

# *Voice in two northern Sulawesi languages*

---

NIKOLAUS P. HIMMELMANN

## 1 Introduction

In this paper, the basic clause structure and some aspects of the verbal morphology of two northern Sulawesi languages, Ratahan and Lauje, are compared with each other and with Tagalog, the language chosen to represent a prototypical Philippine language (which, in fact, it is only with some qualifications). The paper is intended to contribute to the typology of voice-related phenomena in western Austronesian languages by providing primarily descriptive information on two languages for which such information has so far not been available. It should be noted, however, that the grammatical analysis of the two languages is still in its initial stages. Hence, it is almost certain that the information provided here is incomplete, if not erroneous in some respects.

Apart from providing descriptive data, the paper also aims to make some headway in developing a typological grid for western Austronesian languages. Thus, the comparison between the three languages is not limited to the features which come readily to mind when comparing voice-related phenomena in these languages, in particular the basic clause structure (§2), the voice morphology (§3), the paradigms for noun phrase markers (§4) and pronouns (§5), and pronominal prefixes (§6). Instead, it also comprises phenomena which to date have received little attention but may prove to be of major importance for a more elaborate and fine-grained grammatical typology of western Austronesian languages. Among such phenomena are applicative suffixes (§7), the use of voice-marked forms in non-predicative functions, in particular in construction with quantifiers (§8), and voice in stative derivations (§9).

All of these phenomena are further defined and exemplified in the following sections. Their potential typological relevance is discussed in §10, which also provides a summary of the comparison between Tagalog, Ratahan and Lauje.

Ratahan belongs to the Sangiric group of languages and is spoken southwest of the Minahasan area in the province of North Sulawesi. The discussion is based on fieldwork data collected in collaboration with John Wolff in 1996. Lauje belongs to the Tomini-Tolitoli group of languages and is spoken in northern Central Sulawesi. Lauje examples are from my fieldwork data collected in 1988 and 1993. See Himmelmann and Wolff (in prep.) and

Himmelmann (2001) for more background information and full acknowledgments. Tagalog examples are taken from Bloomfield (1917) or my own recordings.

A conscious (and at times quite time consuming) effort has been made to illustrate all grammatical phenomena with examples drawn from realistic spontaneous speech (conversations or narratives). These examples are presented in intonation units, each unit being delimited by a pair of curly brackets (numbers in curly brackets show pause length in seconds, though pause length has not been measured in all instances). In the few instances in which elicited data are given, the available spontaneous data would have required lengthy comments in order to serve as evidence for a given observation.

## 2 Basic clause structure

In Ratahan, in both transitive as well as intransitive clauses, there is always one core argument (the single core argument in the case of intransitive verbs) which is characterised by three features not shared by any other core argument: (a) its position is variable, i.e. it may occur in pre- or post-verbal position; (b) its semantic role is marked by the voice affix on the verb; (c) it is never marked by any kind of phrase marker.<sup>1</sup> This argument is called the *subject* of a verbal clause. The following two intransitive clauses exemplify these characteristics:<sup>2</sup>

- (1) Ratahan uc\_tw 121  
*hairen araq kumukuk e manuk* {}  
 hairen araq um-kukuk Ce manuk  
 later if ACT-cry\_out COMPL chicken  
 'later when the chicken cried out'
- (2) Ratahan uf2\_tw 127  
*araq isé kumukuk e* {1.2}  
 araq isé um-kukuk Ce  
 if 3SG ACT-cry\_out COMPL  
 'when he would cry out'

The following two examples, in which the verb is marked for undergoer voice, illustrate the same characteristics for transitive clauses:

- (3) Ratahan uc\_ew 217  
*naq sawu ntoo Sepus winuno mangasé* {}  
 naq sawu N-to Sepus in-wuno mangasé  
 ?? time LK-NR Joseph REAL(UG)-kill 3PL  
 'at the time they killed Joseph'

<sup>1</sup> Proper nouns occur with the proper noun article in this position (see §4).

<sup>2</sup> Abbreviations used: ACT actor; AND andative; APPL applicative; CAUS causative; COMPL completive; DAT dative; DET determiner; DIST distal; EX exclusive; GEN genitive; GER gerund; IMP imperative; IRR irrealis; LK linker; LOC locative; NR nominaliser; PL plural; PM predicate marker; PN proper noun; PROX proximal; REAL realis; RED reduplication; RPRT reportative; SG singular; SPEC specific (article); STAT stative; UG undergoer; UG.L undergoer.locative; UG.T undergoer.theme.

- (4) Ratahan uc\_tw 583  
*kinambei ne yaq* { }  
 in-kambei ne yaq  
 REAL(UG)-embrace 3SG.GEN 1SG  
 'he embraced me'

In transitive clauses, the non-subject core argument always follows immediately after the verb. If it is a common noun, it may optionally be marked with the genitive marker *nu* (for proper nouns the marker is *ni*). In examples (3) and (4) the non-subject argument is a pronoun (*mangasé* and *ne*, respectively). Examples with full lexical NPs for both actor and undergoer voices:

- (5) Ratahan uf2\_tw 113  
*mamake nu babuq mangasé* {0.9}  
 maN-pake nu babuq mangasé  
 ACT-use GEN slave 3PL  
 'they had (female) slaves'
- (6) Ratahan et\_up 013  
*niraes e nu waa wusak.*  
 ni-raes Ce nu waa wusak  
 REAL(UG)-finish COMPL GEN monkey banana  
 'The monkey had eaten all the bananas.'

For Lauje essentially the same observations hold. There is, however, one major distinction with regard to the use of NP-markers with core arguments, as discussed in §4.

In both Ratahan and Lauje it is very common for core arguments to follow the verb, which is also the most common order in Tagalog. But there is a crucial difference among the three languages with respect to examples in which the subject precedes the verb. Compare the Ratahan examples (2) and (3) above and the following example from Lauje:

- (7) Lauje UF\_LM 150  
*bangkola' ngintugome galaa.* {0.4}  
 bangkola' N-intug-O-me galaa  
 monkey REAL.ACT-look\_for-O-COMPL turtle  
 'The monkey looked for the turtle.'

There is no element corresponding to the Tagalog inversion marker *ay* in these examples. There is also no prosodic break between the clause-initial subject and the following verb. Thus these examples show that SV(X) is indeed an alternative basic word order in Ratahan and Lauje.

The examples of SV(X) order should be strictly distinguished from topicalised constituents, i.e. constituents which are separated from the following segment by an intonational boundary, as in:<sup>3</sup>

<sup>3</sup> Capital letters in Lauje examples (primarily <E> but also <A>) represent paragogic vowels. These vowels occur fairly regularly at the end of phonological words (if consonant-final) but may also have some morphosyntactic function. See Himmelmann (1997) for further discussion.

- (8) Lauje UF\_LD 029f  
*too ntoilapi injéine* {}  
 too N-toilapi injéine  
 person LK-married\_couple PROX  
  
*na'agaad E unga* {}  
 no-'o-gaad E unga  
 STAT.REAL-??-take E child  
 'this couple, (they) had a child.'

Such topicalised constituents are probably possible in all western Austronesian languages, with or without special markers such as Tagalog *ay*. Note that even in Tagalog the use of *ay* is not obligatory:

- (9) Tagalog donat 288  
*yung Eta namán* {}  
 iyón:LK Ayta namán  
 DIST:LK Ayta truly  
  
*takot na takot sa iyó* {}  
 takot na takot sa iyo  
 fear LK fear LOC 2SG.DAT  
 'those Ayta, (they) were very afraid of you'

A prosodic break is sufficient to separate a topicalised constituent from the remainder of the clause. 'True' SV(X) order, on the other hand, is impossible in Tagalog and many other Philippine-type languages.

### 3 Basic voice paradigms

In Ratahan, as in Tagalog, the basic voice paradigm in declarative main clauses consists of four voices and two moods. There is, however, a conspicuous difference with regard to the morphological formatives used to express the voice/mood distinctions. While Tagalog has special formatives for each of the three undergoer voices (i.e. the suffixes *-in* and *-an* and the prefix *i-*), Ratahan has only one formative which exclusively occurs in the undergoer voices, i.e. the suffix *-an*. Compare Tables 1 and 2.

**Table 1:** Basic voice-mood paradigm in Ratahan

	+REALIS	-REALIS
ACTOR	<i>-im-/N-</i>	<i>-um-/M-</i>
UNDERGOER	<i>-in-</i>	<i>-an</i>
UNDERGOER.LOCATIVE	<i>-in--an</i>	<i>-an</i>
UNDERGOER.THEME	<i>-in-</i>	$\emptyset$

Table 2: Basic voice-mood paradigm in Tagalog

	+REALIS	-REALIS
ACTOR	-um-/N-	-um-/M-
UNDERGOER	-in-	-in
UNDERGOER.LOCATIVE	-in--an	-an
UNDERGOER.THEME	i--in-	i-

The three undergoer voices in Ratahan can only be distinguished based on the paradigmatic interaction with the realis-marking formative *-in-*. Thus, in undergoer voice there is no suffix in realis contexts (see (3) above) but the suffix *-an* is used in non-realistic contexts:

- (10) Ratahan uc\_tw 762  
*kau wunón masé { }*  
 kau wuno-an masé  
 2.SG kill-UG 3PL  
 'they will kill you, no doubt about it'

In locative undergoer voice, on the other hand, the suffix *-an* occurs in both realistic and non-realistic contexts:

- (11) Ratahan et\_ww G16a  
*laliwuqan le yaq mangewong buk tiqí!*  
 laliwuq-an le yaq maN-ewong buk tiqí  
 help-UG.L IMP 1SG ACT-carry book PROX  
 'Help me to carry these books!'

- (12) Ratahan et\_ww G16b  
*isé nilaliwuqan ku.*  
 isé ni-laliwuq-an ku  
 3SG REAL(UG)-help-UG.L 1SG.GEN  
 'I helped him.'

In thematic undergoer voice, there is no formative for voice. In realistic contexts, the verb is marked for realistic mood by the infix *-in-*:

- (13) Ratahan et\_ww G2  
*sabel nu winei ku si tuari nu*  
 sable nu in-wei ku si tuari nu  
 machete 2SG.GEN REAL(UG)-give 1SG.GEN PN.LOC younger\_sibling 2SG.GEN  
 'I gave your machete to your younger sibling'

In non-realistic contexts, the verb remains unmarked, i.e. the non-realistic form of the thematic undergoer voice is identical with the basic form of the verb:

- (14) Ratahan et\_up 067  
*wias té wei sisé taawi.*  
 wias teqé wei si =isé taawi  
 rice DIST give PN.LOC=3SG tomorrow  
 'That rice will be given to him tomorrow.'

In Ratahan, then, voice and mood marking are even more intricately intertwined than in Tagalog and other Philippine-type languages. The formal differences between Tagalog and Ratahan are paralleled by semantic ones in that in Ratahan many more verbal roots occur with thematic undergoer voice than in Tagalog. In fact, the thematic undergoer voice in Ratahan appears to be well on its way to becoming not only the formally but also the semantically unmarked undergoer voice (see Himmelmann & Wolff in prep. for more discussion).

The voice-mood paradigm found in Lauje is very different from the one found in Ratahan and Tagalog. To begin with, there are only three basic voices in the Lauje paradigm, as shown by Table 3.

**Table 3:** Basic voice-mood paradigm in Lauje

	+REALIS	-REALIS
ACTOR	<i>N-/(-um-)</i>	<i>M-/(-um-)</i>
UNDERGOER	<i>-in-</i>	<i>no-</i>
UNDERGOER.LOCATIVE	<i>-in--ang</i>	<i>no--i</i>

Furthermore, locative undergoer voice is marked by two different suffixes, i.e. *-ang* in realis mood and *-i* in non-realistic mood:

- (15) Lauje UF\_LM 139  
*tinambunangonyome*                      *nu*   *petu*   *batanganonye* {0.9}  
 in-tambung-ang-O-nye-me              *nu*   *petu*   *batangan-O-nye*  
 REAL(UG)-pile-UG.L-O-3SG.GEN-COMPL GEN soil body-O-3SG.GEN  
 'she covered her body with soil'
- (16) Lauje UF\_LD 264  
*inyaa nrape'i*                      *a'e* { }  
 inyaa no-rape'-i                      *a'e*  
 don't IRR(UG)-close\_by-UGL 1SG  
 'don't get closer to me'

Finally, possibly the most remarkable feature of the Lauje voice paradigm is the prefix *no-* which occurs in the non-realistic forms of the undergoer voices (with the exception of the first person singular, see §6). This feature is remarkable on two counts. First, in typical Philippine-type languages such as Tagalog non-realistic mood is always morphologically unmarked. Second, while a prefix marking non-realistic mood in undergoer voice occurs in several Sulawesi languages (see van den Berg 1996), Lauje and its closest Tomini relatives (Dondo, Tialo and Ampibabo) are, to my knowledge, unique with regard to the segmental shape of this prefix. In southern Tomini languages (Tajio, Taje, Dampelas and Pendau) the functionally equivalent prefix has the shape *ro-* or *ho-* (which in some of the languages undergoes vowel harmonic alternations) and in Kaili-Pamona languages it generally has the shape *ra-*. Apart from its uniqueness, the Lauje form *no-* is also highly conspicuous and somewhat confusing in that *n-*initial formatives in western Austronesian languages generally signal realistic mood (see for example, Tagalog *nag-*, *na-*, *naka-*, *naki-* etc.).

#### 4 NP markers

As in Tagalog, noun phrase markers in Ratahan and Lauje come in two sets, one for common nouns and one for proper nouns. The following discussion pertains primarily to common noun phrase markers. Table 4 lists the forms. Optional markers appear in parentheses.

Table 4: NP-markers

	TAGALOG		RATAHAN		LAUJE	
	common noun	proper noun	common noun	proper noun	common noun	proper noun
ARTICLE	<i>ang</i>	<i>si</i>	–	<i>i</i>	( <i>nu</i> <sub>1</sub> )	<i>si</i>
GEN	<i>ng</i>	<i>ni</i>	( <i>nu</i> )	<i>ni</i>	<i>nu</i> <sub>2</sub>	<i>ni</i>
LOC	<i>sa</i>	<i>kay</i>	<i>su</i>	<i>si</i>	<i>li</i>	–

The most obvious difference between Tagalog and the two Sulawesi languages pertains to the fact that in the Sulawesi languages there is no element corresponding to the specific article *ang* for common nouns. Consequently, common noun phrase subjects in both Ratahan and Lauje may occur without a noun phrase marker, as amply illustrated by the examples in the previous and following sections.

In Ratahan, furthermore, the non-subject argument in both actor- and undergoer-oriented constructions may also remain unmarked. Compare the following two examples with examples (5) and (6) above:

- (17) Ratahan tw 009  
*ngo: yaq mangewong e sinapang { }*  
 ngo yaq maN-qewong Ce sinapang  
 like 1SG ACT-carry COMPL weapon  
 'I'll take the gun'
- (18) Ratahan uc\_tw 022  
*nilaweqan burung yaq {1.3}*  
 ni-laweq-an burung yaq  
 REAL(UG)-lie\_in\_wait\_for-LOC bird 1SG  
 'a bird lay in wait for me'

In this regard, Ratahan differs sharply from Tagalog where the non-subject arguments in these constructions must be marked as genitives.

However, the differences between Ratahan and Tagalog are to a certain degree 'superficial' in that they only pertain to the overt marking of core arguments. The basic distinctions made within the noun phrase marking systems are the same: subjects exhibit coding properties different from all other core arguments; actors in undergoer-oriented constructions are coded as genitives; undergoers in actor-oriented constructions are also coded as genitives (if the undergoer is a proper noun it is marked as oblique); recipients and addressees are coded as obliques in both actor- and undergoer-oriented constructions.

The use of the Lauje noun phrase markers, on the other hand, differs quite distinctly from the one found in Tagalog and Ratahan, despite the fact that essentially the same labels may be used in describing them and that the Lauje formatives are cognate with the Ratahan ones, with the exception of the oblique marker *li*. There are two major differences: first, undergoers

in actor-oriented constructions are not marked as genitives (by the genitive marker *nu*<sub>2</sub>) but rather receive the same marking as subjects (i.e. they may optionally be marked by *nu*<sub>1</sub>, an article-like element). The difference between the two markers of the shape *nu* is further discussed below. Second, recipients and addressees in ditransitive constructions are not marked as obliques but generally remain unmarked.

Evidence for the second difference is relatively straightforward. In Ratahan (and in Tagalog), the additional core argument in ditransitive constructions is marked by the general purpose preposition *su* (unless it is made the subject of the construction):

- (19) Ratahan et\_ww G2  
*tali nu winei ku su anak nu.*  
 tali nu in-wei ku su anak nu  
 rope 2SG.GEN REAL(UG)-give 1SG.GEN LOC child 2SG.GEN  
 'I gave your rope to your child'

In Lauje, it remains unmarked:

- (20) Lauje et\_mt\_g 002  
*'alolongome binee'e unga'e.*  
 'alolong-O-Vme in-bee-'u unga-'u  
 rope-O-2SG.GEN REAL(UG)-give-1SG.GEN child-1SG.GEN  
 'I gave your rope to my child'

Whether other morphosyntactic differences correlate with this difference in noun phrase marking requires further investigation.

The first major difference in noun phrase marking between Ratahan and Lauje mentioned above is in need of a somewhat more elaborate discussion. To begin with, note that the two markers of the shape *nu* are not only distinguished with regard to their use but also differ on a purely morphological level.<sup>4</sup> Thus, genitive *nu* is regularly reduced to *u* after consonant-final words, as in:

- (21) Lauje PF\_DJ 022  
*li lalong u ogo*  
 li lalong nu ogo  
 LOC inside GEN water  
 'in the water'

No such regular alternation occurs with article-like *nu*. Furthermore, there is a special emphatic form '*onu* for article-like *nu* which is found most commonly in clause-initial position:

- (22) Lauje UF\_LM 020  
*'onu bangkola' pogombo'onya nu (0.4) tatambuang E*  
 'onu bangkola' po-gombo'-O-nye:A nu tatambuang E  
 DET monkey GER-word-O-3SG.GEN:A DET bee E  
 'as for the monkeys, they talked to the bees'

No such emphatic form exists for genitive *nu*.

<sup>4</sup> The following account is simplified in that it ignores several variations which occur in spontaneous discourse and often render the two markers indistinguishable. See Yoshimura (in prep.) for a detailed account of essentially the same phenomenon in Tialo, Lauje's northern neighbour.



Article-like *nu* may be used to mark subjects (and topicalised constituents), as in:

- (23) Lauje UF\_LM 122  
*'uminangma nu:: galaa* {0.6}  
 um-'inang-O-me:A nu galaa  
 ACT-eat-O-COMPL:A DET turtle  
 'the turtle already started eating'
- (24) Lauje UF\_LM 059  
*ame goong A nu tatambuang E* {0.3}  
 ame goong A nu tatambuang E  
 like gong A DET bee E  
 'the bees were (arranged) like a gong.'

This use of article-like *nu* is optional,<sup>5</sup> as shown by (7) above and:

- (25) Lauje PF\_DJ 017  
*nomeelu A tuping: { }*  
 noN-peelu A tuping  
 REAL.ACT-inform A eel  
 'the eel said.'

Article-like *nu* is used not only with subjects but also with undergoers in actor-oriented constructions, as in:

- (26) Lauje et\_kl\_e 006  
*tuá'ie nonobong E nu kopi*  
 tuai-'u noN-tobong E nu kopi  
 younger\_sibling-1SG.GEN REAL.ACT-cut E DET coffee  
 'My younger sibling chopped down the coffee tree (that I had planted).'

Again, this usage is optional and, in fact, rather infrequent in spontaneous discourse. Typically, undergoers in actor-oriented constructions remain unmarked, as in (7) above and:

- (27) Lauje UF\_LM 078  
*li'ó mogutu manisanonya o'e*  
 li'oe M-po-gutu manisan-O-nye:A o'e  
 you ACT-??-make honey-O-3.SG.GEN:A DIST  
 'You make that honey'

Actors in undergoer-oriented constructions, on the other hand, are marked as genitives:

- (28) Lauje UF\_LM 089  
*'ininang u pani'i goot E* {0.5}  
 in-'inang nu pani'i goot E  
 REAL(UG)-eat GEN bat many E  
 '(these bananas) had been eaten by many bats.'

<sup>5</sup> The fact that article-like *nu* is optional is the major reason for calling it 'article-like' rather than an article. I take it as a defining characteristic of definite and specific articles that they have to be used whenever a nominal expression is definite or specific. The overall distribution of article-like *nu* is very patchy, with some speakers using it fairly consistently for marking subjects, but others hardly using it at all.

To sum up the presentation so far, there are two important morphosyntactic differences between Lauje and Ratahan which are related to article-like *nu*. First, common noun phrase subjects in Ratahan are never accompanied by a noun phrase marker while common noun phrase subjects in Lauje may optionally be preceded by article-like *nu*. Second, and more importantly, common noun phrase undergoers in actor-oriented constructions may optionally be marked as genitives in Ratahan, but in Lauje they may optionally receive the same kind of marking as subjects, i.e. article-like *nu*.

Note that this difference in noun phrase marking is not mirrored by differences with regard to other aspects of the morphosyntax of core arguments in the two languages. As mentioned in §2 above, subjects in both Lauje and Ratahan are characterised by the fact that their position with respect to the verb is variable and that their semantic role is marked by the voice affix on the verb. Non-subject core arguments in both languages, including undergoers in actor-oriented constructions, have to occur immediately after the verb.

Still, the different noun phrase marking in Lauje is highly significant in that it leads to a change in the markedness relation between actor- and undergoer-oriented constructions. In Ratahan (and in Tagalog) actor- and undergoer-oriented constructions are almost perfect mirror-images of each other, as indicated by the following structural patterns:

- (29) actor-oriented construction:      NP<sub>ACT</sub> ACT-Verb      GEN-NP<sub>UG</sub>  
 undergoer-oriented construction:      NP<sub>UG</sub> UG-Verb      GEN-NP<sub>ACT</sub>

There is no direct evidence for deriving one construction from the other. (Note that this state of affairs only holds for clauses in which the non-subject argument is a common noun. If the non-subject argument in an actor-oriented construction is a pronoun or a proper noun, it is marked as oblique rather than as genitive. Pronouns are further discussed in the following section.)

For Lauje, on the other hand, it may be argued that the actor-oriented construction is slightly less marked than the undergoer-oriented construction (in terms of morphosyntactic marking). In any event, the two constructions are no longer perfect mirror-images of each other:

- (30) actor-oriented construction:      (DET=)NP<sub>ACT</sub>      ACT-Verb      (DET=)NP<sub>UG</sub>  
 undergoer-oriented construction:      (DET=)NP<sub>UG</sub>      UG-Verb      GEN-NP<sub>ACT</sub>

That is, in Lauje the morphosyntactic properties of the two major types of non-subject core arguments (undergoers in actor-oriented constructions and actors in undergoer-oriented constructions) are not identical. They share the fixed post-verbal position, but they differ with regard to noun phrase marking (in a rather subtle way, since the marker in both cases has the basic shape *nu*).

## 5 Pronoun sets

Compared to Tagalog, there are fewer pronoun sets in Ratahan and Lauje, namely two instead of three (for Ratahan plural pronouns there is, in fact, only one set). The distribution of one of the two sets is adequately covered by the label *genitive* (it is used for actors in undergoer-oriented constructions and for possessors). The other set appears in a broad variety of syntactic functions. Since the use of any case label such as *absolute* or *nominative* as a label for this set is prone to lead to misconceptions, the rather nondescript label *free* is chosen here to refer to this set.

**Table 5:** Ratahan pronouns

	SG		PL
	FREE	GEN	FREE
1INCL	<i>yaq</i>	<i>ku</i>	(i) <i>kite</i>
EXCL			(i) <i>kami</i>
2	(i) <i>kau</i>	<i>nu</i>	(i) <i>kumú</i>
3	<i>isé</i>	<i>ne</i>	(i) <i>mangasé</i>

**Table 6:** Lauje pronouns

	SG		PL	
	FREE	GEN	FREE	GEN
1INCL	(li)a'e	-u	(li)'ite	-te
EXCL			(li)'ame	mame
2	(li)'oe	-Vme	(li)'eme	mie
3	(li)io	-nye	jimo	(nu jimo)

The differences between Tagalog and the two Sulawesi languages are not restricted to the inventory of forms. A much more important difference pertains to the fact that Ratahan and Lauje pronouns, unlike Tagalog subject and genitive pronouns, are *not* second-position clitics. The genitive pronouns in Ratahan and Lauje are either suffixes or enclitics, occurring always immediately after their heads (the verb when they are used to express actors in undergoer-oriented constructions, the possessed noun when they express possessors).

Free pronouns in Ratahan and Lauje occur in almost exactly the same positions as full lexical NPs. In particular, they may be used in preverbal position (see, for example, (2) and (49)) and they follow the non-subject argument in postverbal position:

- (31) Ratahan uc\_tw 022  
*nilaweqan*                      *burung yaq* {1.3}  
 ni-laweq-an                      burung yaq  
 REAL(UG)-intercept-UG.L bird 1SG  
 'a bird lay in wait for me'
- (32) Lauje UF\_LD 179  
*nongkoni*                      *bisa io juniaa ine* {}  
 noN-'oni                      bisa io juniaa ine  
 REAL.ACT-carry poison 3SG world PROX  
 'she poisoned this world (lit. she brought poison to this world)'

In Lauje, free pronouns may also be used as non-subject arguments in actor-oriented constructions:

- (33) Lauje PF\_DJ 038  
*li'ó bela jeiang E* {} *monyampuat a'e baoanomá'* {}  
 li'oe bela jeiang E moN-sampuat a'e báoan-O-ma'e  
 youbuddy friend E ACT-save 1SG land-O-AND  
 'you my friend, will rescue me by getting me ashore.'

In Ratahan, however, free pronouns cannot be used as non-subject arguments in actor-oriented constructions. Instead, free pronouns must be marked with the locative proper noun marker *si* when referring to the non-subject argument of an actor-oriented construction:

- (34) Ratahan et\_ww N22  
*apa to namatakuq si kau?*  
*apa to naN-pa-takuq si kau*  
 what NR REAL.ACT-CAUS-afraid PN.LOC 2SG  
 'what frightened you?'

- (35) Ratahan uc\_tw 077  
*nunongkoliwu si yaq {}*  
*nu-nongko-liwu si yaq*  
 REAL.ACT-?-crowd\_around PN.LOC 1SG  
 '(they) crowded around me'

Note that this is the only way in which pronominal undergoers in actor-oriented constructions may be expressed in Ratahan. In this regard, pronominal undergoers differ from common noun phrase undergoers. As mentioned in the preceding section, if the undergoer in an actor-oriented construction is a common noun phrase, it may optionally be marked as genitive. But genitive pronouns in Ratahan can never be used to express the undergoer of an actor-oriented construction.

## 6 Pronominal prefix

In addition to the pronouns shown in Table 6, Lauje has a single pronominal prefix 'u- for first person singular actors in transitive events. This prefix only occurs in the non-realis undergoer voices:

- (36) Lauje PF\_DJ 072  
*moopus E 'u'inang {}*  
*mo-opus E 'u-'inang*  
 STAT-done E 1SG.ACT-eat  
 'since I will devour it completely'
- (37) Lauje UF\_LM 187  
*'udendenima ine {0.5}*  
*'u-dendeng-i-me:A ine*  
 1SG.ACT-hit-UG.L-COMPL:A PROX  
 'I will bang it.'

This is the standard form for non-realis undergoer-oriented constructions involving a 1SG actor. That is, there are no derivations involving a prefix *no-* and a suffix *-u*. Furthermore, 'u- is in complementary distribution with the prefix *no-* which occurs in all the other non-realis forms of the undergoer voices. Thus, strictly speaking, 'u- is also part of the basic voice-mood paradigm in Lauje and should be added as an alternative prefix to the two cells containing *no-* in Table 3.



- (40) Tagalog Bloomfield (1917:30/14)  
*at ang pare at siyá ay nag-hinúy*  
 and SPEC priest and 3SG PM REAL.ACT-wait  
*ng sà-sabih-in ng sundalo.*  
 GEN RED1-statement-UG GEN soldier  
 'and the priest and he waited for what the soldier would say.'
- (41) Tagalog Bloomfield (1917:72/6)  
*na-kà-tanaw siyá ng bahay na ma-ilaw*  
 REAL.STAT-??-in\_sight 3SG GEN house LK STAT-light  
*sa p-in-a-tù-tunguh-an ng kalabáw*  
 LOC ??-REAL(UG)-RED1-direction-UG.L GEN water\_buffalo  
 'he saw a lighted house in the direction toward which the caribou was going.'

They may also be used in construction with the existential quantifier *may* (and other quantifiers):

- (42) Tagalog pep 094  
*ay mayroon palang nagàalaga {0.1}*  
 ay may-doón palá-ng nag-RED1-alaga'  
 PM EXIST-DIST.LOC so!-LK REAL.ACT-RED1-cared\_for  
*doón sa ibun {1.3}*  
 doón sa ibon  
 DIST.LOC LOC bird  
 'there was already someone looking after those birds.'
- (43) Tagalog ahas 096  
*may ipàpakíta ako sa iyo {0.6}*  
 may i-RED1-pa-kita akó sa iyo  
 EXIST UG.T-RED1-CAUS-visible 1SG LOC 2SG.DAT  
 'Come here, I have something to show you.'

In Ratahan and Lauje, voice-marked forms are used primarily in predicative function. They must not be used in construction with the existential quantifier (i.e. there are no examples parallel to the Tagalog examples (42) and (43)). They may be used in argument positions of verbal predicates but such use requires nominalisation by a particle which also functions as a relative clause marker (in Ratahan this is the particle *to*, in Lauje the particle *sau*):

- (44) Ratahan uc\_tw 415  
*waktu itu siningsapen ku e to: (0.3)*  
 waktu itu in-singsap-an ku Ce to  
 time DIST REAL(UG)-answer-UG.L ISG.GEN COMPL NR  
*pinakinak nge si yaq { }*  
 in-pakinak ne si yaq  
 REAL(UG)-ask 3SG.GEN PN.LOC 1SG  
 'at the time I answered what he asked me'

- (45) Ratahan uf\_tw 088f  
*tapi to napók tee { } nangule taa { }*  
 tapi to N-pa-pok teqé naN-qule ta-ná  
 but NR REAL.ACT-?-cut DIST REAL.ACT-return AND-DOWN  
 'But those who split them (the stones) returned'
- (46) Lauje UF\_LM 040  
*ma'o mongintug E sau no'inang E {0.5}*  
 ma'o moN-intug E sau no-'inang E  
 go\_out ACT-look\_for E NR IRR(UG)-eat E  
 '(they went to the river) to look for food.'

In Ratahan, a voice-marked form nominalised by *to* may then also co-occur with the NP-markers *su* or *nu*:

- (47) Ratahan uf\_tw 049  
*nayuma na sto kunaq (0.2) pinangaren Pinatén {0.9}*  
 na-yuma na su-to kunaq in-paN-aren Pinaten  
 REALSTAT-arrive ?? LOC-NR like? REAL(UG)-GER-name Pinaten  
 'they arrived at the place called Pinaten'

## 9 Voice in statives

Philippine-type languages make a basic distinction between two event types, i.e. dynamic and stative. There is no special morphological marking for dynamic events, which are characterised by the fact that they involve actors who are in full control of the event. Statives are morphologically marked by a prefix (in Tagalog *ma-* and/or *ka-*) and comprise a broad range of event classes, ranging from events in which the actor is not fully in control but just manages or happens to do something, to events which do not involve actors.

Statives allow for up to four basic voice-mood derivations, involving at least in part the same morphology as the basic voice-mood derivations for dynamic events. The details of these derivations vary quite substantially, depending on the meaning of the root (for roots denoting perceptions there are different derivational possibilities than for roots denoting emotional states, etc.).

In Ratahan, statives may be marked for voice and mood in essentially the same way as in Tagalog (there are differences with regard to many details and in particular the overall productivity and generality of these formations). For example, with roots denoting emotional states the prefix *ma-* (realis *na-*) marks the theme, i.e. the entity that is or was undergoing a given emotion:

- (48) Ratahan uf\_kera 034  
*tomponú na-awuq e rapa {1.3}*  
 tomponú na-awuq Ce rapa  
 turtle REAL.STAT-annoyed COMPL RPRT  
 'the turtle got angry'

When the same root is marked for stative locative undergoer voice (affix combination *ka--an*), the subject is the cause of the emotion (the place at, and because of which, the emotion is felt), while the theme occurs in the genitive slot:

- (49) Ratahan et\_ww N23b  
*yaq kinaawuqan e*  
*yaq in-ka-awuq-an ne*  
 1SG REAL(UG)-STAT-annoyed-UG.L 3SG.GEN  
 'he got mad at me (I became the reason/object of his anger)'

A similar contrast is expressed by these derivations with roots denoting physical states. Here, however, the subject of the stative locative undergoer derivation is an experiencer:

- (50) Ratahan et\_lw 24\_4  
*nakawus e susu.*  
*na-kawus Ce susu*  
 REAL.STAT-used\_up COMPL milk  
 'The milk is finished/used up.'
- (51) Ratahan et\_lw 24\_2  
*kinakawusan nu susu kami*  
*in-ka-kawus-an nu susu kami*  
 REAL(UG)-STAT-used\_up-UG.L GEN milk 1PLEX  
 'We ran out of milk/the milk ran out on us.'

As far as I can currently ascertain, Lauje does not have voice-marked stative derivations. This assessment, however, needs further investigation and testing.

## 10 Summary and discussion

This section presents a brief summary of the (possibly) voice-related morphosyntactic features discussed in the preceding sections, focussing on the distribution of these features among the three languages under review and putting them into a wider typological perspective. To begin with, let us briefly note some voice-related features which are shared by Tagalog, Ratahan, and Lauje:

- the voice system consists of more than two basic voices.
- all voices involve some kind of morphological marking (with one exception in Ratahan, see Table 9).
- non-subject core arguments regularly occur in immediate postverbal position.
- actors in undergoer-oriented constructions are marked as genitives.

These features have not been discussed at length in the preceding sections since the focus there was on morphosyntactic features which differ among the three languages. Still, these shared features are of major typological import because they are among those features which render the voice system in (some) western Austronesian languages remarkable when compared to the voice systems found in other linguistic areas and language families. Note, however, that though they are widely shared among western Austronesian languages, it is far from clear whether they are shared by all of them.

As for the features which are not shared by Tagalog, Ratahan, and Lauje, there are altogether eleven features which have been reviewed in the preceding sections. For six of these, Tagalog groups with Ratahan, as shown in Table 7.



**Table 7:** Features shared by Tagalog and Ratahan

	Tagalog	Ratahan	Lauje
second core argument in ditransitive construction marked oblique	yes	yes	no
undergoer core argument in actor-oriented constructions marked as genitive (common noun phrases only)	yes	yes	no
voice derivations for statives	yes	yes	no
applicative suffix(es)	no	no	yes
pronominal prefixes in basic voice-mood paradigm	no	no	yes
non-realis prefix	no	no	yes

From a typological point of view, the interesting question is, of course, whether some of the features in Table 7 (and in the following tables) correlate with each other in such a way that the presence of one implies the presence of another. More than just three languages would have to be considered in order to make the search for correlations a typologically interesting one. Nevertheless, some very preliminary observations and suggestions can be derived from the data presented here and the (rather scarce) comparative literature on western Austronesian languages.

For two features, i.e. the presence of pronominal prefixes and applicative suffixes, it has been suggested repeatedly that they correlate with each other. More precisely, the presence of both pronominal prefixes and applicative suffixes is held to be the crucial characteristic that distinguishes Philippine-type languages from Indonesian-type languages.<sup>7</sup> And while it appears to be true that none of the generally recognised Philippine-type languages<sup>8</sup> exhibits both these features, it is not true that the languages of western Indonesia all exhibit both of them. Thus, for example, there are no pronominal prefixes or proclitics in Balinese, a language which otherwise appears to share many characteristics with neighbouring languages such as Javanese and Sasak, including applicative suffixes.<sup>9</sup> Still, the co-occurrence of

<sup>7</sup> To date, the distinction between Philippine-type and Indonesian-type languages (these are the terms used by Wolff (1996)) has been discussed exclusively in historical terms, i.e. as innovations found in Indonesian-type languages with respect to an older, possibly Proto Austronesian or Proto Malayo-Polynesian system which is hypothesised to have been inherited more or less intact by Philippine-type languages (Wolff 1996; Zobel). Note that the historical and the typological perspectives, though overlapping to a considerable degree, differ with regard to the inferences which may be drawn from the same set of data. A particular set of features may not qualify as an innovation from a historical point of view but may still provide a useful typological grid. Here, we are only concerned with typological groupings and generalisations.

<sup>8</sup> That is, in addition to all the Austronesian languages spoken in the Philippines, the languages of Sabah and Sarawak and northern Sulawesi, and Yami.

<sup>9</sup> The typological (and historical) relevance of pronominal prefixes is also somewhat questionable because of the fact that there appear to be pronominal prefixes in several Formosan languages, including Paiwan and Puyuma. But then, the typological position of Formosan languages, though often subsumed without much discussion under the Philippine type, is far from clear.

applicative suffixes with pronominal prefixes is so pervasive in western Indonesia that it surely constitutes a strong typological tendency.

A correlation of similar strength exists, I suspect, between the presence of applicative suffixes and the lack of voice derivations for statives. Rather than having two partially separate voice paradigms (one for dynamic predicates and one for stative predicates), it appears to be the case that in some languages of western Indonesia at least, stative morphology (such as Indonesian *ter-* and Balinese *ka-*) is in complementary distribution with voice marking morphology. However, this topic is hardly ever addressed in the literature dealing with the voice morphology in languages of western Indonesia, and thus certainly needs further study before any serious typological claims can be advanced.

The search for further correlations for two of the remaining features mentioned in Table 7 will be limited by the fact that the morphological category in question does not exist in many languages of western Indonesia. Most of these languages do not mark modal and/or aspectual distinctions on the verb. Thus, the question of whether or not there is a non-realis prefix in the basic voice paradigm does not arise. Similarly, there is no genitive marker in many of these languages. Thus, again, the question of whether or not the undergoer in actor-oriented construction is marked as a genitive does not arise. But in the latter instance it is still possible and useful to investigate the fundamental issue, i.e. whether actors in undergoer-oriented constructions exhibit essentially the same morphosyntactic features as undergoers in actor-oriented constructions (see (29) and (30) above).

Finally, it may be noteworthy that of the six features mentioned in Table 7 four pertain to verbal morphology. This is in contrast with the four features shown in Table 8, three of which pertain to the morphosyntax of nominal expressions. These features are shared between the two Sulawesi languages and distinguish them from Tagalog and other Philippine-type languages which resemble Tagalog.

**Table 8:** Features shared by Ratahan and Lauje

	Tagalog	Ratahan	Lauje
no (or optional) NP-marker for subject	no	yes	yes
subject may occur in immediate preverbal position	no	yes	yes
pronouns are second position clitics	yes	no	no
voice-mood marked derivations can be used with quantifiers	yes	no	no

If Ratahan is taken to be a Philippine-type language – and the features relating to verbal morphology in Table 7 strongly suggest that – then the distribution of the features in Table 8 shows that certain very typical features of nominal and pronominal expressions in Tagalog and most other languages spoken in the Philippines are not necessary correlates of Philippine-type voice morphology. Or, perhaps more productively, it suggests a division of Philippine-type languages into two subtypes: those which are similar to Tagalog in that all noun phrases are marked with some kind of marker, pronouns typically occur in clitic positions, and some kind of prosodic break or inversion marker has to be used when a subject precedes the predicate; and those which are similar to Ratahan in that subjects freely occur in pre-predicate position, pronouns are not second position clitics, and noun phrase marking is either

optional or non-existent. From Clayre's (1996) survey of the languages of Sabah and Sarawak it appears that these languages pattern with Ratahan with regard to these features.

The last feature mentioned in Table 8 — that voice-mood marked derivations can be used with quantifiers — is of interest in that it provides some indication of the degree to which a systematic distinction can be made between nouns and verbs in a given language. The noun-verb distinction in western Austronesian languages is, in general, less clearly developed than in the more familiar European languages. However, these languages certainly differ to the extent to which semantically nominal and semantically verbal expressions may occur in the same morphosyntactic contexts and hence can be more or less sharply distinguished from each other.

For Tagalog, it has been argued that there are actually no morphosyntactic contexts which provide a basis for distinguishing between nouns and verbs (understood as primary morphosyntactic form classes).<sup>10</sup> The fact that in Tagalog putative verbs (i.e. words which express actions and are marked for voice, aspect and mood) may be used with quantifiers without undergoing further derivation (see examples (42) and (43) above) is perhaps the strongest argument for this view. Indonesian-type languages, on the other hand, appear to be much more restrictive with regard to the possibility of directly quantifying voice-marked derivations. Thus, in (standard) Indonesian it is not possible to say *\*ada menjaga anak-anak* for 'someone is looking after the children'. Instead, the phrase *menjaga anak-anak* has to be nominalised with *yang* when it is to be used as the argument of the existential operator *ada* (thus *ada yang menjaga anak-anak* is fine). Unfortunately, for most western Austronesian languages no reliable information is available for this potentially very important typological parameter.

To conclude this discussion of the morphosyntactic features not shared between Tagalog, Ratahan, and Lauje, we may note that there is one feature which is found only in Ratahan (Table 9).

**Table 9:** Feature unique to Ratahan

	Tagalog	Ratahan	Lauje
morphologically unmarked form as part of the basic voice paradigm	no	yes	no

For Philippine-type languages, the occurrence of a morphologically unmarked form as part of the basic voice paradigm may appear to be unusual. Ratahan, however, is not the only Philippine-type language in which this phenomenon is found. According to Table 14 in Clayre (1996:75), it is also widespread among the Philippine-type languages of Sarawak and Arms (1996) reports it for Sindangan Subanen, a language spoken in Western Mindanao. Note also that the thematic undergoer prefix *i-* is quite often dropped in Tagalog spontaneous discourse (and probably in many other Philippine-type languages as well):

<sup>10</sup> See Himmelmann (1991) and Gil (1993) for further discussion and references.

- (52) Tagalog donat 273  
**lalatag**                    mo            doón  
 i-RED1-latag                mo            doón  
 UG.T-RED1-spread\_out 2SG.GEN DIST.LOC  
 'you'd spread it (the sack) out there'

Thus, I would surmise that the occurrence of a morphologically unmarked form as part of the basic voice paradigm is not a feature of particular typological import. This does not preclude the possibility that it is of major historical import (the occurrence of a morphologically unmarked form may lead to a restructuring of the whole paradigm, possibly with further consequences for the overall system of grammatical relations).

## References

- Arms, D.G., 1996, Categories of the Sindangan Subanen verb. In Steinhauer, ed. 1996: 89-114.
- Berg, René, van den, 1996, The demise of focus and the spread of conjugated verbs in Sulawesi. In Steinhauer, ed. 1996:89-114.
- Bloomfield, Leonard, 1917, *Tagalog texts with grammatical analysis*. Urbana: The University of Illinois.
- Clayre, Beatrice, 1996, The changing face of focus in the languages of Borneo. In Steinhauer, ed. 1996:51-88.
- Gil, David, 1993, Tagalog Semantics. *Berkeley Linguistics Society* 19:390-403.
- Himmelmann, Nikolaus P., 1991, *The Philippine challenge to universal grammar*. Arbeitspapier Nr. 15, Köln: Institut für Sprachwissenschaft.
- 1996, Person Marking and Grammatical Relations in Sulawesi. In Steinhauer, ed., 1996:115-136.
- 1997, The paragogic vowel in Lauje (Tomini-Tolitoli): phonology or morphosyntax? In Cecilia Odé and Wim Stokhof, eds *Proceedings of the Seventh International Conference on Austronesian Linguistics*, 81-103. Amsterdam: Rodopi.
- Himmelmann, Nikolaus P., compiler, 2001, *Sourcebook on Tomini-Tolitoli languages*. Canberra: Pacific Linguistics.
- Himmelmann, Nikolaus P. and John U. Wolff, in prep., *Ratahan morphosyntax*. München: Lincom Europa.
- Steinhauer, Hein, ed., 1996, *Papers in Austronesian linguistics No.3*. Canberra: Pacific Linguistics.
- Wolff, John U., 1996, The development of the passive verb with pronominal prefix in Western Austronesian languages. In Bernd Nothofer, ed. *Reconstruction, classification, description: festschrift in honor of Isidore Dyen*, 15-40. Hamburg: Abera-Verlag.
- Yoshimura, Chikao, in prep., A descriptive grammar of Tialo. Manuscript, The Australian National University, Canberra.