

# Split S in the Indonesian Area: Forms, Semantics, Geography

Marian Klamer

Leiden University

[marian.klamer@let.leidenuniv.nl](mailto:marian.klamer@let.leidenuniv.nl)

Split S (SS) systems are found in languages where an intransitive argument S is encoded through case marking, verbal agreement, or both, in the same way as a transitive agent (A), or a transitive patient (P). This paper presents some of the results of a survey of SS systems found in languages of the Indonesian area, covering the larger part of Indonesia, including the Republic of East Timor but excluding the Papuan mainland and the Borneo continent. For the survey, a sample of 39 (28 Austronesian, 11 non-Austronesian) languages was selected on the basis of areal as well as genetic considerations. The sample contains 16 languages with a SS system, and 23 languages without one. Four case studies of languages with SS systems are presented in the paper. Since Acehnese is the best-known Indonesian language with SS, it is taken as a starting point, and then compared with the non-Austronesian languages Klon (Alor island) and Tobelo (N Halmahera), as well as with the Austronesian language Kedang (Flores). For each language, the structural patterns as well as the semantic factors involved in SS are described. The overall conclusion is that a structural or semantic feature that uniquely defines SS in the Indonesian area appears not to exist. In the final section of the paper, the geographical distribution of SS in Indonesia is examined. Although in absolute numbers, most of the languages with SS are located in Eastern Indonesia, this unequal distribution is shown to be statistically insignificant -- it is simply a result of the fact that the number of languages in Eastern Indonesia is four times higher than in Western Indonesia.

## 1. Introduction

Split S (SS) systems are found in languages where an intransitive argument S is encoded through case marking, verbal agreement, or both, in the same way as either a transitive agent (A), or a transitive patient (P).

This paper presents some of the results of a survey of SS systems found in languages of the Indonesian area. The area of research covers the larger part of Indonesia, including the Republic of East Timor, but excluding the Papuan mainland and the Borneo continent. In this area, approximately 400 languages are spoken. Of these, 39 were selected for the sample (28 Austronesian, 11 non-Austronesian). Table 1 contains a list of the sample, with geographical locations, genetic affiliation as well as references. The locations of the islands mentioned in Table 1 are indicated on Map 1.

Although split intransitive phenomena are frequently reported for languages in Indonesia, not all such splits are considered Split S for the purposes of this paper. In section 2, I define what is considered SS here, and what is not. In section 3, I present four case studies of languages with SS in Indonesia, including 2 Austronesian and 2 non-Austronesian (Papuan<sup>1</sup>) languages.

In section 4, a summary of the observed patterns is presented. It appears that the semantics involved in SS in the Indonesian area are very similar to the patterns found elsewhere in the world. Furthermore, the four case studies suggest that there does not exist a

---

<sup>1</sup> Note that ‘Papuan’ is not a genetic group, for discussion and references, see Foley 1986, 2000, Ross 2000, 2005. In this paper the term is used as an equivalent of ‘Non-Austronesian’.

typical Austronesian or non-Austronesian type of SS, nor a unique semantic or structural feature by which the type of SS systems found in the Indonesian area can be characterised.<sup>2</sup>

Regarding the geographic distribution of SS, the data in Table 1 show that the number of languages with SS in eastern Indonesia is much higher than the number of languages with SS in western Indonesia. How significant is this geographic distribution? Does it suggest that SS is an areal feature of eastern Indonesia? This issue is addressed in section 5.

## 2. SS in the Indonesian area: what it is, and what it is not

A language may be split intransitive, but if it has no multiple alignment of S, and/or if the morphological shape of a verb plays a role in the multiple alignment of S, it is not considered an SS language in the survey reported here. In other words, multiple alignment of S is the first criterion; but this is only an instance of SS if the split occurs with UNDERived verb forms -- verbal derivational morphology should not be involved in the split.

The diagnostic that the variable alignment must occur with morphologically simple verbs is an important one to keep the phenomenon cross-linguistically comparable, since it is very common for Austronesian languages to have morphologically derived intransitives that differ from each other in all kinds of aspectual properties and/or the volitionality/control of their argument. A few examples are Balinese (Arka 2003), the Formosan language Amis (Tsukida 2005) and the Northern Borneo language Begak (Goudswaard 2005). While these languages can be called ‘split intransitive’ because they have more than one type (class) of intransitive verb, possibly mirrored in a split in the alignment of S, they are not diagnosed as split S languages here, on the assumption that the split should not be dependent on verbal derivational morphology alone.

## 3. Four case studies of SS in the Indonesian area

### 3.1. Introduction

Consider the list of 39 languages in Table 1.

[insert Table 1 about here]

A total of 16 of them are reported to have SS. The locations of the languages are indicated by the islands on which they are spoken; these islands are indicated on Map 1. In the present paper, four of the languages in Table 1 are discussed in detail. The locations of these four languages are indicated on Map 2.

[insert Map 1 about here]

[insert Map 2 about here]

The first case study presented here is Acehnese, since this is the best known Indonesian language with SS. The second case study is Klon, a Papuan language spoken thousands of kilometers eastwards, on the island of Alor, and the similarities between Papuan Klon and Austronesian Acehnese in terms of SS are pointed out. Thirdly; Klon is compared with

---

<sup>2</sup> This paper is an abbreviated version of Klamer (to appear), which contains case studies of an additional three languages: Kambara, Larike and Mori Bawah.

another non-Austronesian language, Tobelo. We will see that although both these languages have SS, the formal patterns and semantics are quite different. And finally, I discuss Kedang, a language spoken on Flores island. Though this is an Austronesian language like Acehnese, the SS patterns in both languages are very different. In section 4 I summarise the patterns. The general conclusion is that there are no patterns that may be considered typical for SS systems found in the Indonesian area.

### 3.2. Acehnese

A widely known example of an Austronesian language with SS is Acehnese (Durie 1985, 1987). In an Acehnese transitive clause, A is marked with a verbal proclitic, and P with an optional enclitic, as illustrated in (1):

- (1) a. Gopnyan      ka      lon=ngieng(=geuh)  
           s/he            In      1s=see=3s  
           ‘I saw him/her’ (Durie 1987:369)

Acehnese has three lexical classes of intransitive root verbs: (i) verbs that align S like A, with a proclitic; (ii) verbs that align S like P, with an optional enclitic, and (iii) verbs that show fluid SS and align their argument like A when it is in full control, and like P when it is not.

The first class of verbs includes verbs of motion and posture with an animate argument (*jak* ‘go’, *döng* ‘stand’, *beudöh* ‘get up’, *iem* ‘be still’ (Durie 1985:63)), verbs of bodily activity (*khêm* ‘laugh/smile’, *klik* ‘cry’, *muntah* ‘vomit’), verbs of speech and thought or mental activity (e.g., *marit* ‘talk’, *kira* ‘think’, *pham* ‘understand’), as well as some emotion verbs (e.g. *chên* ‘love/feel sympathy for’, *têm* ‘want, like’) (Durie 1985:64). The S of these verbs must be animate, and is marked like A, with a proclitic. An illustration is (2):<sup>3</sup>

- (2)    Geu= jak gopnyan  
        3s    go s/he  
        ‘S/he goes’ (Durie 1987:369)

The second class are events and states with arguments that need not be animate (*rhët* ‘fall’, *reubah* ‘topple over’, *jeuet* ‘become’, *trôh* ‘happen/arrive’), many emotion verbs (*ku’eh* ‘envy’, *seugan* ‘not want to’, *êk* ‘like/feel inclined’), personal attributes (*beuhë* ‘brave’, *caröng* ‘clever’, *gasien* ‘poor’), and bodily and mental states of animate arguments (*sakêt* ‘sick/hurting’, *gatay* ‘itchy’, *mumang* ‘confused’, *dawôk* engrossed) (Durie 1985:64-66). The argument of these verbs is aligned like P, with an optional enclitic. An illustration is

---

<sup>3</sup> **Notational conventions and abbreviations.** In the notation of the examples a clitic is separated from its host by [=], an affix by [-]. **Abbreviations:** 1,2,3=person, Abs=Absolutive, Acc=accusative, Ag.focus=agent focus, APassive=antipassive, App=applicative, Art =article, Cnj=conjunction, Compl=completive aspect, Ctr=marker of control sentence, Dat=dative, Dei=deictic element, Dem=demonstrative, Dist=distal, e=exclusive, Emph=emphasis, f=female, Gen=genitive, i=inclusive, Impf=imperfective aspect, In=Inchoative, Inc=inceptive aspect, Irr=irrealis mood, Iter=iterative aspect, Loc=locative preposition, Mod=mood marker, nh=non-human, Nom=nominative, Neg=negation, obj=objective, p=plural, Part=participle, Poss=possessor, Pfv=perfective, Red=reduplication, Rel =relative clause marker, s=singular subj=subjective, I=P-marking paradigm I, II=P-marking paradigm II, IV=P-marking paradigm IV.

- (3) Gopnyan rhët(=geuh)  
s/he fall 3s  
'S/he falls' (Durie 1987:369)

The third class overlaps with the other two – it contains many emotion verbs (*cinta* 'love/favour', *galak* 'like', *beungeh* 'angry'), verbs of thought or mental activity (*syök* 'suspect', *yakin* 'believe/be sincere'), ability (*jeuet* 'able', *keuneuk* 'likely to'), aspect (*mulayi* 'begin', *piyôh* 'stop'), personal attributes or attitudes (*horeumat* 'polite', *kaya* 'rich', *malee* 'shy'), verbs of motion (*ilê* 'buzz off!'), and the verbs *udêp* 'live' and *matê* 'die' (Durie 1985: 66-67). The S of this third verbal class is fluid: it is aligned like A when it is a 'wanting participant', as in (4), and like P when it is the 'ultimately affected participant' of an event (Durie 1985: 55, 56), as in (5):

- (4) Rila ji= matê  
ready 3.(familiar) dead  
'He was ready to go to his death' (Durie 1985:57)

- (5) ... matê(=jih)  
dead 3.(familiar)  
'... he died' (Durie 1987: 376)

In other words, while there is clearly some semantic basis for the lexical distinction between class one and two -- the obligatory vs. optional animacy of the argument --, the demarcation of the third verbal class is rather arbitrary; and various semantic types of verbs are included in this class.

In sum, the alignment of the S in Acehnese is mostly determined by the lexical subcategorisation properties of the verb, i.e., the class a verb belongs to. The verbal classes (i) and (ii) are characterised by the animacy of their S. The third class of intransitives shows fluid S marking—the S of a verb may be marked as either A or P, depending on the question whether S is volitional, and controlled, or not. In Acehnese, the split and fluid marking of S is thus very transparently dependent on the agentive properties of the S.

### 3.3. Klon

Klon (Baird 2005, in prep.) is a non-Austronesian language spoken on the island of Alor, north of Timor Island. Klon belongs to the Trans New Guinea family. A pronominal A in Klon is aligned as a free pronoun and occurs in preverbal position. A pronominal P in Klon is expressed as a prefix or proclitic. The paradigms are given in (6). In general, the choice of P-marking paradigm depends on the lexical specification of the transitive verb. More than 50% of the transitives align P with paradigm II, about 30% align P with paradigm I, and about 4% align P with paradigm IV.<sup>4, 5</sup>

- (6) Klon free pronouns (full & reduced) and pronominal prefix classes (Baird, 2005: 2, 3)

<sup>4</sup> About 10% of the transitives may be prefixed by a choice between to classes of prefixes; in which case the choice is motivated by the semantics of the context of use (Baird, in prep.).

<sup>5</sup> Class III of the P marking bound pronouns is lacking in this overview, because it is a paradigm of proclitics which shows aberrant behaviour. See Baird (2005, in prep.).



volitional and controlling participant, but rather an affected one. This is illustrated in (10b), where S is aligned like P with a prefix from paradigm IV. In contrast to (10a), where S is aligned like A with a free pronoun, S is interpreted as a more affected participant in (10b). For the alignment of S like P, paradigm IV is used most often, although there are some verbs that select paradigm I (Baird 2005:10).

- (10) a. A      kaak  
           2s     itchy  
           ‘You’re itchy.’
- b. E-     kaak  
           2s.IV itchy  
           ‘You’re itchy (and affected).’ (Baird 2005:8)

To conclude, Klon has a split in the alignment of S. In most cases, the marking of an S is determined by the lexical class to which a verb belongs. Only the third verbal class has fluid SS, and the split in the alignment of S in this class is motivated by the (relative) lack of agentive features of S.

Note that S need not be a volitional and controlling participant to be aligned like A, since the argument of ‘to be itchy’ in (10a) cannot really be volitional, nor can it exercise control on the experience of being itchy. Yet it is aligned like A. This is in line with the analysis that the default alignment of a Klon S is like A. Only diverging from the default pattern needs a semantic motivation.

### 3.4. Tobelo

Like Klon, Tobelo (Holton 2003) is also a Non-Austronesian language. Tobelo is spoken in North Halmahera and belongs to the West Papuan family. It is not genetically related to Klon. Like Acehnese and Klon, Tobelo intransitive verbs are divided into three lexical classes. In the first class, S is aligned like A, with a subjective prefix. In the second class, S is aligned like P with an objective prefix. In the third class, there is fluid S marking. Unlike Achenese and Klon, however, the semantic parameter underlying SS in Tobelo is not the (lack of) agentivity of the argument, but the lexical aspect of the verb.

In Tobelo, A, P and S are cross-referenced on the verb with the ‘subjective’ and ‘objective’ prefixes in (11) (Holton 2003: 37-38). Subjective and objective prefixes both occur before the verb, in that order.

(11) Tobelo subjective and objective prefixes (Holton 2003: 38, 39)

	Subjective prefix	Objective prefix
1s	<i>to-</i>	<i>hi-</i>
2s	<i>no-</i>	<i>ni-</i>
3s male	<i>wo-</i>	<i>wi-</i>
3s female	<i>mo-</i>	<i>mi-</i>
1pi	<i>ho-</i>	<i>na-</i>
1pe	<i>mi-</i>	<i>mi-</i>
2p	<i>ni-</i>	<i>ni-</i>
3p male/female	<i>yo-</i>	<i>aa-</i>
3 neutral	<i>i-</i>	<i>a-</i>

In (12), the subjective prefix marks A, the objective P:

- (12) I-hi-goli  
 3subj-1obj-bite  
 ‘It/they bit me’ (Holton 2003: 39)

The intransitive verbs are divided into three lexical classes. Class (i) denotes active, dynamic, and telic events, such as *hioru* ‘paddle’, *hoho* ‘fly’, *oara* ‘run’, *olyomo* ‘eat’, *temo* ‘speak’, *dumumu* ‘dive’, *toimi* ‘shoot’ and *phiki* ‘bathe’, (13), but also include non-volitional, non-controlled verbs such as *ha’ngeru* ‘sneeze’, *wunenge* ‘vomit’, *iete* ‘laugh’, *guroko* ‘snore’, *gehanga* ‘yawn’, *ari* ‘cry’, (14), and *lyahini* ‘drift away’, (14)

- (13) Mo-phiki  
 3f.subj-bathe  
 ‘She bathes.’ (ibid, p.56)

- (14) De i-sobo-oli i-lyahini, ....  
 and 3-depart-Repetitive 3-float  
 ‘And they floated away again...’ (ibid, p. 56)

In other words, the semantic characterisation of class (i) must refer to the notion of eventhood: “events, not just actions, follow the ‘active’ subjective paradigm” (Holton 2003:56).

The second class of verbs in Tobelo aligns S like P. The verbs in this class are statives such as *pehaka* ‘be wet’, *hauku* ‘be hot’, *modongo* ‘angry’, *kuata* ‘strong’, *omu* ‘jealous’, and in addition, the class contains intransitive verbs with a pleonastic, ‘dummy’ subjective 3neutral prefix, preceding the prefix that marks S (Holton 2003: 56, footnote 13). *Bole* ‘be tired’ is such a verb:

- (15) I-mi-bole  
 3.subj-3f.obj-tired  
 ‘She is tired.’ (ibid, p. 57)

The third class of Tobelo intransitive verbs shows fluid SS. When S is aligned like A, it has a more telic, dynamic sense than when S is aligned like P clause. This is illustrated in (16).

- (16) a. To-birahi  
 1s.subj-happy  
 ‘I rejoice’ (ibid, p. 58)
- b. I-hi-birahi  
 3.subj-1.obj-happy  
 ‘I am happy’ (lit. ‘It happies me’) (ibid, p. 58)

Holton reports similar contrasts for the verbs in (17): the alignment of S like A or like P, renders different interpretations for the same verb, as the contrast between the columns headed S=A and S=P shows. The verb in the S=P column is stative, atelic, while the verb in the S=A column is more telic and dynamic.

(17) Tobelo intransitives with fluid SS (Holton 2003: 58)

	S=A	S=P
<i>-eluku</i>	to tell lies	be a liar
<i>-kioko</i>	go to sleep	be asleep
<i>-modongo</i>	fear	be afraid
<i>-hihanga</i>	go astray	be lost
<i>-tikiti</i>	cough	cough continuously
<i>-tohata</i>	angry	evil

In a subclass of these verbs, SS is used to mark a contrast between stative and inchoative interpretations, i.e. *daluku* S=A ‘get drunk’ vs. S=P ‘be drunk’, or *hiri* ‘sick’ in (18):

- (18) a. Mo- hiri                      b. I- mi- hiri  
       3f sick                            3 3f sick  
       ‘She’s getting sick.’        ‘She’s sick.’                      (ibid: 58)

In sum, Tobelo verbs of class (i) are all events, and this determines that their S aligns like A. The verbs in class (ii) align S like P because they denote states. The third class contains verbs that are UNSPECIFIED for lexical aspect. These verbs get their aspectual interpretation by aligning S either like A (the verb is then interpreted as active, telic, dynamic, inchoative) or like P (the verb then gets a stative, atelic interpretation).

Thus, in two classes of Tobelo verbs, the lexical aspect of the verb (event vs. state) DETERMINES the alignment of S, while the verbs of the third class get their aspectual interpretation as a RESULT of the alignment of S.

### 3.5. Kedang

Kedang (Samely 1991) is spoken in the eastern part of the island of Flores. Unlike Acehnese, Klond and Tobelo, it has only fluid S: the variable alignment of S does not depend on verbal classes. The lexical aspect of verbs does not play a role either. In principle, one and the same verb allows its S to align like A or like P.

Kedang pronominal arguments are aligned as free pronouns and/or pronominal clitics.<sup>6</sup> Kedang constituent order is AVP and SV. Kedang has no case marking on NPs, nor on pronouns -- except for the 1sg pronoun, see (19). The remaining pronouns differentiate A/S from P only by position relative to the verb: S/A pronouns precede the verb, P pronouns follow it.

(19) Kedang free pronouns (cf. Samely 1991: 70-72 vv.)(> marks breathy vowels)<sup>7</sup>

	S & A, pre-verbal	P, post-verbal
1s	>ei	>eqi
2s		o
3s		nuo

<sup>6</sup> There is a set of 19 verbs that obligatorily take subject prefixes (S or A) (Samely 1991:94-96). The prefixes are single consonants and attach to vowel-initial verbal stems. Such phonotactically triggered inflection is not considered here.

<sup>7</sup> These are the unmarked pronouns. The language also has special pronoun paradigms, which are not considered here.

1pi	<i>te</i>
1pe	<i>e</i>
2p	<i>me</i>
3p	<i>suo</i>

Kedang has a split in the marking of P: it has two distinct paradigms to mark P, henceforth referred to as paradigm I and II, see (20). Samely (1991: 70) lists both paradigms as synonymous and describes both as having a ‘subjective’ as well as an ‘objective’ function. The ‘objective’ function refers to fact that they mark P, the ‘subjective’ function refers to their S-marking function. From her examples it appears that a Kedang A is always expressed as a free pronoun, cf. (19), and never as an enclitic.

(20) Pronominal enclitics marking P or S in Kedang (cf. Samely 1991: 70-72 vv.)

	Paradigm I (PI)	Paradigm II (PII)
1s	=ku	=u
2s	=ko	=o
3s	=i	=ne
1pi		=te
1pe	=ke	=e
2p		=me
3p	=deq	=ya

The transitive clause in (29) illustrates the alignment of A and P. The A of the *maqo* ‘steal’ and *ehing* ‘deny’ is 3sg *nuo* ‘s/he’, the P of *maqo* is *doiq* ‘money’, the P of *ehing* a bound pronoun following the verb, henceforth referred to as an enclitic.<sup>8</sup>

Turning now to the intransitive clauses of Kedang, we note that S aligns like A when it occurs as a free pronoun in preverbal position. This is illustrated in (21a), where *ei* ‘I’ is the S of *pan* ‘go’ and precedes the verb. However, S may also align like P, and then it occurs as a bound pronoun following the verb, as in (21b). An additional lexical or pronominal NP may mark S in preverbal position, as in (21c).

- (21) a. >Ei pan >owe >ul...  
 I go Dei market  
 ‘I go to the market...’ (ibid., p. 79)
- b. Pan >oteq =o?  
 go Dei 2s.II  
 ‘Going up, are you?’ (ibid., p. 71)
- c. O pan >oteq =o?  
 you go Dei =2s.II  
 ‘Going up, are you?’ [slightly more courteous than (b)] (ibid., p. 71)

The pattern in (21b) is described as ‘typical for most common, somewhat casual speech’ (Samely 1991:71), while (c) is presented as a polite variety of (b). This suggests that the argument is expressed by the clitic, while the additional NP is optionally present for pragmatic reasons (politeness), or for emphasis or disambiguation. The analysis I present here focusses on the distribution of the clitics.

<sup>8</sup> Samely refers to these as ‘suffixes’ (Samely 1991: 70) but their distribution is clitic-like.

From the description in Samely (1991), it is unclear which factors determine the choice between the alignment of S like A or like P. However, non-verbal predicates typically align S like P, as in (22)-(25).

(22) Predicate is a noun:

>Anaq usun tètèq tètè: “kusing =ne.”  
 child small speak say cat =3s.II  
 ‘The children say: “It’s a cat” ’ (ibid., p. 153)

(23) Predicate is an emphatic possessor noun:

Labur nobe koqo =ne  
 dress Dei Poss.Emp 3s.II  
 ‘That dress is mine.’ (ibid., p.77)

(24) Predicate is an adjective:

Labur koqo miteng =ne  
 dress Poss.Emp black 3s.II  
 ‘My dress is black.’ (ibid., p. 77)

(25) Predicate is a location:

Koq lumar >ote =bètè wela =ne  
 1s.Poss field Dei interior 3s.II  
 ‘My field is up there in the interior.’ (ibid., p. 75)

Non-verbal predicates like these have in common that they are non-dynamic by nature -- all of them denote states. The S of such predicates is typically an undergoer, a participant without control or volition. It is thus not surprising to find that the S of such predicates is aligned like P.

Regarding the fluid S marking in Kedang, my tentative hypothesis is that this relates to the interpretation of the argument: when S is aligned like P, it has a less agentive interpretation, when it is aligned like A, it is more agentive. Thus the S in (26a) and (27a) is less agentive than the S in (26b) and (27b).

(26) a. Ebeng boraq bahe nape e bale =ke  
 watch look.at Compl then 1p.Exc return 1p.e.I  
 ‘When we finished watching, we returned’ /  
 ‘After we will have finished watching, we will return.’ (ibid., p. 91)

b. Bahe suo bale =dèq.  
 then they return Pfv  
 ‘Then they returned home.’ (ibid., p. 158)

(27) a. Heri, o kua kueq =ko?  
 Heri you why.2s cry 2s.I  
 ‘Heri, why do you cry?’

b. Nuo kueq oti mawang =i  
 s/he cry Ag.focus 2.harm 3s.I  
 ‘He cries because you harmed him.’

These examples also suggest a relation between S alignment and other grammatical properties of the clause (e.g., irrealis vs. realis, perfective vs. imperfective), but given the scarcity of data, not much more can be said about this. However, it is relevant to note that S=A alignment (and not S=P alignment) in Kedang is often found in combination with various kinds of aspect markers (Samely 1991:92), e.g. the ‘Inceptive’ *dèq mè*:

- (28) >Ei      *bèq*      *pan*      *dèq mè*  
 I            here      go      Inc  
 ‘I am going’ / ‘I will be leaving now’ / ‘I am about to go’ / ‘I will go immediately’

Now we have addressed the alignment of S like A or like P, we continue by studying more details about the alignment of S like P. In Kedang, the split in P is reflected also in the alignment of S: S is either an enclitic from PI, e.g. =*ko* ‘2s.I’ in (27a), or from PII, e.g. =*o* ‘2s.II’ in (21b).

- (29) >Ei >oroq [nuo maqo doiq] [paq nuo ehing =i]  
 I            suspect s/he steal money but s/he deny 3s.I  
 ‘I suspect he steals money but he denies it’ (Samely 1991:73)

What motivates the choice for PI or PII in the alignment of S? When is S marked with PI, and when with PII? In (30) I list the intransitive verbs found in examples throughout the sketch; those in the left-hand column align S like PI, those in the right-hand column align S like PII. Both PI and PII occur with verbs of states, events and processes, so that lexical aspect is clearly not the determining factor. Neither does the marking appear to correlate with certain verbal classes, since the verbs *bute*, *bikil* and *moruq* occur with both PI and PII.

(30) Intransitive verbs in example sentences in Samely (1991); with S marked as PI or PII

S=PI			S=PII		
verb	translation	page	verb	translation	page
<i>nore</i>	exist (‘there are’)	84	<i>tawe</i>	laugh	90
<i>beq</i>	be here	72	<i>pan</i>	go	70, 88, 89
<i>bale</i>	return	91	<i>hamang</i>	dance	93
			<i>pan &gt;oteq</i>	go up	71
			<i>туру</i>	come down	91
			<i>bèyèng</i>	run	91
			<i>moleng diqen</i>	be better (lit. healthy good)	89
			<i>nihon</i>	be light (of day)	74
			<i>mawin</i>	be wet	91
			<i>adaq &gt;alu</i>	behave refined	76
			<i>mate</i>	dead	93
<i>bute</i>	sleep	73	<i>bute</i>	sleep	73
<i>bikil</i>	broken	73	<i>bikil</i>	broken	73
<i>moruq</i>	fall	73	<i>moruq</i>	fall	73

The split marking of S with PI or PII appears to relate to the dynamicity of the predicate, i.e. whether it is a state or an event. In (31), this contrast is illustrated with the verb *bute* ‘sleep’. In the first clause the S is marked with 3s.II =*ne*, in the second sentence, it is a 3s.I =*i*. The contrast is explained as follows: “*bute=ne* conveys the static nature of the action described, implying that the person is either sound asleep, or else has slept for a considerable time. *Bute=i* emphasises the dynamic side of the action, in this case that the person has not slept for long but fell asleep only recently.” (Samely 1991:72)

- (31) Nuo bute =ne, doq-doq nuo hoko =i. Eeh, bute =i watiq.  
 s/he sleep 3s.II suddenly s/he get.up 3s.I excl sleep 3s.I again  
 ‘He slept, (then) suddenly got up. Why, now he has fallen asleep again!’ (ibid., p. 73)

In a similar way, the contrast between =ne and =i in (32) marks a difference in dynamicity: (32a) “describes the state that the flashlight is presently not usable because it is broken”, while (32b) “draws the listener’s attention to the actual breaking as the cause for its present state of being unusable” (Samely 1991:73), i.e. *bikil* gets a more dynamic event reading.

- (32) a. Koq sender bikil =ne *state*  
 1s.Poss flashlight broken 3s.II  
 ‘My flashlight is broken.’ (ibid., p. 73)
- b. Koq sender bikil =i *event*  
 1s.Poss flashlight broken 3s.I  
 ‘My flashlight got broken.’

The same distinction applies in (33). (33a) “stresses the result of the falling of the coconuts: they are now lying on the ground, while [(33b)] focuses on the falling as the prehistory of the present state.” (ibid, 73). I interpret this as (33a) describing a non-dynamic resulting state (‘to have fallen down’), and (33b) as a dynamic event (‘to be/have been falling down’).

- (33) a. Taq muruq =ya *state*  
 coconut fall =3p.II  
 ‘Coconuts fell.’ (or ‘...have fallen down’)
- b. Taq muruq =deq *event*  
 coconut fall 3p.I  
 ‘Coconuts fell.’ (or ‘...are/have been falling down’)

In sum, S is aligned like PII when the predicate indicates a (resulting) state, and like PI when it is part of an event.<sup>9</sup>

To conclude, Kedang lexical NPs and free pronouns follow a plain nominative-accusative system: A and S are marked in the same way, and occur preverbally, P is postverbal. At the same time, Kedang *dependent* pronouns align S like P, as an enclitic. Kedang has a split P, and S goes along in this split, so that S is sometimes marked with PI and sometimes with PII. In this way, a distinction between a stative or a more eventive reading of the predicate is expressed.

At the same time, Kedang *dependent* pronouns align S like P, as an enclitic. Kedang has a split P, and S goes along in this split. As a result, S is sometimes marked with PI and sometimes with PII. In this way, a distinction between a stative or a more eventive reading of the predicate is expressed.

It was also hypothesised that S is aligned like A (with a full preverbal pronoun) when it is a more agentive participant, and that it gets an undergoer-like interpretation when it is aligned like P. This hypothesis needs to be tested on a richer set of data than is available now.

<sup>9</sup> It is unclear how this alignment of S relates to the alignment of P with Paradigm I or II, though it seems that Paradigm I is typically used to mark P in contexts where the agentive features of A are emphasised, (the “Agent” or the “Action” is “in focus” (ibid., p. 81-83)), while Paradigm II is used in unmarked contexts.

#### 4. Summary and conclusion

The semantics that play a role in the SS patterns in the Indonesian area are very similar to those observed to play a role in the similar split S marking of languages elsewhere in the world (see, for example, Mithun 1991, Dixon 1994). There is not a typical Austronesian structural type of SS, nor a typical Non-Austronesian one: the verbal classes and the semantic parameters triggering the split in Acehese and Klon are rather similar -- but Acehese is Austronesian and Klon is not. At the same time, the formal characteristics of the SS pattern in Acehese and Kedang are very different, while both of these languages are Austronesian. Also, the semantic features involved in the Klon split are quite different from those in Tobelo, despite the fact that both languages are non-Austronesian. Furthermore, in Klon and Tobelo, SS is largely determined by which lexical class a verb belongs to; in Acehese the split is partly lexically determined, partly fluid; while in Kedang it is completely fluid. Another difference is that some languages not only have a Split S, but also a Split P (Klon, Kedang), and they mirror this P-split in the P-marking of S. Other languages do not have a Split P (Acehese, Tobelo), resulting in a more uniform P-marking of S. Finally, some of the languages align S with a default type (in Klon, the default alignment of S is like A), while others do not have a default alignment for S (Acehese, Tobelo). In conclusion, there is not a single semantic or structural feature that can function to characterise the SS patterns found in the Indonesian area.

#### 5. The geographical distribution of SS in Indonesia

In table 2 the SS patterns of Table 1 are summarised, and the figures are compared with the actual number of languages reported in the areas involved, taking the SIL ethnologue as source. The number of languages in the sample with SS that are spoken in the eastern part of Indonesia is much higher (14) than the number of those spoken in the west (2). Does this mean that the distribution of SS is significantly different for east and west Indonesia, i.e. that SS is a feature typical for the eastern part of Indonesia?

[insert Table 2 about here]

To test this, our null hypothesis is that the west and east have no statistically significant different distribution of SS. To evaluate this hypothesis, the Fisher's exact test for count data is applied to the data in Table 2, abbreviated in Table 3.<sup>10</sup> The outcome of the calculations is presented in (34).

[insert Table 3 about here]

(34) Fisher's Exact Test for Count Data  
 TABLE = [ 2 , 5 , 14 , 18 ]  
 Left : p-value = 0.3834739717092759  
 Right : p-value = 0.8790509966980553  
 2-Tail : p-value = 0.6776009159910279

---

<sup>10</sup> Since the numbers are so small, a X-square test cannot be used to test significance.

The 2-Tail p-value = 0.678 shows that we cannot reject our nul hypothesis that there is no significant difference in the spread of SS. To reject this hypothesis at the 5% significance level, the p-value would have to be much smaller than the value of 0.678, namely  $p < 0.05$ . Since our p-value is so much larger than that, the hypothesis cannot be rejected. In other words, there is no reason to believe that the different spread of SS in West and East is *not* random.<sup>11</sup>

In sum, the absolute numbers are that SS occurs more often in eastern Indonesian languages than in western Indonesian languages. However, in relative numbers this is not the case, since the number of eastern languages is about four times higher. In general, absolute numbers do not reveal anything about statistic significance. The significance of areal patterns can only be studied with proportional data from a representative sample, and the data must be evaluated using the appropriate statistical tests. In addition, on the basis of the case studies presented above, we concluded that a structural or semantic feature that uniquely characterises SS in the Indonesian area does not appear to exist. It is thus unclear how a putative areal feature of Split S could be defined in a way that would cover the structural and semantic variety of Split S in this area.

---

<sup>11</sup> Donohue's (2004) conclusion that SA is a feature defining Eastern Indonesia as a linguistic area is therefore not supported by the data.

## References

- Arka, I Wayan. 2003. Balinese morphosyntax: A lexical-functional approach. Canberra: Pacific Linguistics 547.
- Baird, Louise. 2002. A grammar of Kéo: An Austronesian language of East Nusantara. PhD dissertation Australian National University.
- Baird, Louise. 2004. Klon pronouns. Talk presented at the Vrijdagmiddaglezing Colloquium, Leiden University, November 5, 2004.
- Baird, Louise. 2005. Doing the Split-S in Klon. In: Jenny Doetjes and Jeroen van de Weijer (eds.), *Linguistics in the Netherlands 2005*. Amsterdam: Benjamins, 1-12.
- Baird, Louise. In preparation. *A grammar of Klon*.
- Berg, René van den. 1989. *A grammar of the Muna language*. Leiden: KITLV.
- Blust, R. A. 1993. Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics* 32, 2: 241-293.
- Bolton, Rosemary A. 1990. A preliminary description of Nuaulu phonology and grammar. M.A. thesis. University of Texas at Arlington.
- Bowden, John. 2001. *Taba, A description of a South Halmahera language*. Canberra: Pacific Linguistics 521.
- Brotherson, Anna. 2003. A Spatial Odyssey: Referring to Space in Makasai. Honours Thesis, Australian National University.
- Brown, Lea. 2000. *A grammar of Nias*. PhD dissertation Australian National University.
- Brown, Lea. 2005. Nias. In: K. Alexander Adelaar and Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, London: Routledge.
- Coward, David and Naomi Coward. 2000. "A phonological sketch of the Selaru language." In Charles E. Grimes (ed.), *Spices from the east: papers in languages of eastern Indonesia*, 9-54. Pacific Linguistics, 503. Canberra: Research School of Pacific and Asian Studies, Australian National University
- Davies, William. 1999. *Madurese*. Languages of the World/Materials. Munchen: Lincom.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Donohue, Mark. 1995. *A grammar of Tukang Besi*. Berlin: Mouton de Gruyter.
- Donohue, Mark. 1996. Bajau: a symmetrical Austronesian language. *Language* 72,4: 782-793.
- Donohue, Mark. 1996. Inverse in Tanglapui. *Language and linguistics in Melanesia* 27 (2): 101-118.
- Donohue, Mark. 2001. Split intransitivity and Saweru. *Oceanic Linguistics* 40, 2:291-306.
- Donohue, Mark. 2004. Typology and linguistic areas. *Oceanic Linguistics* 43,1: 221-239.
- Djawanai, Stephanus Anthonius. 1980. *A study of the Ngadha text tradition: a linguistic investigation of the collective mind of the Ngadha people on the island of Flores, Indonesia*. PhD dissertation University of Michigan, Ann Arbor.
- Durie, Mark. 1985. *A grammar of Acehese on the basis of a dialect of North Aceh*. Dordrecht: Foris Publications.
- Durie, Mark. 1987. Grammatical relations in Acehese. *Studies in Language* 11.2:365-399.
- Engelenhoven, Aone. 2004. *Leti, a language of Southwest Maluku*. Leiden:KITLV Press.
- Foley, William A. 1986. *The Papuan languages of New Guinea*. Cambridge: University Press.
- Foley, William A. 2000. The languages of New Guinea. *Annual Review of Anthropology* 29:357-404.
- Foley, William. 2005. Semantic parameters and the unaccusative split in the Austronesian language family. *Studies in Language* 29,2: 385-430.
- Friedberg, Claudine. 1978. *Comment fut tranchée la liane celeste et autre textes de littérature orale bunaq (Timor, Indonésie)*. Paris: Société d'études linguistiques et anthropologiques de France Selaf.

- Goudswaard, Nelleke. *The Begak (Ida'an) language of Sabah*. PhD dissertation Vrije Universiteit Amsterdam. Utrecht: LOT.
- Grimes, Charles E. 1991. *The Buru language of Eastern Indonesia*. Unpublished PhD dissertation, Australian National University.
- Grimes, Charles E. (ed.) 2000. *Spices from the East: Papers in languages of eastern Indonesia*. Canberra: Pacific Linguistics 503.
- Heuvel, Wico van den, In preparation. *A grammar of Biak*. PhD thesis Vrije Universiteit, Amsterdam.
- Hughes, Jock. 2000. The morphology of Dobel, Aru, with special reference to reduplication. In: Charles E. Grimes (ed.).
- Hull, Geoffrey. 2001a. A morphological overview of the Timoric Sprachbund. *Studies in languages and cultures of East Timor* 4: 98-205.
- Himmelman, Nikolaus. 2004. On Statives and Potentives in Western Austronesian (Mostly Tagalog). In: Paul Law (ed.), *Proceedings of AFLA 11*. [ZASPiL No. 34]. Berlin: Zentrum fuer Allgemeine Sprachwissenschaft, Typologie und Universalienforschung (ZAS), 103-119.
- Himmelman, Nikolaus. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In: K. Alexander Adelaar and Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, London: Routledge/Curzon, 110-181.
- Holton, Gary. 2003. *Tobelo*. Languages of the World/Materials. Munchen: Lincom.
- Keraf, Gregorius. 1978. *Morfologi dialek Lamalera*. Ende, Flores: Percetakan Offset Arnoldus.
- Klamer, Marian. 1998. *A grammar of Kambera*. Berlin: Mouton de Gruyter.
- Klamer, Marian. In press. Differential subject marking in Kambera. In: Helen de Hoop and Peter de Swart (eds.), *Differential Subject Marking*. Amsterdam: Benjamins.
- Klamer, Marian. To appear. Semantic alignment in the Indonesian area: forms, semantics, geography. To appear in: Soeren Wichman and Mark Donohue (eds.), *The Typology of Semantic Alignment Systems*. Oxford: Oxford University Press.
- Klamer, Marian. In preparation. *A grammar of Teiwa*.
- Klamer, Marian. In preparation. *A short grammar of Alorese*.
- Klinken, Catharina van. 1999. *A grammar of the Fehan dialect of Tetun, an Austronesian language of West Timor*. Canberra: Pacific Linguistics C-155.
- Kratochvil, In preparation. *A grammar of Abui*. PhD dissertation, Leiden University.
- Laidig, Wyn D. and Carol Laidig. 1991. *Tarus Sou Rikedu (Larike Grammar)*. Ambon: Universitas Pattimura/Summer Institute of Linguistics.
- Laidig, Carol J. 1992. Syllables, segments and stress in Larike. In: Donald A. Burquest and Wyn D. Laidig (eds.), *Phonological studies in Four languages of Maluku*. Arlington, Texas: SIL/Pattimura University. Mead, David. 2005. Mori Bawah. In: K. Alexander Adelaar and Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*. London: Routledge.
- Mithun, Marianne. 1991. Active/agentive case marking and its motivation. *Language* 67.3: 510-546.
- Muller-Gotama. *Sundanese*. Languages of the World/Materials. Munchen: Lincom.
- Oglobin, Alexander K. 2005. Javanese. In: K. Alexander Adelaar and Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, London: Routledge.
- Owens, Melanie. 2000. Agreement in Bimanese. MA Thesis, University of Canterbury, New Zealand.

- Ross, Malcolm. 2000 [manuscript]. Defining the Trans New Guinea family: preliminary evidence from pronouns. In: Andrew Pawley, Malcolm Ross and Meredith Osmond (eds.) *Papuan languages and the Trans New Guinea phylum*. Canberra: Pacific Linguistics.
- Ross, Malcolm. 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In: Pawley, A., Attenborough, R., Golson, J., and Hide, R. (eds.). *Papuan Pasts*. Canberra: Pacific Linguistics.
- Samely, Ursula. 1991. *Kedang (Eastern Indonesia): Some aspects of its grammar*. Hamburg: Helmut Buske Verlag.
- Shelden, Howard. 1991. Galela pronominal verb prefixes. In Tom Dutton (ed.), *Papers in Papuan Linguistics 1*, 161-75. Pacific Linguistics A, 73. Canberra: Australian National University
- Shelden, Howard. 1998. Transitivity and Galela pronominal reference. SIL Electronic Working Papers 1998-005. <http://www.sil.org/silewp/1998/005/>
- Staden, Miriam van. 2000. *Tidore, A linguistic Description of a Language of the North Moluccas*. PhD dissertation. Leiden University.
- Steinhauer, Hein. 1995. Two varieties of the Blagar language (Alor, Indonesia). In: Connie Baak, Mary Bakker and Dick van der Meij (eds.), *Tales from a concave world: Liber amicorum Bert Voorhoeve*. Leiden University: Department of Languages and Cultures of SE Asia and Oceania: 265-296.
- Steinhauer 2005. Biak. In: K. Alexander Adelaar and Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, London: Routledge.
- Tsukida, Naomi. 2005. Split intransitivity in Seediq and Amis. Paper presented at the conference 'The Typology of Stative-Active Languages', Max Planck Institute Leipzig, May 2005.
- Uhlenbeck, E.M. 1978. *Studies in Javanese morphology*. [Translation of publication in Dutch in 1949]. The Hague: Martinus Nijhoff.
- Verheijen, Jilis A.J. 1986. *The Sama/Bajau language in the Lesser Sunda Islands*. Canberra: Pacific Linguistics D-70.
- Wimbish, Sandra G. 1991. An introduction to Pagu through the analysis of narrative discourse. M.A. thesis. University of Texas at Arlington.

## Appendix

### On the selection of the sample

Several factors determined which languages were included in this sample. First, the languages of Kalimantan and the Papuan mainland were not included to keep the sample's size and scope manageable. Second, to have a genetically heterogeneous sample, I included both Austronesian (AN) and non-Austronesian (NAN) languages (28 AN, 11 NAN).<sup>12</sup> (Note that the sample does not contain any Oceanic language, because all of these are spoken outside the Indonesian area.) The non-Austronesian ('Papuan') languages in the sample are spoken on islands in eastern Indonesia, to the far West of the Papuan mainland (Halmahera, Alor, Pantar, Timor). They were selected as representatives of distinct non-Austronesian families (West Papuan (WP), Trans New Guinea (TNG), and Geelvink Bay (GB)). Third, since I was also interested in the geographical spread of SS in Indonesia, for the statistic tests reported in section 5, I divided the sample into two major areas: a Western part, including Sumatra, Java, Madura, Bali, and Lombok, and an Eastern part: Flores, Bima, Sumba, Sulawesi, Alor, Pantar, Timor and surrounding islands, Halmahera, C/S Maluku, and the islands off the coast of W Papua. The islands in the Eastern part host 313 languages, those of the Western part, 72.<sup>13</sup> To have an proportional representation of both areas in the sample, 32 eastern languages and 7 western languages were included.

---

<sup>12</sup> The Austronesian languages in the sample belong to the three Austronesian subgroups that are present in Indonesia: West Malayo-Polynesian, Central Malayo-Polynesian and South Bird's Head-West New Guinea. Since these subgroups are not relevant for the present paper they are not indicated in Table 1.

<sup>13</sup> See the SIL Ethnologue, [www.ethnologue.com](http://www.ethnologue.com)

Table 1. Semantic alignment in the Indonesian language sample

	-SS	+SS	An/NAN	
<b>Sumatra</b>				
Karo Batak	x		AN	Woollams 1996, 2005
Nias		x	AN	Brown 2000, 2005
Acehnese		x	AN	Durie 1985
<b>Java, Madura</b>				
Javanese	x		AN	Uhlenbeck 1949/1978; Oglobin 2005
Sundanese	x		AN	Muller-Gotama 2001
Madurese	x		AN	Davies 1999
<b>Bali, Lombok</b>				
Balinese	x		AN	Arka 2003
<b>Flores, Bima, Sumba</b>				
Ngadha	x		AN	Djawanai 1983
Bimanese	x		AN	Owens 2000
Keo	x		AN	Baird 2001
Kedang		x	AN	Samely 1991
Lamalera		x	AN	Keraf 1978
Kambera		x	AN	Klamer 1998; in press
<b>Sulawesi</b>				
Muna	x		AN	Van den Berg 1989
Tukang Besi	x		AN	Donohue 1995
Bajau	x		AN	Donohue 1996b, Verheijen 1986
Mori Bawah		x	AN	Mead 2005
<b>Alor/Pantar</b>				
Blagar	x		TNG	Steinhauer 1993/1999
Teiwa	x		TNG	Klamer, in preparation
Alorese	x		AN	Klamer, in preparation
Klon		x	TNG	Baird 2004, 2005
Abui		x	TNG	Kratochvíl, in preparation
Tanglapui		x	TNG	Donohue 1997
<b>Timor archipelago</b>				
Tetun Fehan	x		AN	Van Klinken 1999
Mambai	x		AN	Hull 2001
Kemak	x		AN	Hull 2001
Makasai	x		TNG	Brotherson 2003
Leti	x		AN	Van Engelenhoven 2004
Bunak	x		TNG	Friedberg 1978
<b>Halmahera</b>				
Tidore	x		WP	Van Staden 2000
Taba		x	AN	Bowden 2001
Tobelo		x	WP	Holton 1997
Pagu		x	WP	Wimbish 1991
<b>C/S Maluku</b>				
Buru	x		AN	Grimes 1991
Dobel		x	AN	Hughes 2000
Larike		x	AN	Laidig and Laidig 1991, Laidig 1992
Selaru		x	AN	Coward & Coward 2000
<b>NE of Bird's Head, Papua</b>				
Biak	x		AN	Steinhauer 2005, Van den Heuvel, in preparation
Saweru		x	GB	Donohue 2001

**Table 2. Summary of Table 1 – The areal spread of SS in the sample**

No. of lgs in Indonesia excluding Borneo and Papua	West		East		Total	
		72		313		385
sample languages (see list in Table 1)	7	+SS 2	32	+SS 14	39	+SS 16
		-SS 5		-SS 18		-SS 23

**Table 3. Data from table 1-2, used for Fisher's Exact Test**

	[+SS]	[-SS]
[West]	2	5
[East]	14	18

The preceding document was presented at the Tenth International Conference on Austronesian Linguistics (10-ICAL). To properly reference this work, please use the following format:

<LastName>, <FirstName>. 2006. <PaperTitle>. Paper presented at Tenth International Conference on Austronesian Linguistics. 17-20 January 2006. Puerto Princesa City, Palawan, Philippines.  
<http://www.sil.org/asia/philippines/ical/papers.html>

For other papers that were presented at 10-ICAL, please visit <http://www.sil.org/asia/philippines/ical/papers.html>.