

## Antifunctionality in Shawi split ergativity, a processing analysis<sup>1</sup>

Antifuncionalidade na ergatividade dividida do Shawi, uma análise do processamento

Luis Miguel Rojas-Berscia

Radboud Universiteit Nijmegen, Centre for Language Studies

ORCID: 0000-0002-0492-9429

Corentin Bourdeau

Radboud Universiteit Nijmegen, Centre for Language Studies

ORCID: 0009-0001-5390-115

Stefan Grondelaers

Radboud Universiteit Nijmegen, Centre for Language Studies & Meertens

Instituut

ORCID: 0000-0002-9535-6637

Pieter A. M. Seuren †

Max Planck Institute for Psycholinguistics

ORCID: 0000-0002-5976-4115

DOI: 10.26512/rbla.v15i1.50748

Recebido em maio/2023 e aceito em junho/2023

### Abstract

In this article, as a follow-up of Rojas-Berscia and Bourdeau (2017), we study morphological ergativity in Shawi (Kawapanan) using a dedicated experimental design. Shawi displays ergative-marking in an opposite direction from Silverstein's Nominal Hierarchy (NH) (Silverstein 1976). We claim this pattern to be *antifunctional*, given the lack of internal syntactic cues that explain why ergativity is omitted or completely obligatory in cases where the NH predicts the opposite.

To test this hypothesis, we carried out a grammaticality judgment experiment in the field with 47 Shawi participants from four sites. We found a significant overall effect of the Antifunctional Ergativity Constraint Expectation (AECE): sentences that violated this constraint were in general deemed less acceptable. Finally, we provide a tentative hypothesis on the historical origin of this pattern, resorting to discussions on the origins of ergativity in historical syntax (Gildea 2004; Gildea and Queixalós 2010), the reconstruction of Proto-Kawapanan morphosyntax, and antifunctional patterns in language (Seuren and Hamans 2010).

Keywords: Split ergativity, Nominal Hierarchy, historical syntax, Shawi, Amazonian languages

## Resumo

Neste artigo, como sequência de Rojas-Berscia e Bourdeau (2017), apresentamos um estudo sobre a ergatividade na língua shawi (kawapana), usando um desenho experimental. O shawi mostra a marcação ergativa numa direção oposta da Hierarquia Nominal (Silverstein 1976). Argumentamos que este padrão é *antifuncional*, devido à falta de indicações sintáticas internas que explicam por quê a ergatividade é omitida ou completamente obrigatória nos casos onde a Hierarquia Nominal prediz o contrário. Para testar esta hipótese, realizamos um experimento de juízo de gramaticalidade no campo com 47 participantes shawi de quatro localidades. Encontramos um efeito significativo geral da Expectativa da Restrição da Ergatividade Antifuncional (EREA): as orações que violaram esta restrição foram julgadas de menos aceitáveis. Finalmente, fornecemos uma hipótese tentativa sobre a origem histórica deste padrão, recorrendo à discussões sobre a origem da ergatividade na sintaxe histórica (Gildea 2004; Gildea e Queixalós 2010), a reconstrução da morfossintaxe proto-kawapana e o conceito de antifuncionalidade na linguagem (Seuren e Hamans 2010)

Palavras chave: ergatividade escindida, Hierarquia Nominal, sintaxe histórica, shawi, línguas amazônicas

## 1. Introduction

### 1.1. Split Ergativity and the Nominal Hierarchy

Ergativity is a grammatical pattern in which the subject of an intransitive clause (S) is treated in the same way as the object of a transitive clause (O), and differently from a transitive subject (A) (Dixon 1994: 1). Below, we present an example of prototypical ergative marking, where the A-NP in (1b) is ergatively marked, whereas the S-NP in (1a) remains unmarked:

Basque (isolate)

(1) a. Gizon-a    dator.

man-DET come

‘The man is coming.’ (de Rijk & Coene, 2008: 198)

b. Gizon-a-k    zakurr-a          ikusi          du.

man-DET-ERG                                      dog-DET          see          AUX

‘The man has seen the dog.’ (Hualde and Ortiz de Urbina 2003: 180–81)

Such a regular pattern is rare among the languages of the world. Many languages that display some sort of ergative case-marking present

asymmetries. These make ergativity operate only in particular circumstances. In other circumstances, other non-ergative systems operate. The conditions under which such asymmetries operate are lexically or grammatically determined (McGregor 2009: 486). This phenomenon is referred to in the typological literature as *split ergativity* (Silverstein 1976: 175). Split ergativity can be conditioned by several factors (de Hoop and Narasimhan 2009 for an account on Hindi split ergativity). . Some of these are (a) the nature of the lexical verb, (b) the nature of the A-NP, (c) tense, aspect and/or mood, and (d) whether it occurs on main or subordinate clauses (McGregor 2009: 486). The semantic features of the core arguments of the main verb are one of the most salient. Silverstein (1976: 176) argues that split ergative systems are not random but follow a lexical hierarchy in which NPs are classified according to their ‘inherent lexical contents’ and their ‘semantic naturalness’ for occupying the A function. This is currently known as NOMINAL HIERARCHY (Dixon 1994), henceforth NH, or ‘person/animacy hierarchy’ (Woolford 2009). The two hierarchies are intrinsically equivalent, except that Woolford (2009) takes the ‘number’ parameter into consideration, whereas Dixon (1994) makes an additional distinction between ‘proper nouns’ and ‘common nouns’. Below, we present an adaptation of the hierarchy:

Figure 1. The Nominal Hierarchy, an adaptation based on Dixon (1994) and Woolford (2009)

**1.pl » 1.sg » 2.pl » 2.sg » 3.pl » 3.sg » proper noun » hum.pl » hum.sg » anim.pl » anim.sg » inan.pl » inan.sg**

When constrained by the semantics of the NPs, split ergative systems can be explained on the basis of two principles:

- 1) In transitive clauses, case marking serves as a means for distinguishing A from O.
- 2) Certain NPs are more likely to occur as A than others based on the capacities of their referents to control an event.

The first principle accounts for the existence of a split. If case-marking is deployed to distinguish A and O, then it is only needed when there is ambiguity. Yet, not all NPs have the same probability to occur in the A function, as put forward in the second principle. As such, no special marking is required when the A-NP is a prototypical agent. Ergative markers are used to identify A arguments when they are unexpected. According to

Dixon, “averaging out over all types of verbs, there is no doubt that human NPs are more likely to be in A than in O function, and that inanimates are more likely to be in O function than in A, with non-human animates falling between these” (1994: 84). Dixon also suggests that, in discourse, the participants involved in the interaction, i.e. speaker and addressee, tend to be privileged actors and tend to occur in A function, the first person pronoun being the most likely to refer to an A argument since the speaker regards himself/herself as the quintessential agent (1994: 84). These tendencies are represented in the NH. Therefore, in systems where the split is conditioned by the semantics of the core arguments, the ergative marker is expected to occur on NPs ranked on the right-hand side of the cline. The breaking-point is therefore established somewhere to the left, beyond which NPs remain unmarked.

In this article we assume that the primary function of language is *the establishment of socially binding commitments or appeals with regard to given propositions, whereby facilitation of language acquisition is a further functional factor*, rather than the marking of group loyalties or of differences in social prestige (Seuren 2009; Seuren and Hamans 2010). Therefore, a linguistic feature like the NH seems a useful functional constraint to comply with those goals, i.e. given the unnecessariness of marking due to salience, marking is simply avoided. This is somehow surmised as well in Gildea (2004: 4), where the author suggests that person-based splits seemed to be the only good candidate for cognitively motivated ergativity.

It has been shown that the NH is far from being a universal. Some languages display several breaking-points along the NH. This is the case of Arrernte, a Pama-Nyungan language of Australia:

Figure 2. Arrernte’s ergative-marking pattern<sup>2</sup> (Woolford 2009)

1.pl » (1.sg) » 2.pl » 3.hum.pl » 3.hum.sg » 3.anim.pl » (3.inan)

Ergative marking on third-person inanimate NPs fits the expectations of the NH. However, the NH does not explain the ergative-marking of the first-person pronoun singular, which occurs typically in A function and therefore is expected not to be marked with the ergative.

Hindi is an even more problematic case for the theory (cf. Piepers 2016; de Hoop and Narasimhan 2009). De Hoop and Narasimhan (2009) argue that the ergative case in Hindi appears on prototypical subjects only (to the left of the NH). This suggests that the marking of core arguments

is not just about distinguishing A from O. Instead, de Hoop and de Swart (2009) propose that case-marking can express certain semantic or pragmatic information, such as agentivity and volitionality (2009: 5). The authors also introduce the notion of ‘argument strength’. They claim that strong arguments are likely to be overtly marked in languages displaying differential subject marking. However, the ‘strength’ of the core arguments can also be affected by various factors that go beyond nominal semantics. De Hoop and Narasimhan (2009) correlate the degree of strength of the core arguments with the degree of transitivity, which itself depends on features of the subject NP (e.g. volitionality and potency), of the object NP and of the verb phrase (e.g. telicity, action, realis). The ‘strongest’ subject is thus argued to be the subject of an active, telic transitive or ditransitive clause that is volitional, high in potency, and that co-occurs with an animate and definite object (de Hoop and Narasimhan 2009: 65). In fact, in addition to the degree of animacy of the subject, ergative marking in Hindi is constrained by the lexical verb class — in this case the verb must refer to a volitional process — and aspect. Also, ergative markers in Hindi appear in perfective clauses only, confirming for this case the status of perfectivity as a subject-strengthening feature (de Hoop and Narasimhan 2009: 66).

Although Arrernte and Hindi violate the NH, there are internal syntactic and semantic cues that allow us to understand the limits/constraints behind those violations. Patterns like those found in Arrernte and Hindi seem to be *antifunctional*, thus counterproductive to the primary function of language in speech and to language acquisition. We hypothesise that these changes can probably best be understood as formally abstract and introspectively inaccessible processes<sup>3</sup> that may have been the result of language change in the acquisition process by young or adult learners as a way to set themselves off against adjacent groups and, therefore, generating major internal community cohesion (Seuren and Hamans 2010: 159).

In this article, we present a descriptive and experimental analysis of the split ergativity system of Shawi.<sup>4</sup> We claim that Shawi represents a challenge to the NH, displaying an antifunctional pattern, which not only partially violates the hierarchy, as is the case of Arrernte, but also structurally mirrors it. Given the rarity of this case, we surmised necessary to test earlier observations experimentally. To test our hypothesis, we carried out a grammaticality judgment experiment in the field, with participants from four non-adjacent field sites in the Shawi area: Santa María de Cahuapanas, Pueblo Chayahuita, Balsapuerto, and Jeberos. The article is structured

as follows. In section 2, we present a survey of split ergativity in Shawi. Section 3 presents the methods we used in the field with regard to the design of the experiment and the actual trials with the Shawi participants. Section 4 provides a Bayesian mixed-effects regression analysis of the data. Finally, Section 5 summarises the results and provides a preliminary explanation for the origins of this pattern, resorting to our current knowledge of Kawapangan morphosyntax, historical syntax of ergative languages (Gildea 2004; Gildea and Queixalós 2010), and the notion of antifunctionality in language change (Seuren and Hamans 2010).

## 2. Split Ergativity in Shawi

Ergative NPs in Shawi are marked by means of the suffix *-ri*. Below, we explore several cases in which splits as regards ergativity are found in the languages of the world, such as sentence type, construction type (main vs. subordinate clauses), TMA (q.v. Bourdeau 2015 for a full account), word order, distinguishability of A/O-NPs and information structure.

Sentence type does not condition the use of an ergative marker in Shawi. The ergative suffix *-ri* can be found in assertive (2), negative (3), and interrogative sentences (4):

(2) *Ƙari ni'ni' tehparawe.*

Ka-ri	ni'ni'	tepa-r-aw-ø.
1MIN.EXCL-ERG	dog	kill-N.FUT-1MIN.EXCL.A-3MIN.O

'I killed a dog.' (EM\_CAS\_Chayahuita\_AHT\_143107)

(3) *Ƙari ku tehparawe ina.*

Ka-ri	ku	tepa-r-a(w)-ø-we	ina.
1MIN.EXCL-ERG	NEG	kill-N.FUT-1MIN.EXCL.A-3MIN.O-NEG	3MIN

'I did not kill him.' (EM\_CAS\_Chayahuita\_AHT\_143107)

(4) *Inta' inari keterin atari?*

In-ta' ina-ri ke-te-r-in-ø atari?  
 who 3MIN-ERG take-VM-N.FUT-3MIN.A-3MIN.R hicken  
 'Whom did he offer the hen to?' (EM\_CAS\_Chayahuita\_AHT\_143107)

Shawi also does not display split ergativity opposing main clauses to subordinate clauses. Proper subordinate clauses (5) and relative clauses (6) are compatible with the use of *-ri*:

(5) *Ku nuwantera naman ina wa'washari tehpakasumare'.*

Ku nuwan-te-r-a-ø naman ina  
 NEG want-VM-N.FUT-1MIN.EXCL.A-3MIN.O peccary DEF  
**wa'washa-ri tepa-ka-ø-ø-su'-mare'**  
 child-ERG kill-PURP-3MIN.A-3MIN.O-NMLZ-BEN  
 'I do not want the child/youngling to hunt the peccary.'  
 (GJT\_EM\_Chayahuita\_AHT\_170317).

(6) *Wa'washa kari i'wara yunirawesu' isu'.*

Wa'washa **ka-ri** **i'wara**  
 child 1MIN.EXCL-ERG yesterday  
**yuni-r-aw-su'** isu'-ø.  
 search-N.FUT-1MIN.EXCL-3AUG.O-NMLZ this-3.COP  
 'The children I searched for yesterday are these ones.'  
 (GJT\_stim\_Cahuapanas\_MPE\_160330)

As far as our corpus concerns, there is no way to account for split-ergativity in terms of Tense, Mood and Aspect, henceforth TMA. The

ergative marker can appear with the Non-Future/Future markers (7), with the imperative/subjunctive (8) and dubitative (9) moods, and with progressive aspect<sup>5</sup> (10):

(7) *Tashiraya kemari tehpapun Pitru.*

Tashiraya kema-ri tepa-pu-n-ø Pitru.

tomorrow 2MIN-ERG kill-FUT-2MIN.A-3MIN.O Peter

'Tomorrow you will kill Peter.'(EM\_CAS\_Chayahuita\_AHT\_143107).

(8) *Ka tananke pa'watun kari tehpachi nunu.*

Ka tanan-ke pa'-watu-n ka-ri

1MIN.EXCL forest-LOC go-SEQ-1MIN.EXCL 1MIN.EXCL-ERG

tepa-chi-ø-ø nunu.

kill-SUBJ-1MIN.EXCL.A-3MIN.O monkey

'When I go to the forest, I shall kill a monkey'

(EM\_CAS\_Chayahuita\_AHT\_143107).

(9) *Kiyarinkema tehpamarainkema kampita.*

Kiya-ri-nkema tepa-mara-i-nkema kampita.

1AUG.EXCL-ERG-2AUG.O kill-DUBIT-1AUG.EXCL.A-2AUG.O-PL 2AUG

'We may kill you (pl.)' (EM\_CAS\_Chayahuita\_AHT\_143107).

(10) *Kemapi-ri nishitarin kinan sawinike.*

Kemapi-ri nishi-t-a-r-in-ø kinan sawini-ke.

man-ERG cut-VM-PROG-N,FUT-3MIN.A-3MIN.O manioc.stick machete-LOC

'The man is cutting manioc sticks with a machete.'

(EM\_CAS\_Chayahuita\_AHT\_143107)



In the case of 3>3<sup>6</sup> sentences, Rojas-Berscia and Bourdeau (2017) carried out a small corpus investigation based on narratives, and concluded that ergativity in these cases “[...] is not optional at all, but conditioned by strict syntactic rules” (2017: 51), such as word order, e.g. (11) and (12). Distinguishability of A-NPs and O-NPs, e.g. (13) and (14), and information structure, e.g. (15) also play a role in the occurrence of the marker:

(11) *Trahka shawi kemaru'sari tihkirin.*

Iraka	Shawi	kema-ru'sa-ri	tiki-r-in-ø.
	O	A	V

Formerly Shawi Aguaruna-PL-ERG slaughter-N.FUT-3MIN.A-3MIN.O

‘THE SHAWI, the Aguaruna used to slaughter.’ (2018: 57)

(NA\_Balsapuerto\_Awkarusa\_BY\_P\_120131)

(12) *Iseke kankan nu'wirarin ni'nirari.*

Iseke	kankan	nu'wi-ra-r-in-ø	ni'ni-ra-ri.
	O	V	A

Here wasp yap-PROG-N.FUT-3MIN.A-3MIN.R dog-DIM-ERG

‘Then, IT IS YAPPING AT THE WASPS, the little dog.’ (2018: 58)

(FS\_Balsapuerto\_AM\_142807)

In (11), the canonical AOV order was violated. This has occurred because the O-NP was fronted for focalisation purposes. A similar process took place in (12), where the whole predicate ‘yapping at the wasps’ was fronted. Both cases triggered a non-canonical word order, hence the occurrence of ergative *-ri*.

(13) *Ni'niri nukurarin.*

**Ni'ni-ri** nuku-ra-r-in-ø.

dog-ERG look.at-PROG-N.FUT-3MIN.A-3MIN.O

‘The dog is looking at it (the toad).’ (2018: 58) (FS\_Balsapuerto\_AM\_142807)

(14) *Tururu' nukurarin.*

**Tururu'**nuku-ra-r-in- $\emptyset$ .

toad look.at-PROG-N.FUT-3MIN.A-3MIN.O

'He (the boy) is looking at the toad.'(2018: 58) (FS\_Balsapuerto\_AM\_142807)

A-NPs are likely to be omitted, given that Shawi is a pro-drop language. Therefore, the NP preceding the verb is likely to be the O-NP. However, if the only NP overtly expressed in the sentence is the A argument, it will be ergatively-marked, see (13). On the contrary, if ergative-marking is absent, as in (14), the unique argument is indeed an O.

(15) *Inaran iseke kankani peyararin, ni'niri nu'wirarin kankan.*

Ina-ran iseke kankan-ni peya-ra-r-in- $\emptyset$ ,

3MIN-ABL here wasp-ERG sting-PROG-N.FUT-3MIN.A-3MIN.O

ni'ni'-ri nu'wi-ra-r-in- $\emptyset$  kankan.

dog-ERG yap-PROG-N.FUT-3MIN.A-3MIN.Owasp

'Then, the wasps are stinging it (the dog), the dog is shouting at the wasps.'

(2018: 59) (FS\_Balsapuerto\_AM\_142807)

In example (15), we would expect the second clause not to have an ergative marker on the A-NP, given that it displays the unmarked AVO order. Nevertheless, ergative *-ri* is there to mark a contrast between 'the wasps' and 'the dog'. The new subject 'the dog' is the new protagonist of the event and was thus ergatively-marked.

However, there are cases in which the use of the ergative marker will be either obligatory or forbidden by the grammar. This asymmetry will be henceforth dubbed *Antifunctional Ergativity Constraint* (AEC). When the A-NP is a first person and the O-NP, a second person, i.e. 1>2, regardless of

the number, speakers of Shawi always use an ergative-marker. The absence of the suffix is unanimously rejected:

(16) *Ƙarinke aweranke.*

Ka-ri-(n(ke))<sup>7</sup> awe-r-a-nke.

1MIN.EXCL-ERG-2MIN.O hit-N.FUT-1MIN.EXCL.A-2MIN.O

‘I hit you’ (EM\_CAS\_Chayahuita\_AHT\_143107).

(17) *Ƙrinkema’ tashinanpeitawenkema’*

Ka-ri-(nkema’)tashinanpei-t-a-we-nkema’.

1MIN.EXCL-ERG-2AUG.O prison-VM-PROG-1MIN.EXCL.A-2AUG.O

‘I am imprisoning you (pl.)’ (EM\_CAS\_Chayahuita\_AHT\_143107).

(18) *Ƙiyarisu’ ku kahtawarainkewe.*

Kiya-ri-(n)-ø-su’ ku

1AUG.EXCL-ERG-2MIN.O-3MIN.COP-NMLZ NEG

katawa-r-ai-nke-we.

help-N.FUT-1AUG.EXCL.A-2MIN.O-NEG

‘It was not us who did not help you’ (EM\_CAS\_Chayahuita\_AHT\_143107).

In addition, regardless of the semantics of the verb, the ergative marker remains obligatory. Below we present examples with a prototypical transitive verb (19), a less transitive verb (20), and psychological verbs (21)-(22). These sentences suggest that the degree of transitivity has no impact on the application of the rule.

(19) *Ƙarinke aniiranke.*

**Ka-ri**-(n(ke)) a-nii-r-a(w)-nke.

1MIN.EXCL-ERG-2.O CAUS-jump-N.FUT-1MIN.EXCL.A-2MIN.O

‘I made you jump.’ (EM\_CAS\_Chayahuita\_AHT\_140731)

(20) *Ƙwara kari yuniranke.*

Iwara **ka-ri** yuni-r-a-nke.

yesterday 1MIN.EXCL-ERG look.for-N.FUT-1MIN.EXCL.A-2MIN.O

‘I looked for you yesterday.’ (EM\_RI(trans2)\_DYI\_2014-07-21)

(21) *Ƙari nawaranke.*

**Ka-ri** nawa-r-a-nke.

1MIN.EXCL-ERG miss-N.FUT-1MIN.EXCL.A-2MIN.O

‘I missed you.’ (EM\_CAS\_Chayahuita\_AHT\_140731)

(22) *Ƙari nateranke.*

**Ka-ri**-(n(ke)) nate-r-a-nke.

1MIN.EXCL-ERG-2MIN.O trust-N.FUT-1MIN.EXCL.A-2MIN.O

‘I trusted you.’ (RI(psy)\_VPP\_2014-07-24)

In ditransitivised verbs, when the second person pronoun refers to the recipient and not the patient of the action, the rule also holds, see (23).

(23) *Ƙari keteranken ipi’ nusha’.*

**Ka-ri** ke-te-r-a-**nken** ipi’ nusha’.

1MIN.EXCL-ERG bring-VM-1MIN.EXCL.A-2MIN.O lowland.paca meat

‘I gave you lowland paca meat’ (EM\_CAS\_Chayahuita\_AHT\_140731).

In the contexts that follow, the use of ergative *-ri* is unanimously rejected. This prohibition is related to, first, the type of NP occupying the O function. If the O-NP is a first person, regardless of number, the A-NP is not ergatively-marked:

(24) *Iwara kema nuwiranku.*

Iwara kema(\*-ri) nuwi-r-an-ku.

yesterday 2MIN tell.off-N.FUT-2MIN.A-1MIN.EXCL.O

‘You told me off yesterday’ (EM\_CAS\_Chayahuita\_AHT\_140731).

(25) *Kampita tewateramakui kiya?*

Kampita(\*-ri) tewate-r-ama-kui kiya?

2AUG be.scared-N.FUT-2AUG.A-1AUG.EXCL.O 1AUG.EXCL

‘Are you afraid of us?’(EM\_CAS\_Chayahuita\_AHT\_140731).

(26) *Kemaru’sa yatehpapirinenpuwawe kanpuwa’.*

Kema-ru’sa(\*-ri) ya-tepa-wi-r-in(e)-npuwa’-w kanpuwa’.

Aguaruna-PL want-kill-FRUST-N.FUT-3MIN.A-1AUG.INCL.O-NEG1AUG.INCL

‘The Aguaruna wanted to kill us, in vain.’ (EM\_CAS\_Chayahuita\_AHT\_140731)

Second, if the A argument is different from a first person, it is not ergatively-marked when the O function is occupied by a second person, regardless of number. Below we present some examples:

(27) *Pitru naterinke.*

Pitru(\*-ri) nate-r-in-nke

Peter trust-N.FUT-3MIN.A-2MIN.O

‘Peter trusted you’ (EM\_CAS\_Chayahuita\_AHT\_140731)

(28) *Inapita nuwirinkema kampita.*

Inapita(\*-ri) nuwi-r-in(e)-(nke)ma kampita.

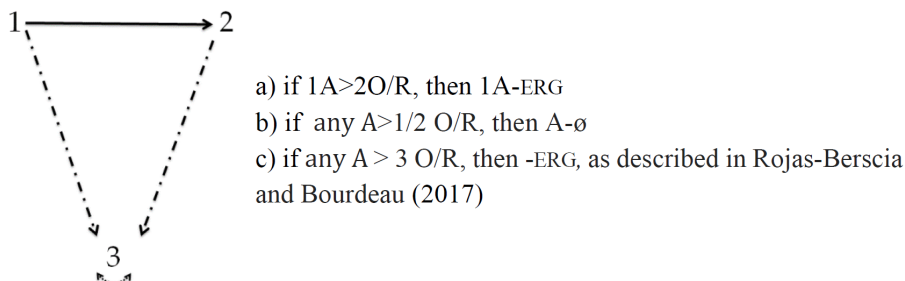
3AUG hate-N.FUT-3AUG.A-2AUG.O 2AUG

‘They hated you’ (EM\_CAS\_Chayahuita\_AHT\_140731).

The two previous conditions are also valid with ditransitived verbs when the first or second person marker refers to the recipient of the action.

A summary of these rules follows:

Figure 3. The Shawi’s split ergativity<sup>8</sup> system



### 3. Methodology

#### 3.1. Introduction

Available census data indicate that 52% of the Shawi over the age of five are non-literate (going up to 63% for the females, see [www.peruecologico.com](http://www.peruecologico.com)). Extracting grammaticality judgments from unschooled adults is a challenge because assessing syntactic quality requires metalinguistic skills non-literate people typically do not possess. Kurvers (2002) for example, found that L1 grammaticality judgments elicited from non-literate participants were based on lexical meaning and social convention rather than on grammar. In Van de Craats & Kurvers (2014), low-literate L2-learners who were asked to rate the grammaticality of the sentence *Mother’s bike is stolen again* replied that the sentence was incorrect since “one should not steal the bike of a mother. She needs the bike to bring her children to school” (cited in van de Craats, Kurvers, and van Hout 2015: 15). The lack of metalinguistic ability in low- or non-literate participants also transpires in problems with deixis (an inability to distinguish the “I” in a text from

themselves) and with syllogistic reasoning (van de Craats, Kurvers, and van Hout 2015: 16).

We addressed the literacy concern by balancing our participant sample in terms of education, and by sacrificing the lexical and contextual richness of our samples to experimental control. We are aware that our method, which heavily relies on judgments of grammaticality or well-formedness, has, in a way, a “back-drop”, which is the absence of a given context.<sup>9</sup> However, in line with Matthewson (2004: 376), had we just relied on text collection we would have probably never found the pattern we deal with in the present study. An exclusive reliance on textual evidence would have probably provided no negative evidence for the phenomenon under discussion.<sup>10</sup> The collection of texts was particularly useful in Rojas-Berscia and Bourdeau (2017) when dealing with the optionality of the ergative marker in certain contexts (mainly 3>3 sentences). In the case of a first person acting upon a second person in a sentence –, which will possibly never occur in a narration, and may scarcely occur in a dialogue – we needed to ask a native speaker to give a judgment. This was the only possible way to obtain the necessary negative evidence.

### 3.2. Stimuli

We constructed sentences that were as short and lexically unchallenging as possible, in order to focus the participants’ attention on the structure rather than on the content of the items. Stimuli consisted of single main clauses and were constructed mainly with Shawi words for items from the Swadesh list, a collection of concepts selected by Morris Swadesh (1971) for their universal, culturally independent availability in as many languages as possible. The sentences were constructed by the first and fourth author of the study together with two Shawi assistants, Moisés Pinedo Escobedo (Cahuapanas) and Abimael Huiñapi Tangoa (Chayahuita). The first author relied on previous studies (Barraza de García 2005; Rojas-Berscia 2013; Bourdeau 2015), and on direct elicitation and translation for the creation of the stimuli. These were then discussed with the two native speakers to avoid the possibility of infelicitousness of the sentences in isolation or any cultural taboo associated with them, and later recorded. The stimuli were recorded using the voice of Abimael Huiñapi Tangoa. His voice was automatically slightly distorted in every trial by Psychopy<sup>11</sup> to avoid any recognition by peers or dialectal bias.

All in all, 19 critical stimuli were constructed in six sets. Set 1 contained three stimuli to test the prediction that first person subjects in transitive constructions with a second person object obligatorily take the ergative suffix *-ri*, whereas the three stimuli in set 2 were designed to verify whether second or third person subjects acting upon a first-person object are mandatorily constructed without the ergative suffix. In set 3, two stimuli were presented in two orders to test whether first or second person subjects acting upon a third person object do indeed take the ergative marker. Set 4 contained three items to test the prediction that a third person subject acting upon a second person object does not take the ergative suffix. Sets 5 and 6 each contained three items to test the prediction that in transitive constructions with a third person subject acting upon a third person object, the ergative is obligatory if unmarked constituent order is altered: set 5 features three stimuli in OVS-order, set 6 three in OSV.<sup>12</sup>

All critical sentences were presented with and without the ergative marker *-ri* in a between-subjects design, to ensure that no participant saw the two versions of the same sentence. The 19 critical sentences were randomised with 8 unrelated filler sentences of variable acceptability. All participants first evaluated four highly ungrammatical test sentences presented in the same order (two intransitives with ergative marking, two with incorrect number morphology). These test sentences were included for the participants to become acquainted with the procedure, and for the experimenters to gauge the quality of the participants (viz. their ability to pass judgment on structural quality).

### 3.3. Participants

Grammaticality intuitions were elicited from 47 Shawi participants from four locations. Extreme flooding during fieldwork (March – May 2017) prevented data collection in the *de facto* Shawi cultural capital Pueblo Chayahuita; two of the three Chayahuita participants in the dataset were interviewed in San Lorenzo (a multi-ethnic market town outside the Shawi heartland), one in Cahuapanas. There were 17 residents from Cahuapana itself, 14 from Balsapuerto (one of whom was interviewed in the multi-ethnic market town Yurimaguas); 13 participants were interviewed in the town of Jeberos, but originated from the neighbouring Shawi communities Jordania and Bethel.

In light of the gender inequality among the Shawi (Dradi 1987), and the



general reluctance of females to talk to strangers when unaccompanied by husband or relatives, we were unable to balance gender in this study: as a consequence, there were only 14 female vs. 33 male participants. The mean age of the participants was 29.8, ranging from 17 to 65. 12 participants reported to be unschooled or only minimally educated, 19 to have (some) secondary schooling including reading and writing Spanish, and 16 to have post-secondary education (for the purposes of this study, we will regard participants in the first group as non-literate, and participants in the second and third group as literate). In order to counter the fact that female Shawi are typically less educated, gender was more or less balanced across the education levels (respectively 25 and 26% of the participants on levels 1 and 2 were female, 37,5 % on level 3 was female).

All participants were paid the Peruvian equivalent of 4 euro's for participation. One participant was excluded from participation on the basis of his inability or unwillingness to evaluate the test sentences.

### 3.4. Procedure

The experiment was implemented in the open-source application *Psychopy* (psychopy.org) and presented on a portable computer complemented with headphones. Participants took the experiment in the presence of the first author (henceforward “the experimenter”), who is a fluent speaker of Shawi, and either Abimael Huiñapi Tangoa or Moisés Pinedo Escobedo, our two Shawi assistants. They explained the task in Shawi. Participants were told that they would listen to a number of short sentences, and that they would have to determine on a ten-point scale whether and to what extent the sentences were “good Shawi”. The left-most grade (1) stood for “not existent/not said by us”. The right-most grade (10) stood for “very good Shawi/said by us”. After the participant had indicated (s)he had understood the task, they proceeded to the four test trials, which were evaluated without the headphones on. All critical trials and fillers were subsequently evaluated with the headphones on. The output of the experiment is available in the Appendix. On average, the experiment lasted about ten minutes. On a number of occasions, participants commented on the experimenter's Shawi in Spanish.

## 4. Results

Our statistical analysis of the data examined factors that influence judgments of grammaticality. We performed three analyses. The first analysis (section 4.1) examined whether the predictions of the Antifunctional Ergativity Constraint (AEC) are borne out empirically in the dataset as a whole and showed that they are. The second analysis investigated whether this pattern of results varies according to the grammatical type of the sentence, and it turns out that that it does. The third analysis examined whether the results vary depending on the level of education of the participant. They do not.

### 4.1. Ergativity and Grammaticality

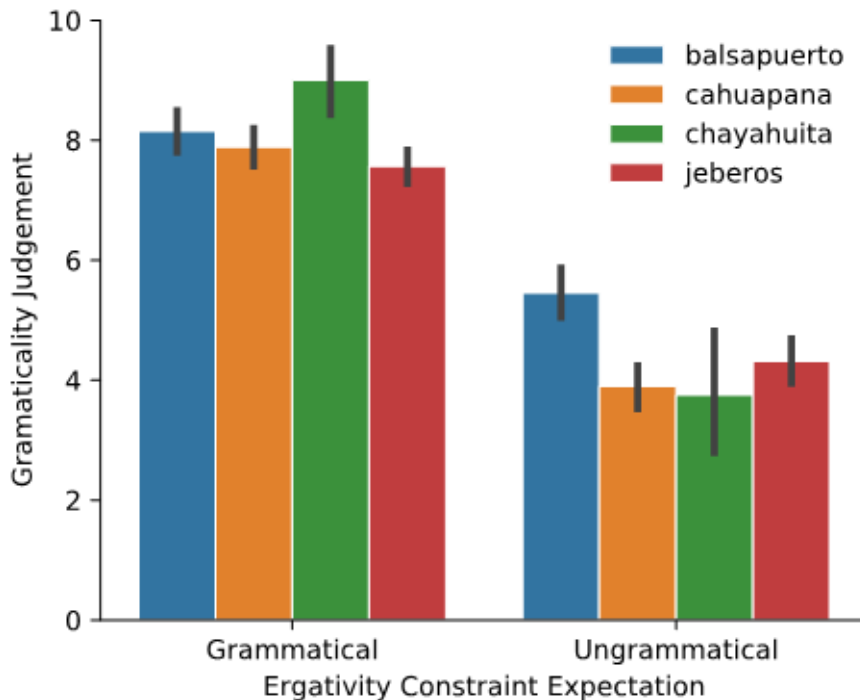
Figure 4 shows the distribution of grammaticality judgments broken down by whether or not the sentence under consideration respects the AEC, and by the origin of the participant. The figure illustrates a clear trend in support of our ergativity-motivated expectations. To examine this pattern, we performed a linear mixed effects regression analysis with *Grammaticality Judgment* as dependent variable. We performed Bayesian inference in the model using the PYMC3 probabilistic programming framework (Salvatier, Wiecki, and Fongesbeck 2016), via the Python-based *Bayesian Model Building Interface (BAMBI)*<sup>13</sup>. In this initial analysis, we focused on the question: did participants evidence knowledge of ergativity by responding differentially to sentences that do and do not violate the AEC?

The statistical model included *Antifunctional Ergativity Constraint Expectation (AECE)*, *participant origin*, and *participant age* (and all their two- and three-way interactions) as fixed effects, with random intercepts for *sentence id*, *participant id*, and *participant gender*. Continuous variables were centred at zero. We obtained posterior estimates from 5000 MCMC samples (minus 1000 burn-in samples) gathered using the NUTS algorithm under standard PYMC3 initialisation and parametrisation. See the supplementary materials for the full posterior summary. We treat any variable for which zero is not in the region of 95% highest posterior density as exhibiting a “significant” effect. The model intercept was significantly different to zero (). We observe a significant main effect of *AECE* (). No significant effects of *participant origin* or *participant age* are observed, but we do find significant effects of the interaction between participant origin and *AECE*: specifically, when compared to the judgments of respondents

from Balsapuerto, respondents from the other three regions, Jeberos (), Cahuapana () and Chayahuita () gave significantly lower grammaticality judgments to sentences that violate the AEC, as is also clear in figure 4 (right, Ungrammatical).

The ergativity constraint predicts grammaticality judgments, controlling for age, gender, origin, and idiosyncrasy among participants and sentences. The prediction holds across all four regions of origin, with participants from Jeberos, Cahuapanas, and Chayahuita providing the strongest judgments that AEC-violating sentences are ungrammatical.

Figure 4. Mean grammaticality judgment across sentence types, according to the prediction made by the Ergativity Constraint and broken down by participant place of origin.



#### 4.2. Ergativity and Grammaticality by Grammatical Subcategory

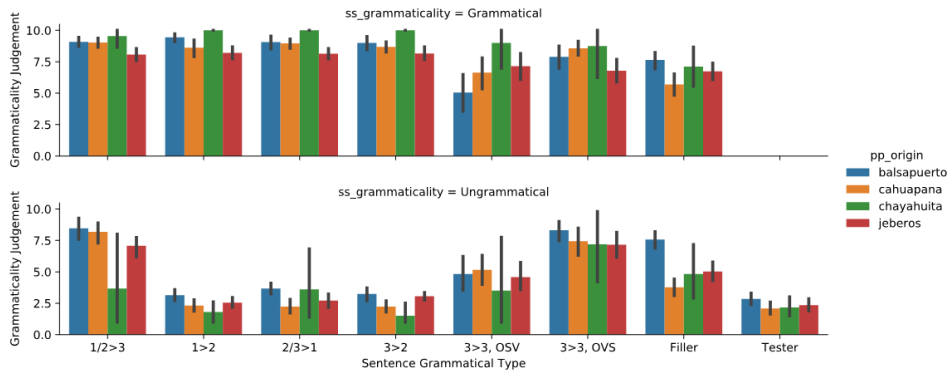
Our second analysis focused on the distribution of grammaticality judgments across specific types of sentences. Figure 5 shows this distribution of responses broken down according to whether the sentence respects the AEC and by the grammatical type of the sentence. We performed a

regression analysis in the same family and using the same tools as outlined in the previous section, again with *Grammaticality Judgment* as dependent variable. Given the findings of our previous analysis, and our focus on sentence type, we examined a simpler model with fewer by collapsing over participant origin, age, and gender. This provides better estimates of the variables we are interested in. This model included *sentence type* and *AECE* (and their interaction) as fixed effects, and a random intercept for *participant id*. Full posterior summaries for these variables can be found in the supplementary materials. The model intercept was significantly different to zero (). Again, we found a significant overall effect of AECE (): sentences that violated the AEC were in general deemed less acceptable.

We see significant variation by sentence type. There were significant effects of the interaction between AECE grammaticality and sentence type for sentence types “1>2” (), “2/3>1” (), and “3>2” (). Sentence type alone did not have a significant effect for any of these sentence types. As is clear from figure 5, these sentence types were all judged less acceptable when they violate the AEC (bottom row), and more acceptable when they adhere to the constraint (top row), but not more or less acceptable than other sentence types when Ergativity is ignored. By contrast, with respect to sentence types “3>3 OSV” (), “3>3 OVS” (), and “Filler” (), sentence type did exhibit a significant effect on grammaticality judgments overall: these sentence types were simply judged less acceptable overall. This pattern is strongest for sentence types “Filler” and “3>3 OSV”. A weaker, but still significant, dispreference was found for “3>3 OVS”. Accordingly, there were no significant effects of the interaction between sentence type and AECE grammaticality for these sentence types. Our analysis did not include Tester sentence types, since no ergativity-respecting versions of these sentences exist, which induces collinearity in model terms (these sentences types are also clearly unanimously rejected, see Figure 5).

In sum, our second analysis confirmed the role of Ergativity (AECE), and isolated the groups of sentences that are driving this effect, statistically speaking: “1>2”, “2/3>1”, and “3>2”. Together, analyses one and two confirm that, statistically speaking, the AEC predicts the grammaticality judgments of all regional groups: most categorically of respondents from Chayahuita, and least categorically of participants from Balsapuerto. Moreover, it is being represented in a manner that is discerning with respect to these different classes of sentence.

Figure 5. The distribution of grammaticality judgments across sentence types, broken down according to whether or not the AEC predicts a grammatical or ungrammatical interpretation, and by region of participant origin.



### 4.3. Ergativity, Grammaticality, and Education

We collected information on the educational level of our consultants. The distribution of education levels across subjects and groups in our data does not provide adequate enough coverage to perform the same level of statistical modelling of this factor. However, we were able to explore the influence of education on participants' decisions by formulating a simpler model that collapsed over regions of origin and sentence type. This allows us to ask: how does participant education relate to participants' accuracy in providing grammaticality judgments, ignoring the participant's region of origin? To address this question, we again performed a mixed effects regression analysis with *Grammaticality Judgment* as dependent variable, controlling for *sentence id* and *participant id* using random intercepts. In this model, we included AECE and participant education (and their interaction) as random effects. This model again confirmed a main effect of AECE (), but suggested that neither participant education, nor its interaction with AECE, had significant effects on participant's responses. Full posterior summaries can be found in the supplementary materials. We stress that these inferences are made with respect to data that collapses over important variables.

## 5. Discussion

Our findings can be summarised as follows:

- 1) Shawi split ergativity can only be described in terms of the AEC,

which is roughly a mirror of the NH, i.e. prototypical first person subject NPs have to be marked .

- 2) This pattern is widespread throughout the entire Shawi area. There is no variation in terms of age or gender of the participant. However, in terms of origin, participants from Pueblo Chayahuita, Jeberos, and Cahuapanas provided lower ratings to sentences violating the AEC. Although Balsapuerto participants also deem these sentences less acceptable, they displayed somewhat higher rankings.<sup>14</sup>
- 3) All Shawi participants considered that a violation of the AEC is unacceptable. The strongest effect was found in 1>2, 2/3>1, 3>2 sentences.
- 4) The 3>3 sentences we used for the experiment were deemed less acceptable overall. This may be due to the lack of a context for the sentences. It could be the case they were considered ‘infelicitous’ rather than ‘ungrammatical’. A more careful analysis, that expands on the findings of Rojas-Berscia and Bourdeau (2017), and that relies on direct elicitation, provided a careful experimental design for contexts, i.e. semantic fieldwork (Matthewson 2004), is necessary.
- 5) Regardless of their level of literacy/education, participants provided reliable metalinguistic judgments, i.e. the AEC in Shawi split ergativity is not a product of literacy or second language acquisition at school.

As mentioned in §4.3, our findings regarding the role of education in the judgments are not categorical. Given the complexity of our field site and complex gender differences in Shawi communities, it is difficult to collect a dataset with even groups from all literacy or education levels. Nevertheless, it was clear that even participants with the lowest level of education, and therefore illiterate, could provide reliable grammaticality judgments. The research reported here can be regarded as a “best practice” report on how to deal with low literacy when accessing linguistic knowledge (cf. Huettig, Kolinsky, and Lachmann 2018; Huettig and Mishra 2014). Indeed, the first author of this article had much experience as regards the question which illiterate participants would not understand an elicitation task of a certain kind. We claim, however, that careful design – including accessible Swadesh vocabulary, short stimuli, and calibration sentences —, preparatory fieldwork, and an in-depth knowledge of the native language of the participants, with a constant assessment of insiders (our Shawi assistants), can overcome the experimental barrier of low literacy. Feeling more comfortable with the researcher undoubtedly makes it easier for

participants to understand a task that involves metalinguistic evaluations and, therefore, access to his/her unconscious linguistic competence.

The AEC seems to be highly salient in the grammatical competence of Shawi speakers. All participants gave an immediate negative reaction when confronted with a sentence that violates the constraint. However, what are the origins of this pattern, and what could have been behind its consolidation in daily speech? We claim that, following Rojas-Berscia and Bourdeau (2017), morphological ergativity in Shawi originates in a passive construction that deployed an instrumental/comitative marker in Proto-Kawapangan. On the other hand, we hypothesise that these passive constructions were progressively relapsing, as it possibly occurred in Shiwilu. However, this process fossilised at some point, probably due in part to the late emergence of the notion of Shawi as a language, in times when the Jesuits established the first *Reducciones* in the north-Peruvian Amazonian foothills, and a new Shawi identity emerged out of the grouping of the Cahuapanas, Parapapura, Munich, and Chayahuita groups.

### 5.1. From instrumental to ergative: towards the origin of *-ri*

In Rojas-Berscia and Bourdeau (2017), following a Semantic Syntax approach, it was suggested that ergative marking in Shawi developed from a passive. However, according to McGregor (2009: 498), the assumption that ergatives originate in passives “is not so popular today”. We concur with McGregor that, in many cases, ergativity does not originate in passives. However, in many other cases it does. Other researchers such as Gildea find the passive > ergative path useful to argue for the motivated creation of ergative patterns (Gildea 2004: 22, who gives a careful assessment of the hypothesis in the Cariban family). For Queixalós and Gildea (2010: 13, bolds are ours), “languages gain main clause ergative patterns through reanalysis of biclausal constructions (especially nominalizations) and **marked voice constructions (especially passives)**”. Our current understanding of comparative Kawapangan morphosyntax allows us to go one step further and find more arguments to support the hypothesis that Shawi ergativity originated in a passive construction too. We claim that the contemporary ergative marker *-ri* and comitative marker *-re* share a common history. This is supported by both formal and functional arguments.

One of the most salient features of ergative marking in Shawi is (optional) object indexation in first person ergatively-marked pronouns:

- (30) **Ka-ri-(nkema’)**                    tashinanpei-t-a-we-nkema’.  
 1MIN.EXCL-ERG-2.O-PL            prison-VM-PROG-1MIN.EXCL.A-2AUG.O  
 ‘I am imprisoning you (pl.)’ (EM\_CAS\_Chayahuita\_AHT\_143107).

In (30), the first person minimal exclusive *ka* is ergatively-marked but also mirrors the verb in terms of object indexation, i.e. both the verb and the subject pronoun carry second person augmented object markers. This only occurs in ergatively-marked first-person pronouns.

A careful look at the Shawi case-marking system shows that this phenomenon occurs not only in ergative constructions, but also in comitative constructions. The formal resemblance between the ergative and the comitative marker is not to be taken lightly:<sup>15</sup>

- (31) *Kema ni’nireken pa’nan.*  
 Kema<sub>i</sub>      ni’ni-**re-ken/nke**<sub>i</sub>            pa’-n-an.  
 2MIN      dog-COM-2MIN.O            go-N.FUT-2MIN.S  
 ‘You left with your dog’ (GJT\_EM\_Chayahuita\_AHT\_170317).

Example (31) is a simple sentence, displaying the intransitive verb *pa’-* ‘to leave’, with the second person singular *kema* as its subject. The sentence also displays a comitative phrase in adjunct position. Just like the ergative, the Shawi comitative *-re* requires the indexation of object/nominal predication marking. In this case, unlike the ergative, the object/nominal predication marker refers to the subject of the sentence. This pattern is still highly productive in the language. A couple of examples follow:

- (32) *Ka tataruku iminke sahkaterawe.*  
 Ka<sub>i</sub>            tata-**ru-ku**<sub>i</sub>            imin-ke      saka-t-r-aw.  
 1MIN.EXCL   father-COM-1MIN.EXCL   farm-LOC   work-VM-N.FUT-1MIN.EXCL  
 ‘I work with my father in the farm’ (GJT\_EM\_Chayahuita\_AHT\_170317).



(33) *Kiya wa'anrukui tehpaterai.*

Kiya<sub>i</sub>            wa'an-**ru-kui**<sub>i</sub><sup>16</sup>            tepa-te-ra-i.

1AUG.EXCL leader-COM-1AUG.EXCL kill-VM-N.FUT-1AUG.EXCL

'We went hunting with the leader' (GJT\_EM\_Chayahuita\_AHT\_170317).

Examples (32) and (33) show cases in which the comitative carries first-person minimal exclusive nominal predication/object markers that are bound to the first person subject pronouns. This is not possible with the Shawi ergative, given the restriction that any argument acting upon a first or second person pronoun cannot be ergatively-marked. The paradigmatic behaviour of both case markers is summarised in Table 1:

Table 1: A comparison of the ergative and comitative cases paradigmatic behaviour<sup>17</sup>

	<b>Ergative</b>		<b>Comitative</b>
<b>1&gt;2</b>	<i>-ri-nke</i>	1	<i>-ru-ku</i>
<b>1&gt;2.pl</b>	<i>-ri-nkema'</i>	1.pl	<i>-ru-kui</i>
<b>1&gt;3</b>	<i>-ri-∅</i>	1+2	<i>-ru-npu</i>
<b>2&gt;1</b>	#	1+2 pl.	<i>-ru-npuwa'</i>
<b>2&gt;3</b>	<i>-ri-∅</i>	2	<i>-re-ken</i>
<b>3&gt;1</b>	#	2.pl	<i>-re-nkema'</i>
<b>3&gt;2</b>	#	3	<i>-re-∅</i>
<b>3&gt;3</b>	<i>-ri-∅</i>	3.pl	<i>-re-∅</i>

The comitative paradigm is complete, i.e. there is a one-to-one mapping between subject and object indexation in the comitative. The ergative paradigm is abridged, due to the AEC. For example, there is no ergative marking such as *\*-ri-nke* for 3>2, or *\*-ri-ku* for 3>1.

How could a comitative case marker become an ergative? Our current understanding of comparative Kawapangan morphosyntax allows us to solve the puzzle. Shiwilu, the sister language of Shawi, also displays ergativity. Ergativity in Shiwilu is marked by means of *-ler* suffixation (Valenzuela 2011). Nevertheless, unlike Shawi, it does not violate the NH. All contexts in which the Shiwilu ergative is used fall under the same constraints established for Shawi (all A > 3 O/R) sentences (Rojas-Berscia and Bourdeau 2017). In addition, the Shiwilu comitative marker *-lek* behaves just like Shawi *-re*, carrying object marking bound to the subject of the sentence. However, *-lek* is not only a comitative marker. It is also an instrumental marker:

(34) *Asu' Mikir utekkunanlek pilli'tulli dunansertaspi*

Asu' Mikir utekkunan-**lek** pilli'-tu-ll-i dunanser-taspi.

this Michael fish.hook-INST grab-VM-N.FUT-3 tambaqui-big

'Michael caught a big tambaqui (*Colossoma macropomum*) with the fish hook.' (Valenzuela et al. 2013: 493)

In (34), *fish hook* is an instrumental adjunct, case-marked by means of *-lek* suffixation. Such a construction in Shawi is impossible by means of *-re* suffixation. Shawi, by contrast, deploys the locative *-ke*:

(35) *Kusi nisiterin pahkaturu iminke sawenike, tanan imianaterinsu', tihkisawatun.*

Kusi nisi-te-r-in-Ø pakaturu imin-ke

Joseph cut-VM-N.FUT-3MIN.A-3MIN.O brushwood field-LOC

saweni-**ke**, tanan imiana-te-r-in-su',

machete-LOC forest burn-V.M-N.FUT-3MIN-NMLZ

tiki-sa-watu-n-Ø.

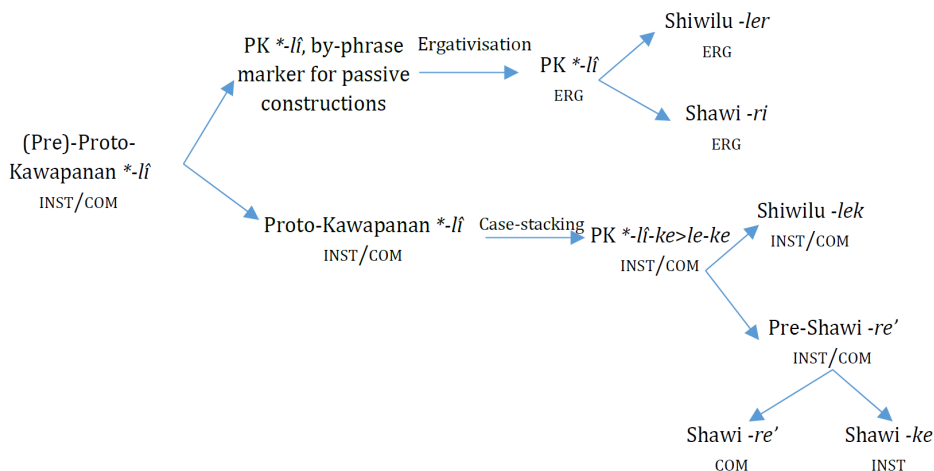
finish-PROG-SEQ-3MIN.A-3MIN.O

'Joseph, after finishing burning [that piece of] forest, cut the brushwood in the field with his machete' (GJT\_EM\_Chayahuita\_AHT\_170317).

We claim that modern Shawi/Shiwilu ergative markers, *-ri* and *-ler* respectively, originate in the Proto Kawapanan comitative/instrumental *\*-lî*. Grammaticalisation patterns such as INSTRUMENTAL>ERGATIVE as well as INSTRUMENTAL/ERGATIVE synchronic bifunctionality patterns have been documented for other languages of the world (q.v. Dixon 1972 for Dyirbal; Butt and Deo n.d. for Indo-Aryan; Fleck 2010 for Mayoruna (Panoan); , also Heine and Kuteva 2002: 180; and Lehmann 2002: 73, for an overview from the perspective of grammaticalisation studies). But, how could the Proto-Kawapanan comitative/instrumental have ended up becoming an ergative?

From the perspective of Kawapanan historical phonology, Shiwilu ergative *-ler* and Shawi ergative *-ri* are cognate. They both originate in Proto-Kawapanan ergative *\*-lî*. On the other hand, Shiwilu comitative *-lek* and Shawi comitative *-re'*, originate in Proto-Kawapanan *\*-le'*. Consequently, the co-occurrence of ergative *\*-lî* <> with instrumental/comitative *\*-le'* must have already existed in the proto-language. We hypothesise that in the early stages, when ergativity was still under development through passivisation by means of *\*-lî* marking, it was necessary to develop a semantic contrast between differential subject marking and instrumental/comitative marking. Therefore, as is the case with several other markers in Kawapanan, the language resorted to case-stacking, adding the locative marker *\*-ke* to the instrumental/comitative by means of suffixation: *\*lî* > *\*lî-ke*. After vowel harmony, PK *\*/i/* became */ə/* <e>, hence *\*le-ke*. The final vowel would have been dropped, as is still the case in Modern Shiwilu, resulting in *-re'* for Shawi and *-lek* for Shiwilu. The process went even further in Shawi, where the instrumental is marked by means of *-ke* suffixation to the NP. The tree below summarises the changes:

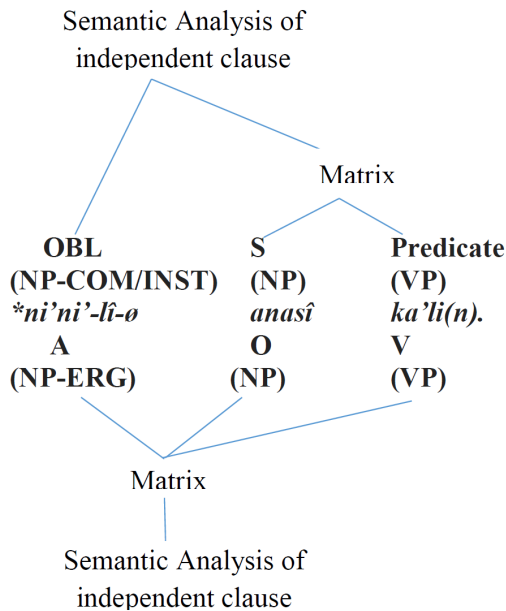
Figure 6: The development of the Kawapanan ergative and comitative case markers



In Shawi, unlike Shiwilu, both the comitative and the ergative retained object marking, which shows their common origin.

From the perspective of historical syntax, two processes must have occurred in the case of the emergence of ergative constructions in Kawapanan: **reanalysis** and **analogical extension**. Given that we do not have records for Kawapanan languages as they were spoken prior to the arrival of the Spaniards, it is difficult to be conclusive on this. However, we can rely on comparative reconstruction, based on our previous observations with regard to the similarity in form and grammatical behaviour of the modern Kawapanan comitative/instrumental and the modern Kawapanan ergative. We assume that a Proto-Kawapanan passive construction existed, which possessed an oblique agent phrase marked by means of the comitative/instrumental marker. The process possibly initiated with 3>3 sentences. This oblique agent phrase would have been initially reanalysed by the speakers, meaning that the semantic representation of the passive construction would have been changed, without involving any modification of the surface manifestation (q.v. Harris and Campbell 1995: 50, also quoted in Gildea 2004: 8) (see Figure 7).<sup>18</sup> Below we include a tree representation of a 3>3 sentence. The tree above represents the passive reading, while the tree below represents the ergative reading post reanalysis.

Figure 7: A tree representation of the development of an ergative construction from a passive construction through reanalysis inspired in Gildea (1998) historical trees and Seuren (2018) for the sentence *Ni'ni'li anasí ka'lin* 'The opossum was eaten by the dog/The dog ate the opossum'



In the case of 3>3 independent clauses, verbal agreement remains the same. Both the passive and the ergative construction have exactly the same verb form (verb-N.FUT-3MIN.S/A), i.e. the same surface structure. For this new semantic analysis to permeate the other types of clauses (1>2, 2>1, etc.), analogical extension was necessary. Here we understand analogical extension in the same sense as Gildea (2004: 11), as a process whereby what is unobservable through reanalysis is made observable. As such, in a 1>2 passive construction in Proto-Kawapaman, the verb would agree with the S-NP, i.e. the second person. However, once reanalysed into an ergative construction, analogical extension would take place, thus triggering a new agreement patterning: the verb would now agree with the new ergatively-marked A-NP, the former OBL. This latter process would still be observable only in modern Shawi (1>2 sentences), since the ergative marker still retains agreement morphology as its comitative/instrumental predecessor.

## 5.2. Hypothetical socio-historical motivations

Language, both a social institution and a cognitive and physiological faculty, is subject to functionality constraints at different levels. The NH (Silverstein 1976; Dixon 1994; Woolford 2009) may well be one of these functionality constraints in language, and possibly cognition (Gildea 2004).

This is easily observed when studying different split-ergativity systems in the languages of the world, whose differential subject-marking is constrained by the hierarchy. However, given the inevitably social nature of language, language is reshaped as an excuse to signal community ascription, in the way of cryptolects, high-class mannerisms (Seuren and Hamans 2010: 136), or discrimination (Cameron 1995). Language is used to recursively construct an identity while speaking (Silverstein 2003).

The notion of Shawi as a community-based unified ethnic group came about only after the settlement of the first Jesuit missions in the Upper Amazon. Daily attacks from the so-called “wild” Mayna prompted different groups of Indians in the Northern Peruvian piedmont to seek refuge in the *Reducciones* established by the Jesuits (Fuentes 1988; Ochoa-Gilone 2007; Ochoa Siguas 2016; González Saavedra 2015). These *Reducciones* were ethnically diverse. They included Tupian groups, such as the Kukama and the Paranapura; Kawapanan groups, such as the Chayahuita, the Cahuapana, the Jebero and the Concho; Jivaroan groups, such as the Awajún; Candoan groups, and Munich groups (Ochoa-Gilone 2007). Although the information is scarce in order to claim the existence of plurilingual *Reducciones*, this seems to have been the case initially. In spite of this multilingual reality, the *lingua franca* of these Jesuit missions seems to have been no other than Mayna-Chawi, the predecessor of Modern Shawi. This is confirmed by the existence of prayers in the language that were used by the priests to conduct catechisation in these missions (Hervás 1787; Beuchat and Rivet 1909; Rojas-Berscia 2015). Thus, speakers of different languages and varieties would have had to switch to Mayna-Chawi at some point, most probably shortly after the settlement in the missions. All these groups would have thereafter been subsumed under the Shawi identity, which encompassed a newly born indigenous *mode de vie*, a syncretism of indigenous customs and western catholic traditions, and a new language.

We hypothesise that the newly born Shawi community was in need of an identity and a “standard” language.<sup>19</sup> This could have been the right scenario for the consolidation of the AEC in Shawi. It has been argued that antifunctionality played a significant role in the development of pervasive Predicate-Raising in German and Dutch, possibly as a community-isolating device, as well as in word-order scrambling in Latin for poetry (Seuren and Hamans 2010, 159). Shawi, for example, violates the NH, but does it not only like Arrernte, which oddly case-marks first person pronouns, but systematically mirroring the NH and creating a unique system, non-existent

in other adjacent languages, and probably foreign to Amerind in general. Grammatical innovations as the one presented here seem to be socially driven. Gildea (2004: 16, bolds are ours) adds:

There is now ample evidence that most grammatical innovation does not arise out of a functional void, such as to ameliorate the inability to express certain semantic distinctions, but rather arises as the conventionalisation of a new and **socially innovative** way to express a concept that could easily be expressed with the existent grammatical resources of a language [...] In fact, usually an innovative construction simply provides a **sociolinguistically distinct** (and often more specific) way to express a distinction already perfectly well expressed in the grammar.

It could have been the case that the development of the AEC in Shawi was a social innovation. Given the irretrievability of this change in terms of linguistic awareness, this is just a socio-historical tentative hypothesis for the development of this pattern in the language.

This is very likely not the last word with regard to this pattern in Shawi. In other parts of the world, it has been found that the expression of ergative distinctions is constrained by use. Ochs (1982), in her study of Samoan ergativity, dubs these languages **SOCIALLY ERGATIVE LANGUAGES**. More research that involves longer stays in the Shawi area and that deals with a larger corpus of spontaneous conversations could possibly shed more light in the case of Shawi. Could the AEC be violated in different speech registers as in Samoan? Further follow-up studies on the acquisition of Shawi by children could also shed light on the processes lying behind the consolidation of the AEC in the grammars of children acquiring Shawi as L1. Would small children follow the NH and violate the adult AEC? These are some of the questions that remain to be answered in future studies.

## 6. Final ideas

This is a first attempt to provide a solid description of the split ergativity system of Shawi, backed-up by experimental testing. Most “descriptive” grammars and sketch grammars are solely based on the intuitions of a few speakers or a key consultant with regard to certain grammatical phenomena. An approach which takes into account the grammaticality judgments of a large sample of the participant population such as the one we sketched in the present article would be of great use to confirm or disconfirm hypotheses regarding the grammatical systems of these languages.

Moreover, future experimental work involving modern psycholinguistic techniques, as well as a detailed follow-up of the acquisition of Shawi by children, would enrich our understanding of this antifunctional pattern in the language.

Although it seems that ergativity is only exclusive to Kawapanan in the area, recent claims made by Sofia de la Torre (p.c.) point to the fact that Cholon, an extinct language once spoken to the south of the Kawapanan area, may have displayed a similar type of split ergativity.<sup>20</sup> More detailed philological studies of this language, and the rediscovery of so far “lost” grammars of neighbouring indigenous languages written during colonial times (Lucero n.d.; Teruel n.d.), will help us enrich our typological understanding of this area.

## References

- Barraza de García, Yris. 2005. ‘El sistema verbal en la lengua shawi’. Tesis de doctorado, Recife: Universidade Federal de Pernambuco.
- Beuchat, H., and P. Rivet. 1909. ‘La famille linguistique Cahuapana’. *Zeitschrift für Ethnologie* 41, 616–34.
- Bourdeau, Coentín. 2015. “Ergativity in Shawi”. MA Thesis, Nijmegen: Radboud Universiteit Nijmegen.
- Butt, Miriam, and Ashwini Deo. n.d. ‘Ergativity in Indo-Aryan’. Accessed 1 May 2018. <http://ling.uni-konstanz.de/pages/home/butt/main/papers/ia-erg.html>.
- Cameron, Deborah. 1995. *Verbal Hygiene*. 1st ed. Routledge Linguistics Classic. London: Routledge.
- Craats, Ineke van de, and Jeanne Kurvers. 2014. ‘10 jaar LESLLA. Onderzoek en educatie, een gelukkige combinatie’. *Les* 187, 22–24.
- Craats, Ineke van de, Jeanne Kurvers, and Roeland van Hout, eds. 2015. *Adult Literacy, Second Language and Cognition*. Nijmegen, The Netherlands: Centre for Language Studies.
- Dixon, R.M.W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.
- . 1994. *Ergativity*. *Cambridge Studies in Linguistics* 69. Cambridge: Cambridge University Press.



- Dradi, María Pía. 1987. *La mujer chayahuita: ¿un destino de marginación?; análisis de la condición femenina en una sociedad indígena de la Amazonía*. Lima: Instituto Nacional de Planificación.
- Fleck, David W. 2010. 'Ergativity in the Mayoruna Branch of the Panoan Family'. In *Ergativity in Amazonia*, 29–63. *Typological Studies in Language* 89. John Benjamins Publishing Company.
- Fuentes, Aldo. 1988. *Porque las piedras no mueren*. Lima: Centro Amazónico de Antropología y Aplicación Práctica.
- Gildea, Spike. 1998. *On Reconstructing Grammar: Comparative Cariban Morphosyntax*. Oxford, UK: Oxford University Press.
- . 2004. 'Are There Universal Cognitive Motivations for Ergativity?' In *L'ergativité En Amazonie*, edited by Francisco Queixalós, 2:1–37. Brasília, DF: CNRS, IRD & Laboratório de Línguas Indígenas, UnB.
- Gildea, Spike, and Francisco Queixalós, eds. 2010. *Ergativity in Amazonia*. *Typological Studies in Language*, v. 89. Amsterdam, The Netherlands ; Philadelphia, PA: John Benjamins Pub. Co.
- González Saavedra, María Luisa. 2015. 'Un Lugar Para Los Shawi En La Historia de Maynas'. *Anthropologica* 33 (34), 249–266.
- Harris, Alice C., and Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hervás, Lorenzo. 1787. *Saggio Pratico delle Lingue*. Cesena: Gregorio Biasini all'Insegna di Pallade.
- Hoop, Helen de, and B. Narasimhan. 2009. 'Ergative-Case Marking in Hindi'. In *Differential Subject Marking*, edited by Helen de Hoop and Peter de Swart, 63–78. Dordrecht: Springer.
- Hoop, Helen de, and Peter de Swart. 2009. 'Cross-Linguistic Variation in Differential Subject Marking'. In *Differential Subject Marking*, edited by Helen de Hoop and Peter de Swart, 1–16. Dordrecht: Springer.
- Hualde, José Ignacio, and Jon Ortiz de Urbina, eds. 2003. *A Grammar of Basque*. *Mouton Grammar Library* 26. Berlin: Mouton de Gruyter.
- Huettig, Falk, Régine Kolinsky, and Thomas Lachmann. 2018. 'The Culturally

- Co-Opted Brain: How Literacy Affects the Human Mind'. *Language, Cognition and Neuroscience* 33 (3): 275–77. <https://doi.org/10.1080/23273798.2018.1425803>.
- Huettig, Falk, and Ramesh K. Mishra. 2014. 'How Literacy Acquisition Affects the Illiterate Mind - A Critical Examination of Theories and Evidence: How Literacy Affects the Illiterate Mind'. *Language and Linguistics Compass* 8 (10): 401–27. <https://doi.org/10.1111/lnc3.12092>.
- Kurvers, Jeanne. 2002. *Met ongeletterde ogen. Kennis van taal en schrift van analfabeten [With illiterate eyes. Awareness of language and writing of illiterates]*. Amsterdam: Amsterdam University Press.
- Lehmann, Christian. 2002. *Thoughts on Grammaticalization. Arbeitspapiere Des Seminars Für Sprachwissenschaft Der Universität Erfurt 9*. Erfurt: Seminar für Sprachwissenschaft der Universität Erfurt.
- Lucero, Juan S.J. n.d. *Gramática y Catecismo de Muchas Lenguas de Quito, y Principalmente de Los Idiomas Paranaपुरo y Cocamo*.
- Matthewson, Lisa. 2004. 'On the Methodology of Semantic Fieldwork'. *International Journal of American Linguistics* 70 (4), 369–415. <https://doi.org/10.1086/429207>.
- McGregor, William B. 2009. 'Typology of Ergativity'. *Language and Linguistics Compass* 3 (1), 480–508. <https://doi.org/10.1111/j.1749-818X.2008.00118.x>.
- Ochoa Siguas, Nancy. 2016. 'Los Piyapi Yamorai o «gente Del Río de La Sal». Los Últimos Proveedores de Sal Del Paranaपुरas, Alto Amazonas, Perú'. *Bulletin de l'Institut Français d'études Andines*, no. 45 (1) (April), 91–108. <https://doi.org/10.4000/bifea.7819>.
- Ochoa-Gilonne, Nancy. 2007. 'Entre plusieurs mondes. Les Chayahuita de l'Amazonie'. PhD Thesis, EHESS.
- Ochs, Elinor. 1982. 'Ergativity and Word Order in Samoan Child Language'. *Language* 58 (3), 646–71.
- Piepers, Joske. 2016. 'Optional Ergative Case Marking in Hindi'. Bachelor Thesis. Radboud University Nijmegen.
- Queixalós, Francisco, and Spike Gildea. 2010. 'Manifestations of Ergativity in Amazonia'. In *Ergativity in Amazonia*, edited by Spike Gildea and Francisco Queixalós. *Typological Studies in Language* 89. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Rijk, Rudolf P. G. de, and Armand de Coene. 2008. *Standard Basque: A Progressive Grammar*. *Current Studies in Linguistics* 44. Cambridge, Mass: MIT Press.
- Rojas-Berscia, Luis Miguel. 2013. “La sintaxis y semántica de las construcciones causativas en el chayahuita de Balsapuerto”. *Licentiate Thesis*, Lima: Pontificia Universidad Católica del Perú.
- Rojas-Berscia, Luis Miguel. 2015. “Mayna, the lost Kawapanan language”. *LIAMES* 15: 393–407.
- Rojas-Berscia, Luis Miguel, y Corentin Bourdeau. 2017. “‘Optional’ or Syntactic Ergativity in Shawi: Distribution and Possible Origins”. *Linguistic Discovery* 15 (1): 50–65.
- Rojas-Berscia, Luis Miguel. 2021. *Pre-Historical Language Contact in Peruvian Amazonia: A Dynamic Approach to Shawi (Kawapanan)*. Vol. 58. *Contact Language Library*. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/coll.58>.
- Salvatier, John, Thomas V. Wiecki, y Christopher Fonnesbeck. 2016. “Probabilistic Programming in Python Using PyMC3”. *PeerJ Computer Science* 2 (abril): e55. <https://doi.org/10.7717/peerj-cs.55>.
- Seuren, Pieter A. M. 1969. *Operators and Nucleus: A Contribution to the Theory of Grammar*. Cambridge: Cambridge University Press.
- . 1972. ‘Autonomous Versus Semantic Syntax’. *Foundations of Language* 8 (2), 237–265.
- . 2009. *Language in Cognition [Language From Within, Vol.1]*. Oxford: Oxford University Press.
- . 2018. *Semantic Syntax*. Second Revised Edition. Leiden: Brill.
- Seuren, Pieter A. M., and Camiel Hamans. 2010. ‘Antifunctionality in Language Change’. *Folia Linguistica* 44 (1). <https://doi.org/10.1515/flin.2010.005>.
- Silverstein, Michael. 1976. ‘Hierarchy of Features of Ergativity’. In *Grammatical Categories in Australian Languages*, 112–71. New Jersey: Humanities Press.
- . 2003. ‘Indexical Order and the Dialectics of Sociolinguistic Life’. *Language & Communication* 23 (3–4), 193–229. [https://doi.org/10.1016/S0271-5309\(03\)00013-2](https://doi.org/10.1016/S0271-5309(03)00013-2).
- Swadesh, Morris. 1971. *The Origin and Diversification of Language*. Edited by

Joel Sherzer. Chicago: Aldine.

Teruel, Luis S.I. n.d. Gramatica de La Lengua Tabalosa Del Peru. 16??

Valenzuela, Pilar M. 2011. 'Argument Encoding and Pragmatic Marking of the Transitive Subject in Shiwilu (Kawapanan)'. *International Journal of American Linguistics* 77 (1), 91–120. <https://doi.org/10.1086/657989>.

Valenzuela, Pilar M., Meneleo Careajano, Emérita Guerra, Julia Inuma, and Fernando Lachuma. 2013. *Diccionario Shiwilu-Castellano Castellano-Shiwilu*. Lima: Federación de Comunidades Nativas de Jeberos (FECONAJE).

Woolford, E. 2009. 'Differential Subject Marking at Argument Structure, Syntax and PF'. In *Differential Subject Marking*, edited by Helen de Hoop and Peter de Swart, 17–40. Dordrecht: Springer.

## FOOTNOTES

1 This study was originally part of the doctoral project of the first author, financially sponsored by the Language in Interaction Research Consortium (LiI), the Max Planck Institute for Psycholinguistics, and the Centre for Language Studies at Radboud University Nijmegen. A postdoctoral fellowship within the Centre of Excellence for the Dynamics of Language (CoEDL) at the University of Queensland also allowed the first author to refine some of the analyses. The implementation on Psychopy was possible thanks to Gary Lupyan, while the Bayesian modelling should be credited to Bill Thompson, who also continuously followed the progress of the experiment with the first author. The authors acknowledge the support of all the Shawi communities involved. Their warm welcome will always be appreciated. In addition, we thank the insightful comments and suggestions from Helen de Hoop, Steve Levinson, Pieter Muysken, Mark Dingemans, Miguel Rodríguez Mondoñedo, Simon Overall, Roberto Zariquiey, Brigitte Bauer, Friederike Lüpke, and the anonymous reviewers. All remaining errors are, of course, solely ours. We also thank Raj Seuren for kindly allowing us to publish this co-authored study of his father posthumously.

The following conventions are used throughout the paper, following, in most cases, the Leipzig Glossing Rules (<http://www.eva.mpg.de/lingua/resources/glossing-rules.php>): 1 ‘first person’, 2 ‘second person’, 3 ‘third person’, A ‘subject of transitive clause’, ANIM ‘animate’, AUG ‘augmented’, AUX ‘auxiliary’, CAUS ‘causative’, COM ‘comitative’, COP ‘copula’, DET ‘determiner’, DIM ‘diminutive’, DUBIT ‘dubitative’, ERG –‘ergative’, EXCL ‘exclusive’, HUM ‘human’, INAM ‘inanimate’, INCL ‘inclusive’, INST ‘instrumental’, LOC ‘locative’, MIN ‘minimal’, NEG ‘negative’, N.FUT ‘non-future’, O ‘object’, PL ‘plural’, PROG ‘progressive’, R ‘recipient’, S ‘subject of intransitive clause’, SG ‘singular’, VM ‘valency modifier’

2 Bracketed tokens indicate they are ergatively-marked.

3 Assuming that a grammar is “a socially-accepted system for the conversion of given semantic thought inputs (consisting of a speech act operator and a propositional content) into recipes for well-defined acoustic or written outputs and to a large extent and also vice versa, from phonologically interpreted input to thought content” (Seuren and Hamans 2010, 141; Seuren 2018).

4 This paper does not focus on 3>3 sentence types (sentences whereby a third person subject acts upon a third person object. That has already been discussed in Rojas-Berscia and Bourdeau (2017). Here we focus on the other sentence types where ergativity occurs.

5 This implies that ergativity happens in the Nucleus or Matrix of the sentence (q.v. Seuren 1969 for this view within Semantic Syntax, 2018)

6 Throughout the article  $x>y$  indicates that person  $x$  is acting upon person  $y$ .

7 Brackets indicate optionality. Object-marking following the suffixation of *-ri* can be found in all the Shawi-speaking territory. Nevertheless, it is not obligatory. Some regions, such as Cahuapanas, prefer the version without *-nke*, and leave just an *-n*. All regions accept the *-n* version. This was the one used in our experiment. Shiwilu does not display this pattern. We suggest this is a fossil of the Passive-to-Ergative process in Kawapanan. This is further discussed in the Discussion section.

8 The arrows indicate the possibility of ergative marking. There is no arrow from 2>1, because the use of the ergative is banned in this case.  $\forall$  in this case stands for ‘any type of NP’.

9 It must be noted, however, that here we are dealing with GRAMMATICALITY JUDGMENTS, not TRUTH VALUE JUDGEMENTS OR FELICITY JUDGMENTS (Matthewson 2004, 399). In this case, we surmise that the absence of a discourse context is not so important to assess the well-formedness of a sentence.

10 It must be noted that this is not the first time the phenomenon is studied. It was partially identified in Barraza de García (2005) and in (Bourdeau 2015, Rojas-Berscia & Bourdeau 2017), after long sessions of direct elicitation, and the use of Spanish as a meta-language for direct translation (q.v. Matthewson 2004 for the relevance of direct elicitation and the use of a meta-language for the study of meaning and grammar).

11 Psychopy is a software package for running behavioural experiments commonly used in experimental psychology.

12 The stimulus set (sentences) is available in the Appendix section. The audio files can be accessed following this link: <https://hdl.handle.net/1839/6699bc85-bf12-4fec-8467-f52760a8c474>

13 <https://github.com/bambinos/bambi>

14 This may have to do with the fact that Balsapuerto is the only Shawi town directly in contact with mestizo cities. The Yurimaguas-Balsapuerto highway was recently completed, allowing Shawi to migrate to the big cities. This situation is not only reshaping the cultural practices of Balsapuerto, but also has an impact on the use of Shawi and the learning of Spanish. Schooling and Spanish acquisition as an L2 may be factors to be taken into account when assessing these results. However, it must also be noted that Cahuapanas speakers are all bilingual, but “respect” the AEC consistently.

15 No other oblique case marker in Shawi carries person-object markers.

16 In (32) and (33), the comitative marker *-re*’ becomes *-ru*, due to assimilation of the following syllable’s vowel quality.

17 We assume that, in  $x > 3$  sentences, ergative *-ri* carries the third person object marker  $-\emptyset$ . This corresponds to comitative adjunct phrases of third person subject sentences, which also carry the third person marker  $-\emptyset$ .

18 The authors, however, deploy the term ‘underlying structure’. We adapted this term to ‘semantic representation’, given our own take on transformational grammar (q.v. Seuren 1972, 2018).

19 This is developed in Rojas-Berscia (2021).

20 Possibly closer to that of Shiwilu (Valenzuela 2011) than to the ANH pattern of Shawi.