Some observations on word order in western Austronesian symmetrical voice languages

Sonja Riesberg
Universität zu Köln & Australian National University

1. Introduction

Western Austronesian languages are well known for their complex and typologically rather unusual voice systems. These consist of two or more basic transitive constructions which are symmetrical in the sense that they are equally morphologically marked and that arguments retain core status in all voices (i.e. no argument demotion as, e.g., in the active-passive alternation, takes place). Symmetrical voice systems thus do not exhibit the same kind of 'syntactic agent prominence' observed in many other languages, in which the agent is usually the default, unmarked choice for the syntactically privileged argument (PSA)[1]. Yet, agent prominence has been claimed to be a universal property in the morpho-syntactic organization in the languages of the world. In addition to occupying the PSA function, agents often occur in prominent position within the clause (e.g. sentence initial), they bear unmarked/less marked case (nominative), and they are often the only argument the predicate agrees with. The well-known semantic role hierarchies that have been postulated in the typological literature in various ways (cf. two versions in (1) and (2)), are one way to capture the observations on the privileged status of agent arguments.

(1) agent > benefactive/goal/experiencer > patient/theme
(e.g. Jackendoff 1972; Givón 1984; Grimshaw 1990; Pesetsky 1995)

(2) agent > patient/theme > benefactive/goal/experience
(e.g. Dik 1978; Larson 1988; Baker 1996)

While there has been quite some variation and disagreement with respect of the ordering in the lower positions on the scale (the two proposals in (1) and (2) already show a divergence whether, e.g., to place the experiencer before the patient, or the other way round), it has never been questioned that the most agent-like argument takes the highest position. Some authors have postulated the reverse order, i.e. patient > agent, for ergative languages (e.g. Dowty 1991; Van Valin & LaPolla 1997), but the assumption that ergative systems are just mirror images of accusative systems have been criticized and refuted in the literature (cf. e.g. Mithun & Chafe 1999 for a detailed argumentation against this view). Recent neurolinguistic experimentation, furthermore, appears to provide support for the claim actors are universally more prominent in language processing, regardless of the grammatical organization of grammatical relations. Thus, Bickel et al. (2015) find a processing bias favouring actors, even in languages with ergative systems, like Hindi.

In this paper, I will investigate the question whether Austronesian symmetrical voice languages – despite their symmetry of actor and undergoer arguments in terms of linking possibilities – show any evidence in favour for the claim that actors are more prominent than undergoers. More precisely, I will look at whether actor prominence in symmetrical voice languages is reflected in word order constraints, in the sense that agent arguments more

[1] The PSA is defined as the syntactic element, that "controls" coding properties, such as agreement, and which is the central element in complex constructions such as relativization, NP deletion, control, etc. In many respects, it thus corresponds to the traditional concept of subjecthood. However, unlike a subject, a PSA is defined construction-specifically. In this paper, we use the terms PSA and SUBJECT in a roughly synonymous fashion, as we assume that subjects also have to be determined language- and construction specifically.

frequently occur before all other arguments, i.e. that there is a tendency for 'agent-first' patterns.[2] Obviously, actor > undergoer, or 'agent-first', is not expected to be in any case a hard constraint in Austronesian symmetrical voice language. Anyone familiar with these languages will easily come up with counter examples to this claim. And in fact, section 2 of this paper will introduce the two most common word order patterns of symmetrical voice languages, showing that both these patterns do *not* exhibit an 'agent-first' pattern.

Our hypothesis is therefore that 'agent-first' word order is manifest in a more subtle way and on a more global (i.e. cross-linguistic) level. That is, it is predicted that if whenever there are deviations from the standard patterns introduced in section 2, such deviations will be biased towards an 'agent-first' order.

The focus of this paper will be on word order preferences of full noun phrases, including free pronouns, in voice marked constructions. Occasional reference will also be made to the ordering of pronominal clitics.

2. Two "standard" patterns of word order – Totoli and Tagalog

In this section, I will use two languages, Totoli and Tagalog, to exemplify two different word order patterns that commonly occur in Austronesian symmetrical voice languages and that therefore can be considered representatives of the two "standard" word order patterns found in these languages. As we will see, both patterns do not reflect any kind of agent prominence in the sense that the actor would always occur in either initial or final (i.e. prominent) position in the clause, or in that the actor would always precede the undergoer.

Totoli and Tagalog are both symmetrical voice languages and can be considered to be representatives of 'Indonesian type' and 'Philippine type' languages, respectively. It should be stressed, though, that these terms have to be used with some caution. The latter (which is probably slightly better established than the former) has been defined to subsume languages which exhibit the following characteristics (Himmelmann 2005a: 113):

- at least two formally and semantically different symmetrical undergoer voices
- at least one non-local phrase-marking clitic for nominal expressions
- pronominal second position clitics

Indonesian-type languages, like Philippine-type language, display symmetrical voice systems, often with two or three symmetrical alternations. As defining characteristics, intensive use of applicative morphology and pro-clitic actor markers in at least some slots of the voice paradigm have been proposed (cf. Himmelmann 2005a: 175), but it stands to reason that this group of languages is actually a very heterogeneous group of languages primarily defined by the fact that they are symmetrical voice languages not meeting the case marking and second position clitic criteria defining Philippine-type languages.

2.1 Word order in Totoli – the flexible subject+VP-pattern

Totoli is a Western Malayo-Polynesian language spoken in the northern part of Central Sulawesi, Indonesia. Given the definitions above, Totoli can be considered an Indonesian-type language. It possesses one actor voice and two undergoer voices. The latter, though formally different, are lexically determined and unlike in Philippine-type languages not semantically distinct, i.e. in both undergoer voices either a patient or a theme argument is linked to subject position. Noun phrases are not marked, and there is no distinct second

[2] Cf. Riesberg & Primus (2015) for a first brief exploration of actor prominence in symmetrical voice languages with respect to reflexive binding, and the use of stative and potentive morphology.

position slot for pronominal clitics. Instead, Totoli exhibits a complex interplay of voice and applicative morphology (cf. Himmelmann & Riesberg 2013) and in the undergoer voice, first person actors are marked by a pro-clitic.

Totoli displays two basic word order patterns, SVO and VSO in both actor and undergoer voice. The verb and the non-subject argument form a VP with fixed ordering[3], while the position of the subject NP is flexible in that it can occur either before or after the verb non-subject complex. Both orders are syntactically equal. Crucially, the pre-verbal subject slot is clause internal, i.e. in these cases the subject NP is not topicalized/left-dislocated, as it commonly happens in strict verb-initial languages, like, e.g. Tagalog (see below). In the following, this pattern will be called the *flexible subject+VP-pattern*. Examples (3) and (4) illustrate the two word order possibilities in the actor voice and the undergoer voice, respectively. In all examples, the actor argument is marked in bold.

(3) a. [***kita***] *majaga* [*isia*]
kita moN-jaga isia
1p AV-guard 3s
'we look after him' [Mansur's_work.1189]

b. *ana* *mogutumo* [*gauan*] [***kita***]
ana mo-gutu=mo gauan **kita**
if AV-make=CPL garden **1p**
'if we want to make a garden…' [monkey_turtle.063]

(4) a. [*buta*] *nibangun* [***tau*** ***pomoo***]
buta ni-bangun **tau** **pomoo**
earth RLS-stand.up **person** **first**
'the earth was cultivated by the ancestors' [tatabua.93]

b. *patei* [***kamuu***] [*tau*]
pate-i **kamu** tau
kill-UV2 **2s** person
'you killed a person' [tau_bentee.206]

As can be seen from the examples above, the flexible subject-VP-pattern and the symmetrical nature of the Totoli voice system result in the fact that agent prominence is *not* reflected in word order. While in the actor voice the agent argument *can* occur in sentence initial position and thus before the undergoer (cf. (3)a.), it can just as well be realized in sentence final position, then following the undergoer argument. In the undergoer voice, on the other hand, the agent can never occur in sentence initial position, as this position is reserved for the undergoer subject. As the agent occupies the function of the non-subject argument, it will always be realized right after the verb. The undergoer subject, if occurring sentence initially can thus precede the agent, as in (4)a., or, occurring sentence finally, follow it (cf. (4)b.).

In both voices we thus find both options: agent before undergoer, and undergoer before agent. It remains to be see how these patterns distribute in discourse and whether a corpus study would reveal preferences for one pattern over the other. In particular, it may very well turn out that in actor voice constructions subject-initial order is preferred (hence actor preceding

[3] With some exceptions not further discussed here, no other constituents can occur between verb and non-subject argument.

undergoer) while in undergoer voice constructions the reverse preference holds, i.e. subject-final order is preferred, hence undergoer following actor.

Totoli has two sets of pronouns, one set occurs as free forms, the other one as pronominal clitics. (3) and (4) already exemplified some of the free forms. Table 1 summarizes the two paradigms under the label 'nominative' and 'genitive'. In actor voice constructions, the nominative set has to be used for both subject and as non-subject arguments (cf.(3)). In undergoer voice constructions, the nominative paradigm has to be used for pronominal arguments in subject function, while non-subject arguments take the genitive paradigm. Note that in the plural, nominative and genitive forms partially overlap, i.e. second and third person non-subject pronouns can be realized either as free forms or as pronominal clitics, and for the first person exclusive no clitic form exists. The clitic *=na* is predominantly used for third person singular, with *sisia* being its plural counterpart. In some instances, however, *=na* can also be found to refer to third person plural actors.

	NOMINATIVE	**GENITIVE**
1SG	*aku*	*=ku; ku-*
2SG	*kau*	*=mu; =ta*
3SG	*isia*	*=na*
1PL EXCL	*kami*	*kami*
1PL INCL	*kita*	*=ta*
2PL	*kamu*	*kamu; =ta*
3PL	*sisia*	*sisia*; (*=na*)

Table 1: Totoli pronouns

The form of referential expressions, i.e. whether an argument is expressed by a common noun, a free pronoun, or a pronominal clitic, does not have an effect on word order possibilities in Totoli. The examples above have already shown that free pronouns show the same distribution as common nouns. The same also holds for pronominal clitics. As the non-subject agent of an undergoer voice construction, they are cliticized to the verb stem. The undergoer subject, again, can occupy the sentence initial slot or the sentence final slot, again gaining both patterns undergoer > actor ((5)a.) and actor > undergoer ((5)b.).

(5) a. *aku* *kodoong* *sukatina*
aku ko-doong sukat-i**=na**
1s POT-want try-APPL2**=3s.GEN**
'I am to be challenged by him' [lelegesan_7.018]

b. *kalambotimu* *aku*
ko-lambot-i**=mu** aku
POT-remember-UV2=**2s.GEN** 1s
'you remember me.' [farming_2.2037]

2.2 Word order in Tagalog – the verb initial+final subject-pattern

Tagalog exhibits all the above mentioned characteristics of a Philippine-type language. It exhibits one actor voice and three semantically distinct undergoer voices: In the patient voice, the PSA usually bears the semantic role of the patient, in the so-called locative voice a local, typically a goal or source, argument is linked to PSA function, and in the so-called conveyance voice the PSA argument can either be a theme, a benefactive, or an instrument.

Common nouns are preceded by one out of three case markers: *ang* always marks the PSA, *ng* [nang] and *sa* occur with non-PSA arguments[4]. The examples in (6) illustrate the four different voices and the case marking of arguments. Again, actor arguments are marked in bold.

(6) a. *Bumabasa* [*ng* *diyaryo*] [***ang*** ***titser***].
-um-RDP-basa ng diyaryo **ang** **titser**
-AV-RDP-read GEN newspaper **NOM** **teacher**
'The teacher is reading a newspaper.' [Schachter/Otanes 1972: 69]

b. *kinain* [***ng*** ***pusa***] [*ang* *daga*]
-in-kain-ø **ng** **pusa** ang daga
-RLS-eat-PV **GEN** **cat** NOM rat
'the cat ate the rat'[5] [Kaufman 2015: ??]

c. *iniabot* [***ng*** ***manggagamot***] [*sa* *sundalo*] [*ang* *itlog*]
-in-i-abot **ng** **manggagamot** sa sundalo ang itlog
-RLS-CV-reach **GEN** **doctor** DAT soldier NOM egg
'the physician handed the egg to the soldier.' [Himmelmann 2008: 265]

d. *kinainan* [***ng*** ***pusa***] [*ng* *daga*] [*ang* *pinggan*]
-in-kain-an **ng** **pusa** ng daga ang pinggan
-RLS-eat-LV GEN cat GEN rat NOM plate
'the cat ate the rat on/from the plate' [Kaufman 2015: ??]

Unlike Totoli, Tagalog word order is strictly verb initial, i.e. it does not provide for a clause internal argument slot that precedes the verb. Arguments follow the verb and while the order of NP arguments can be considered to be free to a certain extent, there are strong tendencies: In pragmatically unmarked contexts, the *ang* phrase always occurs in sentence final position. The dative-marked phrase occurs before the *ang*-phrase; and the genitive-phrase usually immediately follows the predicate (Himmelmann 2005b: 357). I will call this pattern the *verb initial+final subject-pattern*. Again, just like in Totoli, the combination of word order constraints and the symmetry of the voice system prevents any obvious reflection of agent prominence in the ordering of NPs. In the different undergoer voices, the agent usually precedes all other arguments and could thus be considered be in prominent position, but in the actor voice, the agent argument follows the semantically lower ranked arguments.

Tagalog shows different constraints, though, for (free) pronouns, which are second position clitics, and proper nouns, which do reflect agent prominence. That is, if the actor is realized as a pronoun, it will always occur in direct post-verbal position. In these cases, the actor thus precedes all other arguments, irrespective of the voice of the construction, as shown in the examples in (7) (all examples taken from Himmelmann 2005b: 366f.). The same also holds, if the agent is expressed by a proper name, as illustrated in (8)a. for an actor voice construction.

[4] The choice of the *ng* and *sa* determined by the semantic role of the case marked argument: *ng* is used for marking actors, experiencers, patients and themes, *sa* marks goals, recipients/benefactives, and locatives. In this paper, *ang*, *ng*, and *sa* will be glossed as NOMINATIVE, GENITIVE, and DATIVE, respectively.

[5] Kaufman discusses this example under the hypothesis that the different voices in Tagalog involve different participant nominalizations. He therefore translates the examples in (6)b. and (6)d. as, 'the rat was the eaten one of the cat' and 'the plate was the cat's eating place of the rat', respectively.

Note, however, that in this case this is a tendency only and that the reverse pattern is also possible, as illustrated in (8)b.

(7) a. *nag-da-dala* ***sila*** *ng* *sarili* *nila-ng* *banda* *ng* *musika*
RLS.AV-RDP-bring **3p.NOM** GEN own 3s.POSS-LK band GEN music
'they bring their own band'

b. *p<in>atay* ***natin*** *ito-ng* *dalawa-ng* *Hapon*
<RLS(PV)>dead **1pi.GEN** PRX-LK two-LK Japan
'we killed these two Japanese'

c. *ibinalík* ***nila*** *ang* *bata*
i-in-balik **nila** ang bata
CV-RLS-return **3p.GEN** NOM child
'they returned the child'

d. *tinirhan* ***ko*** *ang* *bahay* *na* *ito*
in-tira-an **ko** ang bahay na ito
RLS-dwelling-LV **1s.GEN** NOM house LK PRX
'I stayed at this house'

(8) a. *bumili* ***si*** ***Rosa*** *ng* *bigas*
b-um-ili **si** **Rosa** ng bigas
-AV-buy **NOM** **PN** GEN rice
'Rosa bought some rice' [Schachter/Otanes 1972: 81]

b. *bumili* *ng* *bigas* ***si*** ***Rosa***
b-um-ili ng bigas **si** **Rosa**
-AV-buy GEN rice **NOM** **PN**
'Rosa bought some rice' [Schachter/Otanes 1972: 81]

3. Common deviations from the two standard patterns

The previous sections described two basic patterns of word order restrictions commonly found in Austronesian symmetrical voice languages, which I called the *flexible subject+VP-pattern* and the *verb initial+final subject-pattern*. Representatives of these two types are Totoli, and Indonesian-type language of Sulawesi and Tagalog, a Philippine-type language of the Philippines. In both patterns, there no general preference for agent first word order, though we saw that in Tagalog pronominal and proper name agents (tend to) precede all other arguments.
This section will now focus on patterns that deviate from the two patterns described for Totoli and Tagalog, starting with the former, i.e. with the deviations from the *subject+VP pattern*.

3.1 Deviations from the flexible subject+VP-pattern

3.1.1 Allowing for two agent first-positions in AV

One deviation from the two 'standard' patterns illustrated in the preceding section, can be observed e.g. in Bajau (Sabah), Kelabit (Sarawak) and Pitu Ulunna Salu (Sulawesi). These

languages allow for three different orderings in actor voice constructions two of which put the actor before the undergoer. That is, in addition to the ordinary SVO (and VOS) pattern, these languages also exhibit VSO order in actor voice constructions, as illustrated in (9) for Bajau, in (10) for Pitu Ulunna Salu, and in (11) for Kelabit (Sarawak).

(9) BAJAU (Sabah)

a. *boi* *moo* ***Amzi*** *bua'* *nangka'* *e*
CPL AV.bring **PN** fruit jackfruit DEM
'Amzi brought the jackfruit.' [Miller 2007: 150]

b. *boi* *nguse'* ***iyo*** *kerita'* *Amzi*
CPL AV.clean **3s.II** car PN
'he cleaned Amzi's car' [Miller 2007: 151]

(10) PITU ULUNNA SALU (Sulawesi)

um-batta-m-äq ***kao*** *bittiq-ku*
AV-cut-PRF-1s **1s** foot-1s
'I cut my foot' [Campbell 1989: 130]

(11) KELABIT (Sarawak)

ne-kuman ***la'ih*** ***sineh*** *buaq* *kaber*
PFV-AV:eat **man** **DEM** fruit pineapple
'the man ate pineapple' [Hemmings 2016: 448]

Importantly, no additional VSO order is reported for undergoer voices, which would position the undergoer subject before the actor.

Artawa (1998: 19) and Mayani (2013: 172) give examples for VSO order in actor voice constructions for Balinese and Tajio respectively. However, both argue, that this order is marked and only possible, if there is contrastive focus on the verb. In the Tajio example below, the speaker thus apparently wants to stress the fact that they *bought* the durian, instead of, e.g. stealing or selling it (Mayani 2013: 172). These pragmatic cases of word order variability will not be further considered in this paper.

(12) TAJIO (Sulawesi)

nongoli ***sisia*** *teruriang*
noN-oli **sisia** te=ruriang
AV.RLS-buy **3p** NM=durian
'they *bought* durian' [Mayani 2013: 172]

3.1.2 Ban of VOS in AV

Another deviation from the flexible subject+VP-pattern that can be observed in a few languages (mostly in Sulawesi?) is the ban of VOS order in the actor voice. That is, in these languages (cf., e.g., Tondano and Rampi'), the symmetrical word order pattern between actor voice and undergoer voice is given up, and in the actor voice, we now find a strict agent-first pattern, i.e. only SVO is possible. In the undergoer voice(s), the same two options as in Totoli, i.e. SVO and VOS are available (cf. the Tondano examples in (13)).

(13) TONDANO (Sulawesi)[6]

a. ***sioki'ku*** *mateles* *raaren* **AV: SVO**
si=oki'=ku ma-teles raaren
AN.SG=small.NOM=1s.GEN AV-buy vegetable
'my child buys a/some vegetable(s)' [Brickell 2014: 140]

b. *toto'* *nipèrèt* *kinaan* ***nitim*** **UV: SVO**
toto' ni=pèrèt k<in>aan **ni=Tim**
breast.NOM AN.SG.GEN=bat <PST>eat **AN.SG.GEN=PN**
'Tim ate the bat's breast' [Brickell 2014: 148]

c. *wo* *tu* *pemurkimurkiten* ***nitim*** *kokong* **UV: VOS**
wo tu peN-CVCV-purkit-en **ni=Tim** kokong
and then DYN-RDP-turn.over-PV **AN.SG.GEN=Tim** head
'and then Tim is turning over the head' [Brickell/Schnell 2017: 184]

Gayo (Sumatra) is another language that does not allow for SOV in actor voice clauses, while having both choices SOV and VOS available in the undergoer voice (Eades 2005: 120), as illustrated in (14). However, this constraint is restricted to highly transitive clauses. In less transitive contexts where the undergoer receives a generic reading, or is expressed by a prepositional phrase, the actor subject can occur after the VO complex, or in direct post-verbal position (see (15)a. and (15)b. respectively). In the latter case, actor again precedes undergoer.

(14) GAYO (Sumatra)

a. ***aku*** *munemahè*
aku mun-emah=è
1s AV-make=3.GEN
'I make/am making it' [Eades 2005: 174]

b. * *munemahè* *aku*
mun-emah=è aku
AV-make=3.GEN 1s
for: 'I make/am making it' [Eades 2005: 174]

(15) GAYO (Sumatra)

a. *munyuen* *kepile* ***kami*** *i* *uken* *so*
AV:plant sweet.potao **1pe** LOC upstream yon
'we are planting sweet potatoes upstream' [Eades 2005: 173]

b. *muninget* ***aku*** *kin* *masa* *Jepang*
mun-inget **aku** kin masa Jepang
AV-remember **1s** DAT era Japanese
'I recall the era of the Japanese (occupation)' [Eades 2005: 173]

[6] Brickell does not provide any ungrammatical examples in his grammar of Tondano. Yet, he explicitly states that "In AV transitive clauses the NOM argument has a fixed pre-predicate position. In UV transitive clauses the NOM may either precede or follow the predicate…" (Brickell 2014: 94). We can thus assume, that a constructions like **mateles raaren sioki'ku* for 'my child buys a/some vegetable(s)', is ungrammatical.

For other languages which exhibit both, SVO and VOS in actor voice, it is sometimes stated that SVO is the 'preferred' pattern (cf. e.g. Clayre 2014: 132 for Lundaye).

3.1.3 Ban of VOS and two agent-first positions in AV

In other languages, we find a combination of the two ordering restrictions described in the two sections above, i.e. a ban of VOS in actor voice constructions, *and* the availability of two agent-first positions in AV. Begak (Sabah) is a representative of this type, which thus displays different patterns in AV and in UV: while both voice constructions allow for the verb medial ordering, i.e. SVO, the alternative order in AV is VSO but VOS in UV. Goudswaard describes Begak as basically displaying two word orders: "(i) The verb-initial word order is semantically based and is Verb-Actor-Undergoer, irrespective of the voice marking of the verb. (ii) The subject-initial or verb-medial word order is syntactically based and is Subject-Verb-Object, irrespective of voice marking of the verb" (Goudswaard 2005: 125). The data in (16) illustrate these patterns for both AV (a. & b.) and UV (c. & d.).

(16) BEGAK (Sabah)

a. ***Pius*** (*da*) *gədagang* *pait* *di'* *Dəngon*
Pius da gə-dagang pait di' Dəngon
PN PR AV-buy fish LOC PN
'Pius is buying fish in Dəngan'

b. (*da*) *gədagang* ***Pius*** *pait* *di'* *Dəngon*
da gə-dagang **Pius** pait di' Dəngon
PR AV-buy **PN** fish LOC PN
'(…) Pius is buying fish in Dəngan, (…)'

c. *pait* *ino* *degang* ***Pius*** *di'* *Dəngon*
pait ino -i-dagang **Pius** di' Dəngon
fish yonder -CPL-buy:UV **PN** LOC PN
'this fish was bought by Pius in Dəngan'

d. (*bay*) *degang* ***Pius*** *pait* *di'* *Dəngon*
bay -i-dagang **Pius** pait di' Dəngon
PRF -CPL-buy:UV **PN** fish LOC PN
'this fish was bought by Pius in Dəngan' [Goudswaard 2005: 126]

Goudswaard states, that SVO, rather than VSO is the preferred word order in actor voice constructions, and the only possible order that can be used "as an opening sentence of a story or conversation" (2005: 126). In the undergoer voice, on the other hand, VOS is the preferred choice, while SVO order is the marked choice. Just like in Tondano (and the other languages mentioned in section 3.1.2), Begak thus displays a strict agent-first in AV, and also in the UV there seems to be a preference to use actor > undergoer rather than the other way round. Like in Totoli, pronouns reflect case distinctions (nominative vs. genitive), but do not adhere to any specific ordering constraints.

3.2 Deviations from the verb initial+final subject-pattern

3.2.1 Strict agent-first

Dilon states for Tatana' (Sabah) that "the normal word order within a clause is: Verb Genitive Nominative Dative" or "Verb Actor Subject Object/Oblique" (Dilon 1994: 65). He goes on explaining that "because the actor is the subject in actor voice, it takes the nominative case and there is no genitive", resulting in the pattern 'Nominative Dative (Dative)', i.e. actor > undergoer. Most examples Dilon cites involve either pronominal or proper name actor subject. As illustrated for Tagalog in section 2.2 these types of nominal expressions often behave differently from common nouns in their ordering properties. Yet, note the example (17)c. which involves three common noun arguments and thus provide evidence for strict agent-first word order in Tatana'.

(17) TATANA' (Sabah)

a. *mananda'* ***isio*** *do* *mija'*
mon-tanda **isio** do mija'
AV-make **3s** DAT table
'he is making a table' [Dilon 1994: 43]

b. *mopoguli'* ***i*** ***Gaman*** *do* *buuk* *di* *amai* *no*
mopo-guli' **i** **Gaman** go buuk di amai no
AV.CAU-return **NOM** **PN** DAT book DAT uncle 3s.GEN
'Gaman is giving a book back to his uncle' [Dilon 1994: 65]

c. *moporatu'* *nio* ***tukang*** *do* *papan* *intad* *do* *sowat*
mopo-ratu' *nio* ***tukang*** *do* *papan* *intad* *do* *sowat*
AV.CAU-fall there **craftsman** DAT board from DAT top
'the carpenter is going to drop the planks from the top' [Dilon 1994: 69]

Other languages (of the Philippine type) that exhibit strict agent first word order are Limos Kalinga (cf. Ferreirinho 1993: 59), Takivatan Bunun (De Busser 2009: 101, but section 3.2.3 for Takbanuad Bunun), Mansaka (Svelmoe & Svelmoe 1974), and Agat (Healey 1950).

Ibaloy shows strict agent first order in the undergoer voices (Ruffolo 2004: 417), while in the actor voice ordering seems to be more flexible. Ruffolo provides one example, to show that the nominative argument, if a full NP, "has a relatively free word order with respect to other complements" (Ruffolo 2004: 370).

(18) IBALOY (Philippines)

a. *dimaw* ***i*** ***daki*** *chi* *Kabayan*
<im>law **ʔi** **laki** di Kabayan
<AV.PFT>go **NOM** **man** LOC PN
'the man went to Kabayan' [Ruffolo 2004: 370]

b. *dimaw* *chi* *Kabayan* ***i*** ***daki***
<im>law di Kabayan **ʔi** **laki**
<AV.PFT>go LOC PN **NOM** **man**
'the man went to Kabayan' [Ruffolo 2004: 370]

However, also for AV clauses, Ruffolo states that "preferred constituent order is for the Nominative to follow the verb and precede the E complements (extension-to-core complement, SR)" (Ruffolo 2004: 414). Thus, even though word order is more flexible in AV than in UV, actor > undergoer seems to be the unmarked order in both constructions. Other languages that show agent-first order in UV only (but not in AV) are, e.g., Chotabato Manobo (Philippines, Kerr: 1988).

3.2.2 Obligatory post-verbal actor clitic in AV

A variant of the strict agent-first pattern described in section 3.2.1 is the obligatory occurrence of post verbal actor clitic in the actor voice. For Tboli (Philippines), for example, Forsberg reports a fixed word order of "verb actor (goal) object" (1992: 56). This order holds for all undergoer voices, as illustrated in the examples in (19).

(19) TBOLI (Philippines)

a. *hnebel* ***yem*** ***maen*** *yem* *libun* *leged*
h-n-eben **yem** **maen** yem libun leged
-PV-look.for **that** **father:3s.GEN** that girls industrious
'his father is looking for an industrious girl' [Forsberg 1992: 72]

b. *benlay* ***Ben*** *ou* *snafang*
b-en-lay **Ben** ou snafang
-BV-give **PN** 1s gun
'Ben gave a gun to me' [Forsberg 1992: 79]

c. *ofok* ***Walan*** *du* *asay*
ø-ofok **Walan** du asay
IV-chop.down **PN** it ax
'Walan chopped it down with an ax' [Forsberg 1992: 81]

In actor voice constructions, pronominal subjects take the post-verbal slot, in a similar way as we have seen it for Tagalog, as shown in (20)a. Common noun actor subjects have to be placed in sentence final position, just like their undergoer subject counterparts in the undergoer voices. However, they have to be obligatorily cross-referenced by a co-referential pronoun, which occurs in direct post-verbal position. The first slot in the ordering of arguments is thus always occupied by an element that refers or cross-references to the actor.

(20) TBOLI (Philippines)

a. *mbel**e*** *sotu* *lata* *halay*
m-bel=**e** sotu lata halay
AV-look.for=**1s** one can unhulled.rice
'I'm looking for one can of unhulled rice' [Forsberg 1992: 64]

b. *smakay* ***le*** *ówóng* *yó* ***kem*** ***ngà***
s-m-akay **le** ówóng yó **kem** **ngà**
-AV-ride **3p** airplane that **p** **child**
'the children rode in that airplane' [Forsberg 1992: 63]

3.2.3 Obligatory actor agreement in UV (and optional in AV)

A slightly different form of actor agreement from the one just described for Tboli, can be observed in the Formosan language Puyuma (Taiwan). While in the actor voice we find the "standard" order (as described for Tagalog), i.e. VOS (see (21)a.), the undergoer voice shows deviation from the Tagalog pattern, in that the actor argument is obligatorily cross-referenced by a genitive proclitic. A co-referent actor NP can optionally be realized, following the undergoer subject, which occurs in post-verbal position, as in (21)b.

(21) PUYUMA (Taiwan)

a. *trakaw* *dra* *paisu* ***i*** ***Isaw***
<AV>steal ID.OBL money **NOM.S** **PN**
'Isaw stole money' [Teng 2008: 109]

b. ***tu=****padrek-aw* *i* *temutaw* ***kana*** ***walak***
3.GEN=carry.on.back-PV NOM.S grandparent:3.GEN **DF.OBL** **child**
'the child carried his grandmother on his back' [Teng 2008: 150]

In Tsou, another Formosan language, word order is fixed VOS, in both actor voice and undergoer voice constructions. In both voices, the matrix verb is usually preceded by a preverb of which there are two sets: one for actor voice constructions (*mio, mi, moh, moso*), and one for undergoer voice constructions (*i, o, os, oh*) (cf. Tung 1964: 52). These markers carry tense information, and if the agent (in both AV and UV) is realised pronominally, it is cliticised to this preverbal element. With pronominal actors, Tsou thus displays strict agent-first ordering (cf. (22) a. and b.). If the agent is realised as a full NP, the additional use of the pronominal actor enclitic is obligatory in the undergoer voice, but not in the actor voice (compare (22)c. with d. and e.).

(22) TSOU (Taiwan)

a. *mi-****ta*** *mimo* *ta* *cxumu*
AV:NON.PST-**3s** AV:drink OBL water
'he is drinking some water' [Zeitoun 1992: 11]

b. *i-****ta*** *ima* *si* *cxumu*
UV:NON.PST-**3s** UV:drink NOM water
'the water has been drunk by him' [Zeitoun 1992: 11]

c. *mo* *mimo* *ta* *cxumu* ***si*** ***amo***
AV:NON.PST AV:drink OBL water **NOM** **father**
'my father is drinking some water' [Zeitoun 1992: 11]

d. **i* *ima* *ta* *amo* *si* *cxumu*
UV:NON.PST UV:drink OBL father NOM water
for: 'the water has been drunk by my father' [Zeitoun 1992: 11]

e. *i-****si*** *ana* ***to*** ***amo*** *ʔo* *tacUmU*
UV:NON.PAST-**3s** UV:eat **OBL** **father** NOM banana
'father has eaten a banana' [Zeitoun 1992: 4]

3.2.4 Agent-first if case marking is absent

Again another variation of the agent-first pattern can be found (in the Takbanuad dialect of) Bunun (Taiwan), where word order is flexible if noun phrases are marked by case particles, but where order is fixed to actor > undergoer if case marking is absent. Thus, in the following examples, the nominative marked actor *a Paiʔan* can either precede or follow the accusative marked undergoer *i titiʔ* in (23)a. and b. For the bare nouns, however, the argument in post-verbal position has to be interpreted as the actor, and thus (23)c. is ungrammatical for the reading 'Paiʔan is eating meat'.

(23) TAKBANUAD BUNUN (Taiwan)

a. *ma'un* ***a*** ***Paiʔan*** *i* *titiʔ*
AV:eat **NOM** **PN** ACC meat
'Paiʔan is eating meat' [Jeng 1977: 285]

b. *ma'un* *i* *titiʔ* ***a*** ***Paiʔan***
AV:eat ACC meat **NOM** **PN**
'Paiʔan is eating meat' [Jeng 1977: 285]

c. *ma'un* ***Paiʔan*** *titiʔ*
AV:eat **PN** meat
'Paiʔan is eating meat' [Jeng 1977: 284]

d. **ma'un* *titiʔ* ***Paiʔan***
AV:eat meat **PN**
for: 'Paiʔan is eating meat' [Jeng 1977: 285]

4. Conclusion and discussion

This paper investigated word order in Austronesian symmetrical voice languages. It has been claimed in the literature that actors are universally more prominent than undergoers and that there is a processing bias favouring actors over undergoers. At first sight, Austronesian symmetrical voice systems do not seem to confirm these observations. Due to their symmetrical relation between actor and undergoer voice, many languages allow for both orders, actor > undergoer and undergoer > actor as basic word order patterns. Totoli and Tagalog have been used to exemplify two patterns, the *flexible subject+VP-pattern* and the *verb initial+final subject-pattern*, that are widely found among western Austronesian symmetrical voice languages, and which show no evidence for agent-first in basic word order patterns.

However, the claim of this paper has been, that whenever there is a deviation from these two patterns, there will be a tendency for actors the be realised before undergoers. Looking at a wider range of languages, this has been borne out. Tables 1 and 2 summarize the different deviating patterns, and the languages in which these patterns are attested. Obviously, a larger set of languages is needed to make more reliable claims about word order preferences and a potential bias towards agent-first position in western Austronesian symmetrical voice languages.

Deviations from the subject+VP-pattern	
1. Allowing for two agent-first positions in AV (i.e. SVO & VSO)	*Bajau, Pitu Ulanna Salu, Kelabit*
2. Ban of VOS in AV (i.e. strict agent first in AV)	*Tondano, Rampi', Gayo*
3. Ban of VOS + two agent-first positions in AV (i.e. strict agent-first in AV)	*Begak*

Table 1: Deviating patterns from the subject+VP-pattern

Deviations from the verb initial+final subject pattern	
1. Strict agent-first in UV	*Ibaloy, Chotabato Manobo*
2. Strict agent-first in all voices	*Tatana', Tikavatan Bunun, Mansaka, Agat*
3. Obligatory post-verbal actor clitic in AV	*Tboli*
4. Obligatory cross-referencing with the actor in UV	*Puyuma, Tsou*
5. Agent-first if case marking is missing	*Takbanuad Bunun*

Table 2: Deviating patterns from the verb initial+final subject pattern

Note also that all observations and generalizations made in this paper refer to the behaviour of full noun phrases and free pronouns only. A closer investigation of the ordering of pronominal clitics might well show different preferences and patterns. Likewise, a quantitative study on the actual word order preferences in natural discourse remains a topic for further research.

References

Artawa, I Ketut. 1998. Ergativity and Balinese syntax (Parts 1, 2, and 3). *NUSA Linguistic Studies of Indonesian and Other Languages in Indonesia*, 42, 43, and 44.

Baker, Mark C. 1996. *The polysynthesis parameter* (Oxford Studies in Comparative Syntax). New York: Oxford University Press.

Bickel, Balthasar, Witzlack-Makarevich Alena, Choudhary Kamal K., Schlesewsky Matthias, Bornkessel-Schlesewsky Ina (2015) The Neurophysiology of Language Processing Shapes the Evolution of Grammar: Evidence from Case Marking. *PLoS ONE* 10(8): e0132819. doi:10.1371/journal.pone.0132819.

Brickell, Timothy C. 2014. *A grammatical description of the Tondano (Toundano) language.* PhD dissertation: La Trobe University.

Brickell, Timothy C. and Stefan Schnell. 2017. Do grammatical relations reflect information status? Reassessing Preferred Argument Structure theory against discourse data from Tondano. *Linguistic Typology* 21.1.

Clayre, Beatrice. 2014. 'A preliminary typology of the languages of Middle Borneo'. In Peter Sercombe, Michael Boutin & Adrian Clynes (eds.) *Advances in research on cultural and linguistic practices in Borneo*, 123-151. Phillips, Maine USA: Borneo Research Council.

Campbell, Philip J. 1989. *Some aspects of Pitu Ulunna Salu grammar: A typological approach.* M.A. dissertation: University of Texas at Arlington.

De Busser, Rik L. J. 2009. *Towards a grammar of Takivatan Bunun. Selected topics.* PhD dissertation: La Trobe University.

Dik, Simon C. 1978. *Functional grammar*. (Publications in Language Sciences 7). Dordrecht: Foris Publications.

Dillon, John A. 1994. *A grammatical description of Tatana'*. M.A. dissertation, University of Texas at Arlington.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67. 547-619.

Eades, Domenyk. 2005. *A grammar of Gayo: A language of Aceh, Sumatra.* Canberra: Pacific Linguistics.

Ferreirinho, Naomi. 1993. *Selected topics in the grammar of Limos Kalinga, The Philippines.* Canberra: Pacific Linguistics.

Forsberg, Vivian M. 1992. *A pedagogical grammar of Tboli.* Studies in Philippine Linguistics, Volume 9, 1-110.

Givón, Talmy. 1984. Syntax: *A functional-typological introduction*. Volume I. Amsterdam: John Benjamins.

Goudswaard, Nelleke. 2005. *The Begak* (*Ida'an*) *language of Sabah.* Uetrecht: LOT.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA & London: The MIT Press.

Healey, Phyllis M. 1960. *An Agat grammar.* Manila: Bureau of Printing.

Hemmings, Charlotte. 2016. *The Kelabit language, Austronesian voice, and syntactic typology.* PhD dissertation, SOAS, University of London.

Himmelmann, Nikolaus P. 2005a. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*, 110-181. London & New York: Routledge.

Himmelmann, Nikolaus P. 2005b. Tagalog. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*, 350-376. London & New York: Routledge.

Himmelmann, Nikolaus P. 2008. Lexical categories and voice in Tagalog. In Peter K. Austin & Simon Musgrave (eds.), *Voice and grammatical relations in Austronesian languages*, 247-293. Stanford: CSLI Publications.

Himmelmann, Nikolaus P. and Sonja Riesberg. 2013. Symmetrical voice and applicative alternations: Evidence from Totoli. *Oceanic Linguistics* 52. 396-422.

Jackendoff, Ray S. 1972. *Semantic interpretation in Generative Grammar*. Cambridge, MA: The MIT Press.

Jeng, Heng-hsiung. 1977. *Topic and focus in Bunun.* Taipei: Academia Sinica.

Kaufman, Daniel. 2015. Lexical category and alignment in Austronesian. In Jessica Coon, Diane Massam and Lisa D. Travis (eds.) *The Oxford handbook of ergativity.* Oxford: Oxford University Press.

Kerr, Harland. 1988. *Cotabato Manobo grammar.* Studies in Philippine Linguistics, Volume 7, 1-123.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19. 335–391.

Mayani, Anik. 2013. *A grammar of Tajio. A language spoken in Central Sulawesi*. PhD dissertation, Universität zu Köln.

Miller, Mark T. 2007. *A grammar of West Coast Bajau.* PhD dissertation, The University of Texas at Arlington.

Mithun, Marianne and Wallace Chafe. 1999. What are S, A, and O? *Studies in Language* 23. 569–596.

Pesetsky, David. 1995. *Zero syntax. Experiencers and cascades*. Cambridge, MA & London:

The MIT Press.Teng, Stacy Fang-Ching. 2008. *A reference grammar of Puyum. An Ausronesian language of Taiwan.* Canberra: Pacific Linguistics.

Tung, T'ung-ho. 1964. *A descriptive study of the Tsou language, Formosa.* Taipei: Academia Sinica.

Riesberg, Sonja and Beatrice Primus. 2015. Agent prominence in symmetrical voice languages. *STUF – Language Typology and Language Universals* 68(4). 551–564.

Ruffolo, Roberta. 2004. *Topics in the morpho-syntax of Ibaloy, Northern Philippines.* PhD dissertation, Australian National University.

Schachter, Paul and Fe T. Otanes. 1972. *Tagalog reference grammar*. Berkely, New York & London: University of California Press.

Svelmoe, Gordon and Thelma Svelmoe. 1974. *Notes on Mansaka grammar.* Language Data, Asian-Pacific Series No. 6. Huntington Beach: SIL.

Van Valin, Robert D. and Randy J. LaPolla. 1997. *Syntax. Structure, meaning and function*. Cambridge: Cambridge University Press.

Zeitoun, Elizabeth. 1992. *A syntactic and semantic study of Tsou focus system.* M.A. dissertation, National Tsing Hua University.