tr>
<tr>
<td>2SG</td>
<td>'u-mo'aro/nu-mo'aro</td>
<td></td>
</tr>
<tr>
<td>3SG</td>
<td>no-mo'aro/o-mo'aro</td>
<td></td>
</tr>
<tr>
<td>1PA</td>
<td>ko-mo'aro</td>
<td></td>
</tr>
<tr>
<td>1PL</td>
<td>to-mo'aro</td>
<td></td>
</tr>
<tr>
<td>2PL</td>
<td>ki-mo'aro</td>
<td></td>
</tr>
<tr>
<td>3PL</td>
<td>no-mo'aro/o-mo'aro</td>
<td></td>
</tr>
</tbody>
</table>

As can be seen there is no distinction between singular and plural for third person, though there are two variants for this person, with and without the n. The second person has two variants in the realis, 'n- and nu-; the second person also distinguishes two numbers, singular and nonsingular (= plural), while first person distinguishes a singular, paucal and plural category. There are different prefixes for realis and irrealis for all but the 1SG S, and the relationship between the realis and irrealis forms is not consistent; in some cases there is a vowel alternation, while in others the consonant of the prefix varies. This is clearly a difference in subject marking based on the mood of the clause. Nonetheless, all Ss show the same behaviour, with the same number of distinctions maintained in both realis and irrealis.

The fact that all monovalent verbs show the same coding options means that we are not dealing with a 'split-S' system, in which the single arguments of some monovalent predicates are marked in one way (the same as the A of a bivalent verb), and some are marked differently (the same as the P of a bivalent verb). In fact, as we shall see, the prefixal agreement coding strategy seen with Ss is the same as the agreement found with As.

2.2 Marking As

In (9) a bivalent clause with similar verbal prefix and nominative case for subject is found. (10) shows that the order of the two core arguments is important in the bivalent clause.

(9) No-'ita to ana (na ia)
3R.S/A-see CORE child NOM 3SG
'S/he saw a child.'

V A P

(10) No-'ita na ia te ana

The same prefixes that are used to mark Ss are used with As. As with Ss, different prefixes are used for realis and irrealis clauses, as shown in (11) with the verb manga 'eat'. The paradigm is identical to that used to mark agreement with Ss.

<table>
<thead>
<tr>
<th>Realis</th>
<th>Irrealis</th>
<th>'ate/eats.'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>ku-manga</td>
<td></td>
</tr>
<tr>
<td>2SG</td>
<td>'u-manga/nu-manga</td>
<td></td>
</tr>
<tr>
<td>3SG</td>
<td>no-manga/o-manga</td>
<td></td>
</tr>
<tr>
<td>1PA</td>
<td>ko-manga</td>
<td></td>
</tr>
<tr>
<td>1PL</td>
<td>to-manga</td>
<td></td>
</tr>
<tr>
<td>2PL</td>
<td>ki-manga</td>
<td></td>
</tr>
<tr>
<td>3PL</td>
<td>no-manga/o-manga</td>
<td></td>
</tr>
</tbody>
</table>

The free DP na ia is grammatical in this sentence, but is unlikely to appear without a strongly contrastive pragmatic context. See 2.4.

For monovalent verbs with experiencer subjects there is an alternative agreement option using the genitive enclitics that will be described in 2.4. This will be discussed in relation to table 5.
The same prefixal coding is found regardless of the semantic role of the A, as shown in the following different predicates:

agent A

(12) No-

\[
\text{manga te } bae na \text{ ana}\nonumber \\
\text{3R.S/A-eat} \text{ CORE rice} \text{ NOM child}\nonumber \\
\text{ 'The child ate rice.'}
\]

recipient A

(13) No-

\[
\text{awa te } bae na \text{ ana}\nonumber \\
\text{3R.S/A-receive} \text{ CORE rice} \text{ NOM child}\nonumber \\
\text{ 'The child received the rice.'}
\]

experiencer A

(14) No-

\[
\text{ita te } bae na \text{ ana}\nonumber \\
\text{3R.S/A-see} \text{ CORE rice} \text{ NOM child}\nonumber \\
\text{ 'The child saw the rice.'}
\]

instrument (effector) A

(15) No-

\[
\text{motiti=e na } kau \text{ te } 'oloo\nonumber \\
\text{3R.S/A-dry=3P NOM wood} \text{ CORE sun}\nonumber \\
\text{ 'The sun dried the wood.'}
\]

The enclitic = 'e in the last example will be explained in the following section.

2.3 Marking Ps

In addition to the bivalent clauses seen in 2.2 an alternative bivalent clause coding option exists. In this clause type the verb shows the same prefixal agreement for the A that was seen in 2.1 and 2.2, but also has an enclitic that shows agreement with the P. The case marking is also changed, the nominative na that was found with subject Ss or As now appearing with the P, and the (non-nominative) core case te that appeared with Ps in 2.2 is associated with the A in the double-agreement clause type. Word order is freer in this clause type than in the clause type without P agreement seen in section 2.2.\(^9\)

\[
\begin{array}{ccc}
V & A & P \\
\text{(16) No-} & \text{ita=e (te ia) na ana}\nonumber \\
\text{3R.S/A-see=3P CORE} & \text{3SG NOM child}\nonumber \\
\text{ 'S/he saw the child.'}
\end{array}
\]

\[
\begin{array}{ccc}
V & P & A \\
(17) & \text{– No-} & \text{ita=e na ana te ia}
\end{array}
\]

\(^9\) Note the use of a definite article with 'child' in the English translation of (16), as opposed to the indefinite in (9). There is a consistent difference between clauses with, and without, P enclitics in that those with P enclitics show greater definiteness, specificity, telicity, affectedness, or a combination of these features.

Unlike the prefixes used to mark agreement for an S or A, there is no variation in the form of the P enclitics for realis or irrealis mood, as can be seen in (18) in which the A is kept constant as a third person, marked with a third person prefix no-

\[
\begin{array}{cc}
\text{Realis} & \text{Irrealis} \nonumber \\
(18) & \text{They missed / miss ...} \\
1SG & \text{no-moniasi=aku} \nonumber \\
2SG & \text{no-moniasi=ko} \nonumber \\
3SG & \text{no-moniasi=e} \nonumber \\
1PA & \text{no-moniasi=kami} \nonumber \\
1PL & \text{no-moniasi=kita} \nonumber \\
2PL & \text{no-moniasi=komi} \nonumber \\
3PL & \text{no-moniasi=e} \nonumber \\
\text{na-moniasi=aku} & \nonumber \\
\text{na-moniasi=ko} & \nonumber \\
\text{na-moniasi=e} & \nonumber \\
\text{na-moniasi=kami} & \nonumber \\
\text{na-moniasi=kita} & \nonumber \\
\text{na-moniasi=komi} & \nonumber \\
\text{na-moniasi=e} & \nonumber \\
\end{array}
\]

In addition to these regular P enclitics, which may be used to mark the P of any bivalent verb, an additional set of enclitics are optionally found with Ps that are recipients or beneficiaries. (19) shows that the verb hoti 'donate (food or clothing) to someone charitably' can take either the P enclitics described above, or the 'dative' enclitics. Most other verbs, such as siasia 'beat' cannot appear with the dative enclitics, shown in (20).

\[
\begin{array}{cc}
\text{Regular P enclitics} & \text{'dative' P enclitics. They donated to ...} \\
(19) & \text{'dative' P enclitics. They donated to ...} \nonumber \\
1SG & \text{no-hoti=aku} \nonumber \\
2SG & \text{no-hoti=ko} \nonumber \\
3SG & \text{no-hoti=e} \nonumber \\
1PA & \text{no-hoti=kami} \nonumber \\
1PL & \text{no-hoti=kita} \nonumber \\
2PL & \text{no-hoti=komi} \nonumber \\
3PL & \text{no-hoti=e} \nonumber \\
\text{no-hotinak} & \nonumber \\
\text{no-hotinso} & \nonumber \\
\text{no-hotine} & \nonumber \\
\text{no-hotinsami} & \nonumber \\
\text{no-hotingkita} & \nonumber \\
\text{no-hotingkomu} & \nonumber \\
\text{no-hotine} & \nonumber \\
\end{array}
\]

\[
\begin{array}{cc}
\text{Regular P enclitics} & \text{'dative' P enclitics. They beat ...} \\
(20) & \text{'dative' P enclitics. They beat ...} \nonumber \\
1SG & \text{no-siasia=aku} \nonumber \\
2SG & \text{no-siasia=ko} \nonumber \\
3SG & \text{no-siasia=e} \nonumber \\
1PA & \text{no-siasia=kami} \nonumber \\
1PL & \text{no-siasia=kita} \nonumber \\
2PL & \text{no-siasia=komi} \nonumber \\
3PL & \text{no-siasia=e} \nonumber \\
\text{*no-siasia=aku} & \nonumber \\
\text{*no-siasia=ko} & \nonumber \\
\text{*no-siasia=e} & \nonumber \\
\text{*no-siasia=kami} & \nonumber \\
\text{*no-siasia=kita} & \nonumber \\
\text{*no-siasia=komi} & \nonumber \\
\text{*no-siasia=e} & \nonumber \\
\text{*no-siasia=ne} & \nonumber \\
\end{array}
\]

This dative enclitic option is considered by most speakers as archaic, but is used by many speakers productively even with verbs that only have a beneficiary P by virtue of applicative derivation, as seen in (21). Here the same verb siasia is derived with a benefactive applicative = ako, and the beneficiary P is marked by the dative enclitics. It is also possible for the regular P enclitics to be used, as in (22).
Examples of Ps of various semantic roles with enclitic agreement, and the conditions on their use, are shown in (23) - (36). The dative enclitics may only be used with Ps that bear beneficiary or recipient roles, as mentioned above.

agent P (derived only)

(23) No-wila-ngene=aku te ana
3R.S/A-go=APPL-3G.P CORE child
'The child went with me.'

(24) *No-wila-ngene=naku te ana
3R.S/A-give=APPL-3G.P CORE child
beneficiary P

(25) No-hoti=aku te ana
3R.S/A-give.charitably=3G.P CORE child
'The child gave (food or clothing) to me.'

(26) No-hoti=naku te ana
3R.S/A-give.charitably=3G.DAT CORE child

beneficiaries are also found in derived verbs derived by applicatives, such as was seen in (21). Recipient Ps are only known to occur in two verbs, hu 'give' and kaku 'send'.

recipient P

(27) No-hu'u=aku te boku te ana
3R.S/A-give=3G.P CORE book CORE child
'The child gave me a book.'

(28) No-hu'u=naku te boku te ana
3R.S/A-give=3G.DAT CORE book CORE child

While instruments are rare as the lexical P of a verb they are found in a small number of verbs for which they are inherent, and are productively found when an instrumental applicative (identical in morphological form to the benefactive applicative) is used with a verb (as seen in (31) and (32)).

(intermediary) instrument P

(29) No-heka-batu=e na po'o te ana
3R.S/A-INTENS-stone=3P NOM mango CORE child
'The child threw a mango (as if it were a stone).'

(30) *No-heka-batu=ne na po'o te ana
3R.S/A-INTENS-stone=3DAT NOM mango CORE child

10 The morphologically complex verb pa-mate 'kill' (literally, 'cause to die', with the generic causative prefix pa) is used in cases where the person has been revaved, medically or magically following death. A permanent or irreversible death will be marked with the verb hoko-mate 'kill, murder', showing the same root mate 'die, be dead' prefixed with hoko-, the factive prefix. See Donohue (1999:205-211).
2.4 The inflectional patterns of Tukang Besi

The agreement, case-marking and word orders seen in 2.1-2.3 are summarised in the schemas shown in Table 1.

Table 1. Basic agreement, word order and case marking in Tukang Besi

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Variants</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>VAP (- VPA)</td>
<td>te na</td>
</tr>
<tr>
<td></td>
<td>P not indexed on V</td>
<td>A-V</td>
<td>VPA</td>
<td>na te</td>
</tr>
<tr>
<td>Monovalent</td>
<td>S-V</td>
<td>S-V</td>
<td>S</td>
<td>na</td>
</tr>
</tbody>
</table>

The full set of pronominal affixes and free forms is given in Table 2. The agreement prefixes that index S or A have variant forms depending on the mood, realis or irrealis, of the clause, while the other major verbal paradigm, marking a P, are invariant (and much closer phonologically to the independent pronouns, which may appear in any syntactic role). The dative P enclitics are clearly related to the generic P enclitics with the addition of a nasal component; in some cases the k of the generic P enclitics is cognate with an s, showing the sound change that is also found in the 1SG,GEN form =su, a sound change that has only marginally affected Tukang Besi in these morphemes, but which is widespread in other languages of western Buton. The genitive enclitics in the second last column have not been introduced yet, but will become relevant in Table 5, where they appear in a variety of minor clause types to indicate arguments of the verb, and in an optional role presented later in this section. While there are three different sets of pronominal enclitics, a single word may only host one enclitic.11

Table 2. Pronominal forms in Tukang Besi

<table>
<thead>
<tr>
<th>Position:</th>
<th>Pre-root</th>
<th>Post-root</th>
<th>Post-root Dative P</th>
<th>Post-root nominal POS1 R subordinate A, S, P</th>
<th>Independent (any)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role: (Mood): realis</td>
<td>SA</td>
<td>P</td>
<td>Dative P</td>
<td>nominal POSS1 R subordinate A, S, P</td>
<td></td>
</tr>
<tr>
<td>1SG</td>
<td>ku-</td>
<td>ku-</td>
<td>=aku</td>
<td>=naku</td>
<td>=su</td>
</tr>
<tr>
<td>2SG</td>
<td>'i=mu</td>
<td>ko</td>
<td>=o</td>
<td>=u</td>
<td>iko'o</td>
</tr>
<tr>
<td>3SG</td>
<td>no-lo</td>
<td>na-a</td>
<td>=?</td>
<td>=?</td>
<td>no</td>
</tr>
<tr>
<td>1PA</td>
<td>ko</td>
<td>ka</td>
<td>=ami</td>
<td>=sami</td>
<td>=nto</td>
</tr>
<tr>
<td>1PL</td>
<td>to</td>
<td>ta</td>
<td>=kita</td>
<td>=ggita</td>
<td>=mami</td>
</tr>
<tr>
<td>2PL</td>
<td>i</td>
<td>ki</td>
<td>=komu</td>
<td>=ngkomu</td>
<td>=miu</td>
</tr>
<tr>
<td>3PL</td>
<td>no-lo</td>
<td>na-a</td>
<td>=?</td>
<td>=?</td>
<td>no</td>
</tr>
</tbody>
</table>

The only TAM-conditioned split involves the use of different sets of prefixes, as shown above. All Ss are marked on the verb with the same prefixes that are used to

mark As, and none of them use the P enclitics. Compare (7), repeated below, with the ungrammatical clauses in (39)-(41), which show the same predicate with P, dative P, and genitive agreement markers.

(7) No-tunu na kau 3R.S/A-burn NOM wood
"The wood is burning."
(39) *Tunu=a e na kau burn=3P NOM wood
(40) *Tunu=ne na kau burn=3DAT NOM wood
(41) *Tunu=no na kau burn=3GEN NOM wood

While the most P-like Ss, those with theme or patient roles, show no difference in marking compared to the most A-like Ss, we do find some alternative coding with experiencer Ss. In addition to the normal prefixal agreement option, shown in (6) repeated below, an experiencer S may also be marked on the verb by means of genitive enclitics, as in (42).12

(6) No-malino na ana 3R.S/A-lonely NOM child
"The child is lonely."
(42) Malino=no na ana lonely=3GEN NOM child
The use of the genitive agreement option on verbs is only found with Ss, not experiencer As. It is worth noting, however, that a number of bivalent psych predicates, such as gau 'desire' and hada 'want', are more frequently used in nominal clauses than in verbal clauses, in which case the indication of the experiencer is found by means of genitive clitics. In (43) and (44) we can see that a regular verbal clause is not grammatical with genitive agreement marking on the verb.

Verbal clause with prefixes
(43) Ku-hada te po'o 1SG.S/A-want CORE mango
'I want a mango.'

Verbal clause with genitive enclitics
(44) *Hada=sa te po'o want=1SG. GEN CORE mango
'I want a mango.'

11 This means that polyvalent verbs will show agreement for at most two arguments, the A and the first P. Tukang Besi is a primary object/secondary object language.

12 P-enclitic or dative enclitic marking on the verb remains ungrammatical: *malino=a.e, *malino=ne. This genitive predication is only found with experiential verbs, not with any other kinds of predicates.
In (45) the nominal clause, of the form [te NP te NP], takes genitive marking on hada to indicate the experiencer of ‘want’. It is not grammatical to use verbal prefixes in a nominal clause, as seen in (46).

Nominal clause with genitive enclitics
(45) Te hada-su te po'ō
    CORE want=1SG.GEN CORE mango
    ‘I want a mango.’ (Literally, ‘My want(ing) is a mango.’)

Nominal clause with prefixes
(46) *Te ku-hada te po'ō
    CORE 1SG.S/A-want CORE mango
    ‘I want a mango.’

While this is not the same use of genitive enclitics to mark an argument of a verbal clause as was the experiencer S option, it indicates a preference, or at least consistent tendency, to avoid the prefixal forms with experiencers. Note that, even allowing for the genitive coding option available with experiencer Ss, we would not want to classify the Tukang Besi agreement system as showing a ‘split-S’ or ‘split-intransitive’ pattern, since genitive marking is not a feature of the agreement system for Ps. It is true that both the I agreement paradigm and the genitive paradigm are enclitic, but they are formally distinct sets of morphemes, and the use of P enclitics to index an S is ungrammatical, as shown in (39) and (40).

An agreement subset more similar to a split-S system is found when motion verbs are serialised with numeral verbs. Either of these verb types on their own can only show agreement by means of the prefixal agreement markers. The grammaticality of prefixal agreement with a motion verb is shown in (47) - (50).

Motion verb with pronominal prefixes
(47) Ko-mai=no
    1PA.S/A-come=PFV
    ‘We have come.’

Motion verb with pronominal enclitics
(48) *Mai=kami
    come=1PA.P
(49) *Mai=nsami
    come=1PA.DAT
(50) *Mai=mami
    come=1PA.GEN

Analogously to the motion verb, the prefixal agreement pattern of a numeral verb are shown in (51) - (54).

Motion verb with pronominal prefixes
(51) Ko-gana=no
    1PA.S/A-be, four=PFV
    ‘We are now four.’ / ‘There are now four of us.’

Motion verb with pronominal enclitics
(52) *Gana=kami
    be, four=1PA.P
(53) *Gana=nsami
    be, four=1PA.DAT
(54) *Gana=mami
    be, four=1PA.GEN

When these verbs are serialised together the only grammatical agreement pattern involves the use of both the prefixes and P enclitics; the enclitics may be taken from the set of plain P enclitics shown in (55), or from an additional set of enclitics that are found only in these constructions, illustrated in (56) and shown in full in (57).

The verb + numeral verb serialisation is found only with nonsingular numbers, and so there are no singular enclitics for 1SG, 2SG or 3SG.1 The dative or genitive enclitics may not be used in this construction. Other instances of verb serialisation do not require (or allow) different prefixal and enclitics.

Motion verb and numeral verb serialised
(55) Ko-mai-gana=kami
    1PA.S/A-come=be, four=1PA.P
    ‘The four of us came.’
(56) Ko-mai-gana=ngkami
    1PA.S/A-come=be, four=1PA.INTRADIR.P
    ‘The four of us came.’

Regular P enclitics
(57) 1PA ko-mai-gana=kami
    ko-mai-gana=ngkami
1PL to-mai-gana=kita
    to-mai-gana=ngkita
2PL i-mai-gana=komiu
    i-mai-gana=ngkomiu
3PL no-mai-gana=e
    no-mai-gana=e

‘Four of... came’

A possible explanation for the use of both prefixal and enclitic agreement markers for these predicates can be found in the nature of motion verbs: motion verbs, and some posture-assuming verbs, are predicates in which the one participant is simultaneously the instigator of the event, the agent, and the participant that undergoes a change of location, a theme. It is thus not completely unexpected that the marking for such a predicate should involve both an element that is typical of agentive agreement, the prefix, and an element that is associated with a P, a syntactic role that is most typically associated with a theme or a patient. The unexplained aspect is why this double-marking pattern should emerge only when the

1 In order to emphasise that only a single participant was involved in the action the adverb ala'a ‘only’ or pe'e 'self' + enclitics may be used, as in (i) and (ii).
(i) 'U-mai
    ala'a
    2SG.R.S/A-come only
    'Just you came.' or 'You just came.'
(ii) 'U-mai
    pe'e
    2SG.R.S/A-come self=2SG.GEN
    'You came on your own.'
motion verb is serialised with a numeral verb. For this conundrum I have no explanation, except to hypothesise that an earlier general double-marking pattern with intradirectives has been preserved only with the serial verb construction, while the simple motion verb construction has normalised to the general nominative-accusative pattern.

From the data in sections 2.1-2.4 we summarise the different agreement paradigms in Tukang Besi and their functions:

(58) prefixal sets  
  S,A agreement prefixes, obligatory on verbs  
  P enclitic  
  agreement clitics, optional on most bivalent verbs  
  dative enclitic  
  optionally used in place of the accusative agreement markers, only when marking a P that is benefactor or recipient  
  genitive enclitic  
  with main clause verbs only used to show agreement with an intransitive experiencer  
  prefix + (intransitive)  
  verbs  
  used with motion verbs that are serialised with numeral

The case marking system for core arguments is quite simple as shown in table 1: an S is marked with na, while in a polyvalent clause na can mark either the A, if the verb shows agreement only for the A and not for the P, or the P, if there is a pronominal enclitic on the verb.

Structurally we can determine that an S or an A are always external to the VP, while a P is external to the VP only if the verb is marked with P agreement enclitics. A locative adjunct can only follow a VP, and not intrude in it. With a monovalent predicate, in which the S is external to the VP, this means that a location can intrude between an NP subject and the verb. In addition to (4), repeated below, we can also find the grammatical (59).

V  S  Location  
(4) Ku-waliñu=mo (na iaku) i kampo  
  1SG/5/A-return=PFV NOM 1SG OBL village  
  "I returned to the village."  
  V  Location  S

(59) Ku-waliñu=mo i kampo (na iaku)  
  1SG/5/A-return=PFV OBL village NOM 1SG  
  "I returned to the village."

With a bivalent verb with single agreement for A, but not P, we find that a locative adjunct can appear in the following positions, based on (9). The adjunct may appear following or preceding the A, but may not intrude between the P and the verb.

When a bivalent verb is marked for agreement with both the A and the P a different set of possibilities is found. The adjunct may appear anywhere following the verb, not respecting the close constituency between the verb and the P that was evident in (60)-(62).

V  A  P  Location  
(60) Ku-ita te ana (na iaku) di kampo  
  1SG/5/A-see NOM child NOM village  
  "I saw a child in the village."  
  V  P  Location  A

(61) Ku-ita te ana di kampo (na iaku)  
V  Location  P  A

(62) *Ku-ita di kampo te ana (na iaku)

A further difference between Ps in the two different bivalent clause types can be seen in the grammatical positions available for time expressions. A time expression must always follow an S, or a na-marked A, but may precede or follow a na-marked P. This results in the following grammaticality judgements for clauses in which the verb lacks agreement for P.

V  P  A  Time  
(69) Ku-ita te ana (na iaku) dinggawi  
  1SG/5/A-see CORE child NOM 1SG yesterday  
  "I saw a child yesterday."  
  V  P  Time  A

(70) *Ku-ita te ana dinggawi (na iaku)  
V  Time  P  A

(71) *Ku-ita dinggawi te ana (na iaku)
When there is agreement for the P on the verb, the following judgements are found. Once again there is complete grammaticality for the time adverbial in any position following the verb.

\[
\begin{array}{cccc}
V & A & P & \text{Time} \\
(72) & \text{Ku'ita'e} & (\text{te iaku}) & \text{na ana dinggawi} & \text{ISG,5/A-see=3P CORE:ISG NOM child yesterday} \\
& V & A & P & \text{Time} \\
(73) & \text{Ku'ita'e} & (\text{te iaku}) & \text{dinggawi} & \text{na ana} \\
& V & \text{Time} & A & P \\
(74) & \text{Ku'ita'e} & \text{dinggawi} & (\text{te iaku}) & \text{na ana} \\
& V & P & A & \text{Time} \\
(75) & \text{Ku'ita'e} & \text{na ana} & (\text{te iaku}) & \text{dinggawi} \\
& V & P & \text{Time} & A \\
(76) & \text{Ku'ita'e} & \text{na ana} & \text{dinggawi} & (\text{te iaku}) \\
& V & \text{Time} & P & A \\
(77) & \text{Ku'ita'e} & \text{dinggawi} & \text{na ana} & (\text{te iaku}) \\
\end{array}
\]

These data concerning locative and temporal adjuncts, taken together, imply roughly the following structural positions.\(^{14}\)

\[
(78)
\]

???

\[
\begin{array}{ccc}
& \text{IP} & \text{te DP}_{S,A}, \text{na DP}_{P} \\
\text{VP} & \text{na DP}_{S,A} \\
V & \text{te DP}_{P} \\
\end{array}
\]

Further tests for the constituency of the P and the non-agreeing verb can be found in Donohue (1999, 2004).\(^{15}\) Note that an alternative morphological realisation of the nominative argument is available. It may appear preverbally, yet not in a topic position external to the clause. In this case it is marked with \textit{te}, the case marker that otherwise marks non-nominative core arguments. This split in marking emphasises the more general nature of \textit{te}, and the more restricted function of \textit{na} in the linker inventory of Tukang Besi.

We have now thoroughly examined the morphological and structural properties of core arguments in verbal clauses in Tukang Besi, finding considerable differences between marking strategies, revealing a nominative-accusative pattern in verbal agreement and a complex set of patterns in case marking and phrase structure. In the next section I shall show how these patterns are related to the question of subjecthood.

3. DETERMINING THE IDENTITY OF THE SUBJECT

The problems inherent in identifying a subject have been recognised by many authors (for instance, Keenan 1976, Schachter 1976, 1977, Foley and Van Valin 1977, 1984, Andrews 1985, Van Valin and La Polla 1997, and many others). While languages such as English and many western European languages display a subject that can be identified on the basis of argument structure positions ("select the argument that ranks highest in the thematic hierarchy" is a good rule-of-thumb way of identifying the argument that counts as subject in English), (western) Austronesian languages show a syntax that is more directly sensitive to pragmatic factors, and a split in 'subject properties' between the highest-ranked argument and the one that is morphologically marked as unique. Having examined the morphosyntax used to mark the core arguments, subjects and objects, in Tukang Besi we must now turn to the question of identifying the subject. It seems unproblematic that the sole argument of a (typical) monovalent verb is the subject of that predicate, and I shall take some of the syntactic properties of this type of argument as a guide in evaluating the grammatical status of the arguments of bivalent verbs. The constructions that shall be examined here include two that are useful in identifying the subject, and one, restrictions on bounding external relative clauses, that is dependent on the position of the argument in a thematic hierarchy, and so has no bearing on the identification of the subject but which is methodologically relevant to illustrate the different marking afforded to subjects of subordinate clauses.

When we examine the syntactic properties of the single subcategorised argument of a monovalent predicate, we find a number of tests that we can apply to determine the status of the arguments of polyvalent predicates. The constructions listed here are only a subset of those that are discussed in Donohue (1999), but serve to illustrate the argumentation.\(^{16}\)

1. Control floating quantifiers

Only the S in a monovalent predicate can control a floating quantifier, as seen in the singular interpretation of \textit{saba'ane} in (79).

\(^{14}\) The identity of the "P" is not certain. Topicisation is found to the left of the clause, and there are other reasons to feel that this is not a CP in the sense of a CP being an extra-clausal projection. Nonetheless, there are certain similarities between a CP and the "P" constituent shown here.

\(^{15}\) Topicisation also affects the case marking strategy, in that the nominative case may only appear preverbally. A nominative argument is a preverbal position must be marked with the underspecified case \textit{te}.

\(^{16}\) From here on I shall omit the gloss "/A/" with the verbal prefixes, letting their unique position differentiate them from the three onelitic pronominal paradigms, which shall continue to be labelled as "P", "DAT" and "GEN".
(79) [QUANT Saba’an] no-wila na mia kua kampo
all 3r-go NOM person ALL village
‘All of the people went to the village(s).’
*‘The person/people went to all of the villages.’

2. Control conjunction reduction
Conjunction reduction preferentially applies between core arguments in adjacent clauses. A textual example is presented later in this section, but the following sentences illustrate the principle.

(80) No-wila=mo na mia kene b’obu, maka no-helawe
3r-go=NOM person COMIT dog and then 3r-rest
‘The people went with the dog(s), and then (they), rested.’
*‘The people went with the dog(s), and then (they), rested.’

3. External relative clauses
An external relative clause is marked by the infix <um> in the verb, which shows no agreement. Only the S of the clause may appear as the head of the relative clause.

(81) te mia [RC w<um>ila kua kampo]
CORE person go.SI ALL village
‘the person/people who went to the village(s)’

(82) *te kampo w<um>ila
CORE person go.SI
‘the village(s) where (the people) went’\(^\text{17}\)

When we apply these tests to the two types of polyvalent clauses, we find the following results for subjecthood:

1. Controlling floating quantifiers
Only the A in a polyvalent predicate without P agreement, and only the P in a polyvalent predicate with P agreement, can control a floating quantifier, as seen in the unambiguous interpretation of Saba’an in (83) and (84).

Floated quantifier referring to an A
(83) [QUANT Saba’an] no-lemba te kaluku na amai
all 3r-carry CORE coconut NOM 3PL
‘All of them carried coconuts.’
*‘They carried all of the coconuts.’

Floated quantifier referring to a P
(84) [QUANT Saba’an] no-lemba=e na kaluku te amai
all 3r-carry=3P NOM coconut CORE 3PL
‘They carried all of the coconuts.’
*‘All of them carried coconuts.’

Only the argument that is eligible to be marked with nominative case, na, may control floating quantifiers.

2. Controlling conjunction reduction
The A in a polyvalent predicate without P agreement, and the P in a polyvalent predicate with P agreement, is the preferred controller and target of conjunction reduction in adjacent clauses.

Coreference between S and A, both nominative
(85) No-waliako=mo di kampo, maka no-lta te ana
3r-return=NOM OBL village and then 3r-see CORE child
‘I returned home, and then I saw a child.’

Coreference between S and P, both nominative
(86) No-waliako=mo di kampo, maka no-lta=e (te ana)
3r-return=NOM OBL village and then 3r-see=3P CORE child
‘I returned home, and then (a child/I) saw I.’

Only the argument that is eligible to be marked with nominative case, na, may control or be the target of zero anaphora between clauses.

3. Heading external relative clauses
The relative clause marked with the infix <um> may only be used with an A head, not a P. As with an S head, the verb is not marked with nominative agreement prefixes, but may optionally use P enclitics.

(87) te mia [RC k<um>oho te kau ]
CORE person chop.SI CORE wood
‘the person/people who chopped the wood’

(88) te mia [RC k<um>oho=e na kau ]
CORE person chop.SI=3P NOM wood
‘the person/people who chopped the wood’

(89) *te kau k<um>oho
CORE wood chop.SI
‘the wood that (the person) chopped’

(90) te kau [RC i-koho ]
CORE wood PP-chop
‘the wood that was chopped’

\(^{17}\) To be expressed grammatically the verb must be derived with a locative applicative, and then the P head relative clause prefixes used, as in (i).
Only the A may head a relative clause marked with the infix; a P may only head a relative clause marked by prefix, shown here in (90) as i-. The P-headed relative clauses are discussed in more detail in table 5.

The data on floating quantifiers, conjunction reduction and external relative clauses may be summarised as follows:

(91) external relative clause shows the same nominative-accusative distinctions as are made with the verbal agreement sets, and does not show a correlation with case marking
floating quantifiers follows the distinctions in case marking in that the argument marked with nominative case na is the argument that can be the restriction of a floating quantifiers
conjunction reduction follows the distinctions in case marking in that the argument marked with nominative case na is the argument that can control or be the target of conjunction reduction

Other tests will also single out the na-marked argument as being syntactically privileged (Donohue 1999, 2004). Most importantly, while constructions such as external relative clauses can be accounted for by appealing to argument-structure positions, the fact that the selection of the na-marked argument is divorced from considerations of argument structure hierarchies, either of the arguments of a bivalent predicate being eligible for this marking and this set of syntactic properties, means that we can only account for its syntactic status by assuming a privileged position in an ordered grammatical functions hierarchy. The conclusion, therefore, is that the na-marked argument is the subject of the clause it is in. Note that the notion of subject in Tukang Besi is not one that can be defined in terms of the labels A, S, and P. It is true that an S will be the subject of its clause, as the sole core argument in the subcategorisation frame, but in a bivalent clause we cannot uniquely identify subject with either A or P: rather than being syntactically predetermined, it is determined by pragmatic considerations, and monitored morphologically on both the verb and the choice of case marker.

Interestingly, but not importantly for the purposes of this paper, we should note that the diathesis found in bivalent clauses does not involve demotion: both the single-agreeing verbs described in 2.2 and the doubly-agreeing verbs detailed in 2.3 have two core arguments, a subject and an object. The syntactic situation is identical to that of the better-described Philippine-type languages such as Tagalog, a point to which I return in section 5.18

18 Evidence for the argument status of both A and P in both clauses described in 2.2 and those seen in 2.3 can be found through an examination of reflexive constructions:
(i) Ku-‘ita te karanamu
1SG-see CORE self=1SG.GEN
‘I saw myself.’
(ii) Ku-‘itad-e na karanamu
1SG-seesP NOM self=1SG.GEN
‘I saw myself.’

We can therefore draw the following correlations between agreement, case and structural position, on the one hand, and grammatical function and syntactic role on the other.

### Table 3. Coding properties of SUBJECTs in different syntactic roles

<table>
<thead>
<tr>
<th>Syntactic role</th>
<th>Case marking</th>
<th>Verbal agreement</th>
<th>Word order / structural position</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>na</td>
<td>prefix</td>
<td>VP-external/structural position</td>
</tr>
<tr>
<td>S</td>
<td>na</td>
<td>prefix</td>
<td>VP-external</td>
</tr>
<tr>
<td>P</td>
<td>na</td>
<td>enclitic</td>
<td>VP-external</td>
</tr>
</tbody>
</table>

### Table 4. Coding properties of OBJECTs in different syntactic roles

<table>
<thead>
<tr>
<th>Syntactic role</th>
<th>Case marking</th>
<th>Verbal agreement</th>
<th>Word order / structural position</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>te</td>
<td>prefix</td>
<td>VP-external</td>
</tr>
<tr>
<td>S</td>
<td>na</td>
<td>(prefix)</td>
<td>VP-external</td>
</tr>
<tr>
<td>P</td>
<td>te</td>
<td></td>
<td>VP-internal</td>
</tr>
</tbody>
</table>

Objects Ss are only found in passives, as discussed in Donohue (2005b). The implications for these facts, syntactic and morphological, on a discussion of differential subject marking will be presented in the next section.

4. IMPLICATIONS

We have seen that the agreement morphology in Tukang Besi is essentially nominative-accusative, grouping S and A together as opposed to P. At the same time the case marking system tracks the grammatical subject, monitored by the amount of agreement marking on the verb, and correlating roughly with the structural position of the argument. We have, therefore, different marking for subjects in terms of the different inflectional components, agreement and case marking, not aligning together, and in terms of the paradigm used to mark the subject. In a bivalent verb with agreement marked for both A and P the agreement paradigm that marks the subject is the set of P enclitics, and the P nominal is marked with the na case. On the other hand a bivalent verb with agreement marked for the A alone will show agreement with subject by means of the prefixes; the case marking na on the A nominal is consistent in marking subject.

While the agreement markers are not consistent in their marking of subject or object, they are consistent in marking argument structure positions, marking an S or A in the case of the prefixes and a P in the case of the enclitics. The discrepancy arises in the assignment of grammatical functions to argument structure positions,

(For ‘S as object’, see the description of passives in Donohue (2005b). Non-passive Ss do not show any evidence of being objects)

(Some speakers, mainly younger ones, prefer onggo ‘body’ as a reflexive nominal.) See Arka and Manning (1998), or Donohue (1999), for discussion of the significance of these data.)
the principles for which are shown in (92) and which correlate with the appearance of enclitics on the bivalent verb (following suggestions in Manning 1996). This makes clear the difference between argument structure positions and grammatical functions: while there is a relationship, generally that shown in (92a), in all languages, it is possible for the inverse mapping to be found, as in (92b). In this case the P (theme, lowest role) is the subject. This is found locally in Tukang Besi, but is apparently the default mapping found in languages such as Mam (England 1983) and certain Eskimo languages (see Manning 1996 for detailed argumentation of this point).

(92) A-structure / GF-mapping for the pronominal voice alteration

a. No enclitics \( \langle -A, -P \rangle \)
   SUBJ OBJ
b. P enclitics \( \langle -A, -P \rangle \)
   SUBJ OBJ

An approach such as that advocated in Sells (2001), in which the P is raised in a clause with P enclitics on the verb, would more formally capture the generalisations illustrated here, and would agree with the evidence for constituency shown in 2.4. Note that, in Tukang Besi, rather than being covert the pronominal incorporation on the verb posited by Sells is overt, in the form of the P agreement clitics.

The fact that the determination of subject in a bivalent clause depends on pronominal features and their marking means that the complications associated with pronominal arguments in terms of their position on the animacy hierarchy come into play. While we have seen that there are bivalent clause types with agreement for the P and others lacking this agreement, when we deal with local persons it is almost unheard of to not show agreement on the verb. This means that, in a clause with local persons, a local person will always be subject, and if there are two local persons the subject will always be the P.

It is not ungrammatical for an pronominal argument that is indexed on the verb to also be represented by a free pronoun in the same clause; this is extremely rare, but is attested in texts. Such constructions are usually associated with pragmatically marked information, and are so judged as being more natural if the case-marked pronoun appears preverbally, as in (94) and the textual (96) (see the discussion following (78) for the licensing of case in preverbal positions). Nonetheless, examples such as (93) and (95), with postverbal na-marked subjects, are not judged to be ungrammatical, though they are unusual, and contain a sense of contrastive focus.

(93) No-’ita=aku te ana (#’ina iaku)
   3r-see=1sg.p CORE child NOM 1sg
   ‘The child saw me.’

(94) Te iaku no-’ita=aku te ana
   CORE 1sg 3r-see=1sg.p CORE child
   ‘The child saw me.’

(95) #’To-waliako=mo i kampo na ikita
   1pl.r-return=pfv OBL village NOM 1pl

(96) … te iaku habantu ()u-hu=aku te kuli=mo …
   CORE 1sg in.fact 2sg.r-give=1sg.p CORE skin=3GEN
   ‘…in fact you’ve just given me the peel (of the bananas) …’

Sentences such as (97), in which a local person is not indexed on the verb, are often judged to be grammatical, but are at best marginally felicitous, and are not part of any corpus of naturally-occurring speech. This implies a constraint requiring local person Ps to be the subject of their clause; morphologically, local persons should be marked on the verb. This does not apply to all local persons, regardless of syntactic role; (98) shows an instance of a non-nominative local person A; the choice to code loka ‘bananas’ as nominative forces the object coding of the A. The fact that te iko’o ‘you’ is overtly coded as a DP in the clause is mildly surprising, and probably reflects contrastive focus.

(97) *#’No-iita te iaku na ana
   3r-see CORE 1sg NOM child
   ‘The child saw me.’

(98) ‘Oho, ’u-pidi=’e te iko’o na loka saba’ane=’e.”
    yes 2sg.r-rubbish=3p CORE 2sg NOM banana all=3p
    ‘Yes, you’ve rubbish all of the bananas.’

Other evidence for the origins of the grammatical subject of Tukang Besi lying in pragmatic conditions can be found by examining sentences with question words. We find that there is a condition on the case marking (and thus clause type) that may be associated with question words, that can be expressed as in (99).

Constraints on question word case marking

(99) A question word cannot be cast syntactically in a way that would lead to it being assigned nominative case (na) in a verbal clause; that is, a question word may not be the grammatical subject.

The implications of this are easily illustrated. Question words appear in situ, as in the questioned non-nominative P in (100).

(100) No-sampi te paira na amai?
   3r-pick CORE what NOM 3pl
   ‘What did they pick?’

Questioning the A in this clause is only possible if the verb shows P enclitics, that is, casts the A in the non-nominative case, te. It is not possible to question an A that bears the nominative case.20

20 That this is a syntactic, and not simply morphological, restriction can be shown by examining the argument in a preverbal position, commonly used with focussed information. Here we find that the
Cleft questions for As and Ps

(107) Te e mai na s<un>sampi te loka
CORE: who NOM pick,SI CORE banana
‘Who picked the bananas?’ (= ‘Who is it that picked the bananas?’)

(108) Te pai na di-sampi?
CORE: who NOM PP-pick
‘What was picked?’ (= ‘What is it that was picked?’)

The simplest explanation that can account for these restrictions is one that assumes that the pragmatic function FOCUS, with which question words are inherently associated, is incompatible with the specification of the nominative case/subject, which implies that some element of the pragmatic function TOPIC adheres to the grammatical function SUBJECT. Arguments against treating the no-marked element as simply being a topic can be found in Donohue (1999, 2004), and are similarly described for Tagalog in Kroeger (1993).

In short, we have a diathesis that involves the alternation of the grammatical functions subject and object, without demotion. This does not affect the relative positions of the arguments in argument structure, but the pragmatic and syntactic consequences are reflected in the morphosyntax of the language.

Table 5 shows the different agreement strategies used to mark subjects in different constructions. A dash indicates that the syntactic role in question is not eligible to appear in that construction.

Table 5. Differential subject agreement marking

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>A</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>default main clause</td>
<td>prefix</td>
<td>prefix or genitive</td>
<td>P enclitic</td>
</tr>
<tr>
<td>experience</td>
<td>prefix or genitive</td>
<td>prefix</td>
<td>P enclitic</td>
</tr>
<tr>
<td>motion-numeral SVC</td>
<td>prefix + P enclitic</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>existential</td>
<td>(prefix)/(P enclitic)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>passive</td>
<td>(prefix)</td>
<td>-</td>
<td>Ø</td>
</tr>
<tr>
<td>nominalisation</td>
<td>genitive</td>
<td>genitive</td>
<td>genitive</td>
</tr>
<tr>
<td>adverbial temporal</td>
<td>genitive</td>
<td>-</td>
<td>genitive</td>
</tr>
<tr>
<td>P-relative clause</td>
<td>-</td>
<td>genitive</td>
<td>genitive</td>
</tr>
</tbody>
</table>

21 This is similar to the analysis presented of Chichewa agreement in Bresnan and Mehombo (1987).

22 There are cases in which these principles are violated. We have already seen some cases of atypical verbal marking in the optional genitive marking found with experiencer subjects, seen in 2.4, and with the double marking patterns found with motion-numeral serial verb constructions, also seen in 2.4. Additionally, there is no clearly discernible ‘subject in existential, passive and metatopical clauses, and there is no distinction between different arguments in nominalised clauses and clauses that use nominal morphology, such as P-relative clauses and adverbial temporal clauses. See Donohue (1999:75-78, 274-281, 379-385, 411-413, 472-473, 478-480; 2005b) for further details.

23 A secondary P may appear as a core argument in a passive of a trivalent verb (as in ‘they were given the books’), but cannot be marked on the verb with P-enclitic agreement. This contrasts with the closely related Muna (van den Berg 1989), which allows two suffixal agreement positions.

24 The enclitic ‘-m’ is often not marked in the enclitic form, and the enclitic form is not optional, but must be marked.

Less frequently a non-subcategorised participant may be questioned in a cleft, following applicativisation, but more commonly adjuncts are questioned in situ. Questions with in situ question words are shown in (i) - (iii), while grammatical and ungrammatical clefts are shown in (iv) and (v).

(i) No-kede di ‘orma? (ii) No-kede di paina?
3r-sit OBL where 3r-sit OBL what
‘Where did they sit?’ ‘What did they do?’

(iii) No-kede-mi te paina? (iv) *Te pai na di-kede-mi?
3r-sit-LOC APPL CORE what CORE what NOM PP-sit=3GEN
‘What did they do?’ ‘What did they do?’

(v) Te pai na di-kede-mi-no?
CORE what NOM PP-sit-LOC APPL=3GEN
‘What did they do?’
In addition to the major split in subject marking found in Tukang Besi, that between marking a subject that is an S or A, by prefix, or a subject that is P, by enclitic, we have also examined three minor sentence types, those involving experiencer Ss, motion verbs serialised with numeral verbs, and existential clauses, and seen that there are additional patterns of marking. Experiencer verbs may index an S by means of genitive clitics, a marking strategy that dominates the subordinate clause marking. Motion verb + numeral verb serialisations show a double marking pattern in which both S-agreement and P-(like) agreement marking is found. The existential construction is the least fixed, with agreement only optional, and potentially marked by either the prefixes or the enclitics, depending on pragmatic conditions, or by the multifunctional ke(n) ‘comitative /’instrumental / ‘and’). The variability of marking options in the existential construction correlates with the fact that there is no grammatical subject in this construction, and this has been shown in table 5 by the bracketing about the data. The verb in a passive construction, which also lacks a subject, only optionally shows agreement for the S.

Table 6 shows the case marking strategies available to the different kinds of subjects (and Ss).

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>A</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>default main clause</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>experiencer</td>
<td>na</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>motion+numeral SVC</td>
<td>na</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>existential</td>
<td>(na or ke ‘comit’)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>passive</td>
<td>na</td>
<td>-</td>
<td>te ‘core’</td>
</tr>
<tr>
<td>nominalisation</td>
<td>geniüve</td>
<td>genitive</td>
<td>genitive</td>
</tr>
<tr>
<td>adverbial temporal</td>
<td>geniüve</td>
<td>-</td>
<td>geniüve</td>
</tr>
<tr>
<td>P-relative clause</td>
<td>-</td>
<td>genitive</td>
<td>genitive</td>
</tr>
</tbody>
</table>

Case marking is more consistent than agreement marking, with the ‘nominative’ *no* appearing to mark subjects in all main clause types, and the genitive *mu* in all subordinate clause types. This is true regardless of the marking strategy employed, again the existential showing variation, correlating with the lack of subject in this construction. The passive, by contrast, shows the case marking patterns expected of an S subject, even though the S argument in this construction lacks subject privileges.

The final section explores the issues raised in section 4 from a more cross-linguistic perspective.

---

25 A secondary P may appear in a P-relative clause, not heading the relative clause but as a genitive complement of it. An example of this can be found in (144).

5. **MODULES OF GRAMMAR**

Tukang Besi is the south-easternmost language in the Western Malayo-Polynesian branch of Austronesian, and is thus the last language to be (relatively) closely genetically related to the Philippine-type languages, famous for their unusual and controversial subject properties. It is also one of the westernmost languages in Indonesia to show pronominal head-marking, a feature that is widespread in eastern Indonesia, but absent in the west (and in the Philippines). It is worth comparing the morphosyntax of the Philippine-type languages, and of a more familiar language such as English.

In Tagalog a clause with a predicate that selects for two arguments, such as *kita* ‘see’, allows for either of those arguments to be coded as subject, indicated by the use of nominative forms, in these examples the determiner *ang*, and the choice of prefix on the verb. See Kroeger (1993) for explicit arguments about the status of these arguments as subjects, building on discussion in Schachter and many others.

(109) Naka-kita ng aso ang bata
     AV-see GEN dog NOM child
     ‘The child saw a dog.’

(110) Na-kita ng bata ang aso
     PV-see GEN child NOM dog
     ‘The child saw the dog.’

While the alternation in (109) and (110) could be (and has been) analysed as an instance of a ‘familiar’ voice alternation such as that seen between an active and a passive clause, or an active and an antipassive, there are substantial differences, which may be summarised as follows:

- in the Philippine case seen here in Tagalog the amount of morphological marking is equal for both voice choices; no voice selection is morphologically unmarked with respect to the other.
- in the Philippine case the voice alternation does not involve ‘demotion’ of the Agent when the Patient is coded as subject. In both voice selections we have two core arguments, one of which is subject.
- in the Philippine-type voice systems there are voice choices for more arguments than are subcategorised for by the predicate. This can be seen in (109) and (110), which show that a monovalent predicate like *lunsad* ‘alight’ may select a participant that is not subcategorised for by the verb, if the verb is marked with the appropriate morphology, in this case -an ‘dative voice’.

I am simplifying (though not misrepresenting) the facts of Tagalog morphology somewhat dramatically for the purposes of exemplification here, since the argumentation is secondary to the main discussion of this chapter. In particular the degree of lexical specification found in affinal possibilities is not mentioned at all in this chapter; McFarland (1976) and De Guzman (1978) provide discussion on this topic.
by the case marking, thus showing that the functions of the two subject-marking systems are quite distinct.

The fact that there is no dedicated voice-monitoring morphology might come as a surprise to linguists acquainted with the voice systems of western Eurasia. There are, however, analogous voice systems found elsewhere, showing a similar lack of dedicated morphology. Lango (Noonan and Bavin Woock 1978; Foley and Van Valin 1984; Noonan 1992) has a voice system without any dedicated voice morphology, as shown in (115) and (116). Here there is consistent agreement on the verb for both arguments, but positional variation correlates with subject status.

Active subject object
(115) Dákó ɗ-jwát-ɗ lócà woman 3SG-hit-3SG man
'The woman hit the man.'

Passive subject object
(116) Lócà dákó ɗ-jwát-ɗ man woman 3SG-hit-3SG
'The man was hit by the woman.'

In Palu'e and Manggarai, both languages from Flores, in southern Indonesia (Donohue 2005a; Arka and Kosmas 2005), a voice alternation operates with no verbal morphology. In Manggarai there is a change in the VP-final agreement clitic that marks subject, and prepositional marking of the Agent, in the passive construction, shown in (118). The active construction in (117) shows a VP that agrees with the Agent as subject.

Active subject object
(117) Aku [vp cero latung] k
1SG fry corn=1SG
'I fry/am frying corn.'

Passive subject oblique
(118) Latung hitu [vp cero l=aku t=] corn that fry by=1SG=3SG
'The corn is (being) fried by me.'

In Palu'e, as in Lango, word order alone codes the difference in grammatical functions. There is no change in the marking (or absence of marking) on the verb, and no change in the use of bare NPs.

Active subject object
(119) Ata wai lie ata laki person woman see person man
'The woman saw the man.'

As with Tukang Besi, the case marking option ('nominative') is the same for all subjects, but the choice of morphology used on the verb to index that argument varies depending on the semantic and syntactic role of that argument. While the Tagalog voice morphemes are reflected in Tukang Besi, they are not found in main clause uses, being relegated to subordinate functions. In their place, the pronominal agreement system has been adopted, showing an (eastern) Indonesian tendency towards head marking replacing the older Austronesian system.27 This is not such a radical departure from the earlier; recall that Sells (2001) analysed the voice system in Tagalog as involving (abstract) pronominal elements.28 In Tukang Besi we have simply seen the replacement of these abstract pronominal elements with overt ones in main clauses. In doing so we have ended up with a system in which the paradigm used for marking subject on the verbs varies with the voice of the clause. The syntactic role, S and A on the one hand or P on the other, is monitored consistently by the agreement marking, while the grammatical function is indicated consistently

---

27 See Donohue (2002) for discussion. Of the morphemes seen here, Tagalog <um> ‘actor voice’ is cognate with Tukang Besi <um> S-(and A)- infix (in subordinate clauses), and Tagalog -an ‘dative voice’ is cognate with Tukang Besi -a ‘nominative’.

28 A development of this analysis, which treats the ‘voice morphology’ on the verb as pronominal, is currently being pursued by the present author, in collaboration with Alastair Butler.
Passive subject oblique
(120) A ta laki a ta wai lie
person man person woman see
"The man was seen by the woman."

The Manggarai case, in which there is at least marking for non-core status of the Agent in the passive clause, has a direct parallel in the non-literate passive constructions of many Chinese languages. The Hokkien examples in (121) and (122) illustrate the active-passive alternation, showing the use of the multifunctional morpheme *ho to mark the passive. (123) shows that the agent in (122) is an oblique, not an adjunct, since it is not optional in the passive construction.

Active subject object
(121) I phah liau hit-e lang
3SG hit PERF that-MOD person
"He hit that person."

Passive subject oblique
(122) Hit-e lang ho i phah liau
that-MOD person NO3SG hit PERF
"He hit that person."

Comparing with English, or other languages of the western European model, we are struck by the relative autonomy of the argument-structure determined syntactic roles and the case marking, and this is due to the fact that the voice system of Tukang Besi, and other Western Malayo-Polynesian languages, is non-demotional, as described in section 4. This means that a bivalent clause remains bivalent, regardless of which voice, A-as-subject or P-as-subject (which could just as well be labelled 'direct' and 'inverse'), is selected. This has the consequence of bivalent clauses with subject Ps having a second core argument, the A, which must be thought of as the object of the clause. Since the verbal agreement for this argument is with the same set of prefixes that are used for agreement with Ss and As in all clause types, we can see that, essentially, the same agreement prefixes can be used to mark either a subject or an object, depending on the clause type. By contrast the P enclitics are only used to mark a subject on the verb, since they are not present in a clause that does not have a subject P. In terms of case marking, no is essentially used to mark subjects, and te for objects (with the existential and passive constructions as exceptional cases).

In contradistinction to the materials presented in Aissen (1999), we see that Tukang Besi (as well as many other western Austronesian languages) does not have a high-ranked constraint of the form *Obl|AGR, and that this non-subject *AGR (⇒ A) is case-marked identically to an object *Pat (⇒ P) (taking 'object' to be roughly equivalent to 'non-subject term'). I suggest that the extreme regularity of subject marking in languages such as English, languages which have played the major part in informing modern syntactic theory, is to a large part an artefact of the absence of object As. In a language with a demoting voice system, in which the 'Agrte' is coded as oblique and the 'Pat' is the single core argument of the predicate, there is no mechanism for agreement with the 'Agt'.

Since the 'highest role' feature, the feature that defines S and A as a group, is relevant to the description of all languages in at least some constructions it follows that the Agrte might retain marking as a (core) category regardless of the voice choice, if the voice is non-demotional. This is especially so if the language has a grammatical subject that is closely related to the pragmatic notion of topic, a characteristic that applies to Tukang Besi, as seen in section 4 (and is true, to a greater or lesser extent, in most Austronesian languages, certainly in those west of New Guinea). Unlike a more syntactically-centred grammatical function, one that can be defined in terms of the syntactic roles A, S and P, the Tukang Besi subject is determined pragmatically. This means that syntax does not play a significant role in allowing, through the constructions that are restricted on the basis of grammatical functions, determination of a relationship between the grammatical function subject and the syntactic roles A and P in a bivalent clause, and thus semantic notions such as actor / proto-agent / 'Agrt', and undergoer / proto-patient / 'Pat' cannot be automatically assigned to a particular grammatical function just on the basis of their relative position in argument structure. In a language such as English knowledge of the identity of the subject will automatically give information about the identity of the most agentive core argument in the predicate in Tukang Besi this 'agent'-identifying (or at least 'agent'-confirming) side-effect of the well-constructed subject argument is absent, and so some 'freeing up' of the degree to which subject marking on the verb is performed by a unique paradigmatic set can be expected, since we require more overt morphological means to identify the syntactic grouping that cannot, in this language, be easily identified through grammatical function identity.

In short, the unusual grammatical function selection system of Tukang Besi, with its lack of demotion in the pronominal voice system, sets up the conditions necessary for the differential subject marking system in which a given pronominal agreement paradigm, the S,A prefixes, retain constant reference to the argument they mark, while indicating different grammatical functions (subject or object) depending on the voice construction, direct or inverse, in which they occur.

REFERENCES

29 Constructions such as control structures with verbs like 'wont', and (to a lesser extent) textual cohesion in narratives, show S,A pivots in (nearly) all languages, and so cannot be taken as evidence for the grammatical function 'subject'. For further discussion, see (amongst others) Manning (1996) and Falk (2000a).
Mark Donohue


