Pronominal agreement and the changing face of Austronesian voice: The view from Sipora Mentawai

Emma Keith
Australian National University

I Wayan Arka
Australian National University,
Udayana University

Proceedings of the LFG'25 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2025

PubliKon

lfg-proceedings.org

Abstract

Sipora Mentawai (SM) shows a restructured Malayo-Polynesian Western Region voice system in which inherited Actor Voice (AV) morphology is retained but no longer supports a productive reduced 'Indonesian-type' AV/UV alternation. Instead, the language occupies a semi-alternating position on an alternation continuum: residual voice material persists, alongside TAM and information-structure correlates, but there is no longer an alternation between two grammatical voices; only between AV and non-voice-marked verbs, which in main clauses are, save some edge cases, obligatorily marked for agreement. AV *m-* assigns SUBJ to the ACTOR macrosemantic role; ensures it has a privileged position in the information structure; and assigns REALis mood to verbs on which it appears. Agreement morphology can be either grammatical or anaphoric, and carries no mood assignment. This account presents a detailed and novel case of how reduced 'Indonesian-type' voice systems are subject to attrition into 'semi-alternating' systems.

1 Introduction

The languages of Indonesia, particularly those in the Western Region linkage of Malayo-Polynesian (MP) (cf. Smith 2025), are widely discussed in typological work for exhibiting 'symmetrical voice' systems (Arka 2003b; Foley 1998, 2008; Riesberg 2014; Ross 2002).[†] In this tradition, symmetrical voice refers to systems with (at least) two voice constructions that differ in which core argument is privileged (SUBJ/PIVOT/PSA), without a straightforward active/passive derivational asymmetry; importantly, symmetricality is diagnosed by a bundle of grammatical properties (especially core status and PIVOT effects), not by morphology alone. Sipora Mentawai (SM)[1], however, shows a restructured (and more strongly reduced) profile: an AV prefix *m-*[2] is retained, reflecting the widespread Proto-Austronesian Actor Voice infix **-um-*[3], but SM does not show a productive AV/UV alternation in main clauses of the type familiar from Indonesian-type symmetrical voice languages. Instead, SM has a system of pronominal agreement – including both

[†]Many thanks are due to the audience at the 30th International Lexical Functional Grammar Conference at Universitat Pompeu Fabra, Barcelona, especially Miriam Butt, Ash Asudeh, Joan Bresnan, and Paul Zodl, for their useful feedback and critique. Floofy McFloofkins, Keira Mullan, and anonymous reviewers also offerred comments after the conference which influenced the development of the paper, and Zoe Orlanda Elwyn helped with typesetting. Sentences and grammaticality judgments come from Sumario Tatubeket, Hasael Samangilailai, and Arkelaus Saogo. All remaining errors are our own. This research received financial support from the Australian Research Council grant *Languages of Barrier Islands, Sumatra: Description, Typology, and History* (DP230102019). Emma Keith was supported by an Australian Government Research and Training Program (AGRTP) scholarship.

[1]Sipora Mentawai (SM) (mwv, ISO639-3, Sumatran, Austronesian (AN)) is a Mentawai isolect spoken by around 25,000 people as an L1 on Sipora, one of the Mentawai islands off the coast of West Sumatra within Indonesia's Barrier Islands. It is by far the most widely spoken Mentawai isolect, and enjoys widespread use as an L2 and *lingua franca* by speakers of other, non-mutually intelligible Mentawai isolects on the other islands as well as in the Mentawai diaspora on mainland Sumatra. It is classified as a Sumatran language (Billings & McDonnell 2024), alongside its geographic neighbours Nias and Enggano, and as such sits within the Western Region linkage of the Malayo-Polynesian group of Austronesian languages (Smith 2025).

[2]Here we use 'voice' in the broad cross-linguistic sense of a grammatical category encoding linking/diathesis.

[3]Proto-Austronesian (PAN) is the ancestral proto-language of Proto-Malayo-Polynesian (PMP) as well as numerous Formosan languages; both PMP and PAN are ancestral to SM at different time-depths.

anaphoric and grammatical agreement in the sense of Bresnan & Mchombo (1987) – and verbs that take grammatical agreement do not take AV morphology. They may, however, participate in superficially 'passive-like' orderings where the OBJ is fronted. Embedded clauses, which may show 'bare' (i.e., non-voice-marked and non-agreement-marked) verbs, reveal a number of intriguing patterns about the relation between the relic voice morphology and the mapping of grammatical functions, supporting a view where AV *m-* synchronically identifies the grammatical function SUBJ with the most prominent argument, rather than participating in a symmetrical voice alternation with UV.

This paper presents an LFG analysis of the systems of voice and agreement in SM, drawing on recent fieldwork data. §2 defines a number of terms which are used throughout the remainder of the text, including 'Indonesian-type' (§2.1); 'symmetrical voice' (§2.2) and related terms, before characterising the voice system in SM as 'semi-alternating' (§2.5). §3 describes the interrelated systems of voice, agreement, aktionsart morphology, and mood in SM. The unique category of 'aktionsart' morphology is described (§3.1), along with its ability to 'host' AV *m-*. Agreement morphology is then outlined (§3.2), along with its inability to co-occur with AV *m-* on a verb; then, the complicated interactions between the three categories of aktionsart morphology, voice, and agreement are detailed, first in matrix clauses (§3.3) and then in subordinate clauses (§3.4). §3.5 describes the remarkable freedom in the ordering of constituents in SM, and the ways in which both AV *m-* and agreement morphology serve to selectively restrict possible orderings. Having described the properties we want to model, §4 sets about accounting for these in LFG, through a mixture of Phrase Structure Rules and lexical entries for relevant morphemes.

2 On 'Indonesian-type' voice systems

Before an explanation of the voice system in SM can be undertaken, it is necessary to define a few key terms which are frequently used in the discussion of 'Indonesian-type' voice systems. These are: 'Indonesian-type' (§2.1); 'symmetrical voice' (§2.2); 'voice alternation' (§2.3); 'non-alternating language' (§2.4); and 'semi-alternating voice' (§2.5).

2.1 'Indonesian-type' voice systems

The term 'Indonesian-type', which tends to be used to contrast with a 'Philippine-type' in Austronesianist linguistics (Ross 2002; Chen & McDonnell 2019), refers to a voice system with two (as opposed to four, in Philippine-type) minimal voices: Actor Voice and Undergoer Voice. The notion of 'Indonesian-type' voice systems is often conflated with the development of valence-increasing morphology from reflexes of PMP **i* and **akən* or **an* (Kaufman 2009a). We use 'Indonesian-type' as a typological heuristic for a reduced symmetrical voice profile, not as a claim of homogeneity across 'Indonesian' or Western MP Region languages: Indonesian-type languages are, for our purposes, those which display a minimal alternation between AV and UV. Indonesian-type languages are 'reduced' in comparison to Philippine-type languages in at least two ways: 1) the reduction in the number of grammatical voices, and 2) the greater possibilities for degree-of-symmetry reduction (Arka 2017), where the consequences of the morphological marking across arguments become more uneven.

2.2 Symmetrical Voice

Voice symmetricality[4] refers to the 'equality' in markedness of AV and UV: in a perfectly 'symmetrical' voice language, neither AV nor UV is more marked, either morphologically or in terms of the demotion of some argument to non-core (Riesberg 2014). Symmetricality exists on a continuum, in terms of how Actor and Undergoer in the different voices behave with respect to various tests of 'core'-hood: accessibility to extraction/relativisation; control and binding; and the possibility of OBL-like encoding, among others (Arka 2017). Perhaps the canonical example of a symmetrical voice is Balinese (Arka 2003a): note that in the sentence below the ACTOR *tiang* cannot be elided even though it is the syntactic OBJ.[5]

(1) a. [$_{\text{SUBJ/A}}$ tiang] ny-epak [$_{\text{OBJ/U}}$ cicing-e]
1SG AV-kick dog-DEF

b. [$_{\text{SUBJ/U}}$ cicing-e] ∅-sepak [$_{\text{OBJ/A}}$ *(tiang)]
dog-DEF UV-kick 1SG

'I kicked the dog.'[6] (Artawa 1998: 8) (Balinese)

2.3 Voice alternation

Languages which show 'alternating' voice systems are those in which different grammatical voices alter the relations between SUBJ/OBJ and ACTOR/UNDERGOER, *without* the implication that the relation between the grammatical voices is necessarily symmetrical. Hence, languages with true passive constructions are alternating but not symmetrical. True passives, in contrast to UV constructions, obligatorily demote the Actor to OBL status, typically accompanied by stronger detransitivising and discourse-backgrounding effects than UV. An example of a language with a true passive is Indonesian itself (Arka & Manning 1998):

(2) a. [$_{\text{SUBJ/A}}$ Amir] mem-baca [$_{\text{OBJ/U}}$ buku itu]
Amir AV-read book that
'Amir read the book.'

b. [$_{\text{SUBJ/U}}$ buku itu] di-baca [$_{\text{OBL/A}}$ (oleh Amir)]
book that PASS-read by Amir
'The book was read (by Amir).' (Indonesian)

[4]This term originates with Foley (1998) (published as Foley 2008), and was subsequently developed in the Austronesianist literature, including in Arka (2003b).

[5]The notation *( ) indicates that a bracketed element may not be deleted; ( ) indicates that an element is freely deletable. Abbreviations used in glossing, in addition to those found in the Leipzig rules, are: AV=Actor Voice; DIR=Directional; INC=Inceptive; INCH=Inchoative; PERS=Personal; REAL=Realis U=Undergoer; UV=Undergoer Voice

[6]The use of only one translation for the two sentences reflects the oft-remarked upon (e.g., Arka & Manning 1998: 2; Woollams 1996: 46) fact that it is difficult to capture the 'unmarkedness' of U-initial orderings in symmetrical voice languages in English translation, since the English passive involves demotion of the A to non-core. The use of passives in translations of OBJ-fronted sentences in SM and passives in Indonesian reflects the fact that these orderings are, indeed, pragmatically marked in the source language.

2.4 Non-alternating language

In non-alternating languages, the linking of semantic roles to grammatical functions is typically fixed, with comparable discourse functions such as topicalisation being handled by other resources such as word order. These languages tend to have no overt voice morphemes. An example of an MP language with a non-alternating system is Kambera (Klamer 1996): there is no voice morphology in the language, and a 'passive-like' ordering is created instead through a purely syntactic dislocation where the OBJ is moved into sentence-initial position without any change to the morphology of the verb. Such languages represent the extreme end of the attrition of voice morphology on the continuum of voice symmetricality.

(3) a. ka [SUBJ nyuna] na-tinu-nya [OBJ na lau]
CONJ 3SG 3SG.NOM-weave-3SG.DAT ART sarong
'So that she weaves the sarong.' (lit. 'She she-weaves-it the sarong.')

b. ka [OBJ na lau] na-tinu-nya [SUBJ nyuna]
CONJ ART sarong 3SG.NOM-weave-3SG.DAT 3SG
'So that the sarong was woven by her.' (lit. 'The sarong she-weaves-it she.')
(Klamer 1996: 13) (Kambera)

2.5 Semi-alternating voice

'Semi-alternating voice' refers to systems where voice-related alternations survive only partially, being restricted, uneven, or construction-specific, and no longer form a fully productive, grammar-wide alternation paradigm. In the Austronesian case, this typically means that former AV/UV morphology remains in form, but its productivity and paradigmicity is in some way diminished. Semi-alternating Austronesian languages can contain some remnants of alternating voice systems, but these are not systematic throughout the whole language: productive AV/UV opposition in main clauses is not clearly supported. Semi-alternating systems can develop out of alternating systems through the attrition of voice morphology at both a formal and functional level (Arka 2024), where the presence of all the listed options would imply an alternating system; the presence of none of them a non-alternating system; and some subset of them a semi-alternating system:

1. Formal level: are there morphological forms expressing both AV and UV?
2. Functional level:
   (a) As a selector of the mapping between syntactic (SUBJ/OBJ) and semantic macro-roles (ACTOR/UNDERGOER).
   (b) As a selector of PIVOT function.
   (c) As a marker of discourse-pragmatic prominence.
   (d) As a TAM encoder.

As this paper will argue, SM appears best described as a language of this type: AV *m-* identifies SUBJ with ACTOR; requires SUBJ to be discourse-prominent; and encodes TAM (here, REAL mood), but there is no corresponding UV form, nor evidence for a true UV or passive construction in main clauses where the OBJ is identified with ACTOR morphologically.

3 Voice and agreement in Sipora Mentawai

3.1 Aktionsart morphemes and the form of voice

Understanding voice in SM requires first understanding what we will call for now 'aktionsart morphemes'.[7] These are a number of prefixes which appear on verb stems and perform predictable yet diverse operations in terms of aktionsart. All verbs in main clauses, aside from a small lexically-specified list of exceptions, *must* take either an aktionsart morpheme or an agreement morpheme (§3.2). The aktionsart morphemes are listed in full below.[8]

Intransitive morphemes	Transitive morphemes
ma- 'durative' *pu-* 'inceptive' *tu-* 'inchoative/involitive'	*pasi-* 'transitive' *paN-* 'distributive'
pa- 'middle/reciprocal'	

Table 1: Sipora Mentawai aktionsart morphemes.

AV *m-* only appears when it is 'hosted' by the aktionsart morphology.[9] Aktionsart morphemes which begin with /p/ have their initial segment replaced by /m/ when AV is present - AV *m-* plus *pasi-* yields *m-asi-*; *paN-* yields *m-aN-*; and *pu-* yields *m-u-*. This rather idiosyncratic state of affairs can be traced back to Proto-Malayo-Polynesian (PMP) (Kaufman 2009b), where AV (at this point still part of an alternating system) had the form **-um-*; when it was hosted by morphemes beginning with *p-* this combined with a phonotactic prohibition against sequences of successive labials (Blust 2013, 2023) to produce alternating sets of Actor Voice *mX-* versus non-Actor Voice *pX-*, where *X* is any subsequent string of segments.

3.2 Agreement

Like in the nearby and closely related Enggano (Hemmings to appear; Hemmings & Dalrymple to appear), SM has two sets of pronominal agreement affixes which appear on

[7]This is necessarily a term of convenience. These morphemes may in some cases attach to nominal stems, but always derive verbs of a unified aspectual 'character', as implied by their glosses; their high level of integration and lexicalisation with the stem motivates the use of the term aktionsart, i.e., *lexical* aspect. A full discussion of the status and category of this rather unique morphemic category is, needless to say, well beyond the scope of this paper; they are discussed here only with regard to their capacity to host voice morphology.

[8]The capital N-, as it is conventionally written in Austronesianist linguistics, stands for Homorganic Nasal substitution (Blust 2004), a morphophoneme widespread throughout MP languages which replaces the initial consonant of its stem with a nasal at the same place of articulation or similar, typically with some degree of language-specific idiosyncracy as to exactly what realisations exist in what environments. In SM, its realisations are NC[+cor,-voice]→/n/; NC[-cor,-voice]→/ŋ/; NV→/ŋV/; NC[+voice]→C[+voice].

[9]There are minor exceptions to this: the verbs *ei* 'go' and *oi* 'come', which happen to be the only vowel-initial members of a larger closed subset of frequent verbs which do not require aktionsart morphemes in main clauses, can be host to AV *m-* without any aktionsart morphology, as in sentence (5) and elsewhere. In these instances the morpheme has the form *m-*. These are the only instances where *m-* adds a segment rather than replacing one, and in doing so provides impetus for conceiving of this as a prefix rather than infix or access process synchronically.

verbs. Unlike Enggano, however, where both sets are prefixed to the verb, in SM there is both a prefix and a suffix set.[10]

	1SG	1PL.INCL	1PL.EXCL	2SG	2PL	3SG	3PL
Prefix	*ku-*	*ta-*	*ku- kai*	*nu-*	*nu- kam*	*i-*	*ra-*
Suffix	*-ku*	*-ta*	*-mai*	*-nu*/C_ *-m*/V_	*-mui*	*-na*	*-ra/-da*

Table 2: Pronominal agreement paradigms in Sipora Mentawai

The suffix set is historically older, with each form being traceable to one of the two suffixal genitive pronoun sets of Proto-Malayo-Polynesian (PMP). These pronouns were used both to provide a POSS to NPs and also to provide an ACTOR to verbs, but only in Undergoer Voice. Note, however, that the suffixes in SM show no such restriction to Undergoer Voice, but instead encode the SUBJ in both relative clauses (4) and main clause verbs marked with PRF *-(ng)an* (5). The POSS paradigm in SM is identical in form to this verbal suffix set aside from the member for 3SG, where the verbal suffix set has *-na* and the POSS set has *=nia*, the latter being the expected reflex given the form in PMP. Correspondences are given below (PMP forms from Ross 2006); note that PMP *h is lost in all Sumatran languages (Billings & McDonnell 2024).[11]

	1SG	1PL.INCL	1PL.EXCL	2SG		2PL	3SG	3PL
SM verbal agreement suffixes	*-ku*	*-ta*	*-mai*	*-nu*/C_	*-m*/V_	*mui*	*-na*	*-da*
SM POSS	*=ku*	*=ta*	*=mai*	*=nu*/C_	*=m*/V_	*=mui*	*=nia*	*=ra/=da*
PMP POSS	**=ku*	**=ta*	**=mami*	**=nihu*	**=mu*	**=muihu*	**=nia*	**=da*
PMP set	GEN1	GEN1	GEN1	GEN2	GEN1	GEN1	GEN1	GEN

Table 3: Correspondences of suffixal agreement set with SM and PMP possessive paradigms

(4) katukolobat=nia kele' si=kua-m si=boiki'
meaning=DEF like REL-say-2SG REL=before
'The meaning is as you said before.'

(5) m-ei ekeu ka koat, itco-ku-ngan kapa'
AV-go 2SG LOC sea see-1SG-PRF boat
'Go to the sea, I saw a boat!' (Sipora Mentawai)

The prefix set, meanwhile, seems to be a local innovation: corresponding sets are found in other Barrier Island languages, including Nias (Brown 2001), Sigulai (Akoli et al. 2025), Devayan (Mullan 2025), and Enggano (Hemmings to appear; Hemmings & Dalrymple to appear). The prefix set appears in main clauses, without regard for the TAM marking of the verb.

[10] All plosive-initial suffixes are realised with gemination of the initial consonant in the environment V_; this is understood to be the result of a general morphophonological rule rather than allomorphs requiring individual specification. [r] is a free variant of /d/.

[11] POSS pronoun sets are traditionally written as enclitics in Austronesianist linguistics because these tend to attach to the rightmost constituent of the NP rather than its head. However, the derived verbal set in SM, which only attach to verbs, behave like affixes, for instance appearing between the stem and other verbal morphology.

(6) a-i-ga-gaba enung-an=nia, a-i-alak-an nia
REAL-3SG-PROG-look.for move-NMLZ=3SG.POSS REAL-3SG-take-PRF 3SG
'He's looking for a way out, (but) he's already trapped.' (of a fish in a net)

Unusually (though similar to Enggano, Hemmings to appear), the forms for 1PL.EXCL and 2PL consist not of a single prefix indicating both PERS and NUM, but rather of a prefix which indicates only PERS and the free form of the relevant pronoun postposed to the verb, as in (7).

(7) ku-ei kai ka laggai
1-go 1PL.EXCL LOC village
'We (excl.) are going to the village.

At least in the case of the SM forms, these are all transparent grammaticalisations (with induced monosyllabification) of free forms (Ross's (2006) PMP NOM1 set), with the exception of 2 *nu-*, which seems to have been remodelled after the suffix set. The 1 and 2 prefixes are historically derived from SG forms.

SM verbal agreement prefixes	*ku-* 1	*ta-* 1PL.INCL	*i-* 3SG	*ra-* 3PL
PMP free pronouns	**aku* 1SG	**ita* 1PL.INCL	**ia* 3SG	**sida* 3PL

Table 4: Correspondences of SM prefixal agreement set with PMP free (NOM1) pronouns

In sentences where either a prefix or a suffix can be used (i.e., perfective matrix clauses), equivalent sentences drawing from opposite pronominal sets produce identical translations.

(8) a. ra-abbit-nan abak=nia ka koat
3PL-take-PRF canoe=DEF LOC ocean
b. abbit-ra-ngan abak=nia ka koat
take-3PL-PRF canoe=DEF LOC ocean
'They took the canoe to the ocean.'

In the typology of Bresnan & Mchombo (1987), the agreement markers in SM can perform both grammatical (9-a) and anaphoric (9-b) agreement. In other words, they can either agree with an external SUBJ DP, or act as a pronominal SUBJ in themselves.

(9) a. simanteu i-jigu jo'jo'
male 3SG-hit dog
'The man hit the dog.'
b. i-jigu jo'jo'
3SG-hit dog
'He hit the dog.'

3.3 The interplay between voice, agreement, and mood

All verbs in main clauses must host either AV *m-* or agreement, but cannot host both (hence the ungrammaticality of (10-c) and (11-c)). Since AV only has form when hosted

by aktionsart morphemes on all verbs other than *ei* (11) and *oi*, this means for all other verbs that if AV is present, aktionsart morphology is also present to host it (e.g., *paN-* DISTR in (10-a)).

(10) a. ina m-an-eu' iba
mother AV-DISTR-cook fish
'Mother cooks fish.'
b. i-seu' iba
3SG-cook fish
'She cooks fish.'
c. *i-ma-neu iba

(11) a. m-ei sita ka gerat
AV-go 1PL.EXCL LOC outside
'Let's go out.'
b. ta-ei ka gerat
1PL.INCL-go LOC outside
'We (incl.) have to go out.'
c. *ta-m-ei ka gerat

The aktionsart morphemes *ma-* 'durative' (12), *tu-* 'inchoative' (13), and *pa-* 'middle/reciprocal' (14) never host AV morphology. Note that both *ma-* and *tu-* derive unaccusative verbs, while *pa-* derives verbs with low transitivity and bidirectional action.

(12) a. ma-goiso' abak
DUR-small canoe
'The canoe is small.'
b. *m-ma-goiso' abak

(13) a. lalep nera tu-ragat
house that INCH-collapse
'The house collapsed.'
b. *lalep nera m-u-ragat

(14) a. jo'jo' nera pa-kukru ka gerat
dog that RECP-chase LOC yard
'The dogs were chasing each other in the yard.'
b. *jo'jo' nera m-a-kukru ka gerat

Agreement and aktionsart morphology can co-occur, and in doing so reveal the underlying /p/-initial forms of some of the aktionsart morphemes (*pasi-* TR in (15) and *pu-* INC in (16)).[12] This is most common in contexts involving anaphoric rather than grammatical agreement, and extends to all aktionsart morphemes, including those that derive unaccusative verbs (i.e., *ma-* DUR, as in (17) and *tu-* INCH, as in (18)).

(15) ta' nu-itco' kam i-pasi-taptap bibilet?
NEG 2-see 2PL 3SG-TR-wash clothes
'Don't you (pl.) see her washing the clothes?'

[12]These forms also surface in imperatives, where aktionsart is optional but sometimes appears.

(16) a-i-pu-oni teteu
REAL-3SG-INC-name grandfather
'It is named 'grandfather'.' (lit. 'It has the name 'grandfather'.')

(17) ta-ngena i-ma-ra'
1PL.INCL-wait 3SG-DUR-cooked
'We wait for it to be cooked through.'

(18) lalep=ra a-i-tu-patpat
house=3PL.POSS REAL-3SG-INCH-closed
'Their house is closed.'

The /p/-initial forms are also present when the agreement is suffixal rather than prefixal (e.g., 3SG suffix *-na* in (19)), which shows that this is a grammatical rather than morphophonological alternation.

(19) si Yosep m-asi-bukkup si Yudas, si Yudas
PERS.SG Yosep AV-TR-punch PERS.SG Yudas PERS.SG Yudas
pu-gerei-na-ngan
INC-scream-3SG-PRF
'Yosep hit Yudas, Yudas screamed.'

AV *m-* also interacts with MOOD: verbs which are host to *m-* are REALis (20-a), whereas those without are IRRealis by default (20-b).[13]

(20) a. sia m-asi-matei-ake' sikkoinan
3PL AV-TR-dead-CAUS crocodile
'They killed the crocodile.' REAL
b. ra-matei-ake' sikkoinan
3PL-dead-CAUS crocodile
'They will kill the crocodile.' IRR

3.4 Subordinate clauses

REL clauses are headed by the relativising proclitic *si=*. Verbs in REL clauses can take suffixal agreement (§3.2) and TAM (e.g., *CV-* PROG in (21)), but *not* AV *m-* (hence the ungrammaticality of (21-b)).

(21) a. ra-perep [$_{REL}$ si=pu-ga-gailak]
3PL-sleep REL=INC-PROG-lie.down
'They sleep lying down.'
b. *ra-perep si=m-u-ga-gailak

Where a REL clause is transitive, its SUBJ may be provided through agreement marking, and in this way the controlling GF in the matrix clause may be disambiguated - e.g., (22),

[13]The REALis/IRRealis distinction in SM is a fundamental tool of the grammatical system, and in many ways serves the role played by TENSE systems in Indo-European languages, but it should not be mistaken for a tense system in itself. We reflect REALis as English PAST or PRES in translation and IRR as either English FUT, since REALis (known) events are more likely to be PAST temporally, but this is an inherent limitation of translating from a REAL/IRR language to a tensed language such as English.

where the local pronominal SUBJ is specified as being 3PL by the agreement marker, so *tai kebbukatta* must be the SUBJ to *kau* 'give', since it matches these PERS/NUM values, rather than *nganga siburu'* 'the old language', which is SG.

(22) nganga si=buru' [$_{REL}$ si=kau-ra tai
language REL=old REL=give-3PL PERS.PL
kebbuk-at=ta]
older.sibling-NMLZ=1PL.INCL.POSS
'The old language that was given by our ancestors.'

However, if agreement marking is not present, as in (23) the argument of a matrix clause which controls into a REL clause is often ambiguous, if the semantics of the verb and its arguments do not provide a clue as to which argument is being controlled. Note that the absence of both AV *m-* and agreement, which gives rise to such ambiguities, is unique to subordinate clauses and would be ungrammatical in a main clause.

(23) nera sikkoinan [$_{REL}$ si=pasi-matei-ake' simanteu]
that crocodile REL=TR-dead-CAUS male
'That's the crocodile that killed the man.' OR 'That's the crocodile that the man killed.'

In XCOMP/XADJ clauses, the marking of the verb for voice is more variable, and depends on the argument structure of the verb in the matrix clause: if this does not obligatorily subcategorise for a controlled XCOMP SUBJ as with *oba'* 'want' (24-a), then AV *m-* is obligatory. However, if the matrix verb *does* obligatorily subcategorise for XCOMP SUBJ, as with (25) then the verb within XCOMP may or may not take AV *m-*, with no discernible change in meaning.

(24) a. ku-oba' [$_{XCOMP}$ m-asi-kau bulagat kai Tiur]
1-want AV-TR-give money DIR Tiur
'I want to give money to Tiur.'
b. *ku-oba' pasi-kau bulagat kai Tiur

(25) a. aku m-asi-guglu-ake' si Tiur [$_{XCOMP}$ [$_{SUBJ}$ _]
1SG AV-TR-command-CAUS PERS.SG Tiur
m-asi-kukru jo'jo' nera]
AV-TR-chase dog that
b. aku m-asi-guglu-ake' si Tiur [$_{XCOMP}$ [$_{SUBJ}$ _] pasi-kukru
1SG AV-TR-command-CAUS PERS.SG Tiur TR-chase
jo'jo' nera]
dog that
'I told Tiur to chase the dog.'

It is also worth touching on one structure which is *not* a true subordinate clause, but which resembles it superficially: anaphoric agreement which links to a prior IP. As implied by the name, sentences with anaphoric agreement can link anaphorically to prior sentences in a discourse where the same SUBJ has been referred to. But, essentially, since anaphoric agreement provides a local (pronominal) SUBJ, these are syntactically independent IPs. Thus, a similar meaning can be created by a matrix clause controlling

into an XCOMP (26) and two IPs where the second bears anaphoric agreement (26-a).

(26) a. a-m-ei aku ka pelabuhan [$_{XADJ}$ m-asi-gaba iba s=abeu]
PRF-AV-go 1SG LOC harbour AV-TR-look.for fish REL=large
b. [$_{IP}$ a-m-ei aku ka pelabuhan] [$_{IP}$ ku-gaba iba s=abeu]
PRF-AV-go 1SG LOC harbour 1-look.for fish REL=large
'I went to the harbour to look for big fish.'

3.5 Flexibility of constituent ordering

The ordering of constituents in SM is, particularly for a language of the Western Region of MP, remarkably free. Intransitives can freely appear in both SUBJ-initial and V-initial orderings (cf. (12) and (13)). Additionally, there is an auxiliary-first ordering, where the AUX *ai* inflects for aspect. In the AUX-first ordering, the SUBJ DP must precede the verb. This is true for transitive sentences as it is with intransitives; all verbs except unaccusatives must take AV *m-* rather than agreement marking.

(27) a. m-u-urai nia
AV-INC-sing 3SG
[VS]
b. ai nia m-u-urai
AUX 3SG AV-INC-sing
'She is singing.' [AUXSV]
c. ai-an nia m-u-urai
AUX-PRF 3SG AV-INC-sing
'She was singing.' [AUXSV]
d. *ai m-u-urai nia [*AUXVS]
e. ai nia m-asi-tut simakobu
AUX 3SG 3SG-TR-follow pig.tailed.langur
'He is following the pig-tailed langur.' (NOT 'The pig-tailed langur is following him.') [AUXSVO]
f. *ai m-asi-tut nia simakobu [*AUXVSO]

Transitive sentences without AUX *ai* where the main verb bears AV *m-* can be of the orders SVO or VSO. VSO orders are more common in contexts where the SUBJ has been established anaphorically and is thus low-prominence (backgrounded).

(28) a. aku m-asi-saki-ake' gette'
1SG AV-TR-buy-CAUS yam
[SVO]
b. m-asi-saki-ake' aku gette'
AV-TR-buy-CAUS 1SG yam
'I sell yams.' [VSO]

Transitive sentences where the verb bears agreement instead of AV *m-* are rather different in the orders they permit: SVO; VOS; and OVS are allowed, but not SOV; VSO; or OSV. Like with sentences where the verb bears AV *m-*, the SUBJ-initial order is most common when the SUBJ is topical or contrastive in the discourse, and the other orderings when it

is backgrounded.

(29) a. sikkoinan a-i-sotn-an aku
crocodile REAL-3SG-bite-PRF 1SG
[SVO]

b. a-i-sotn-an aku sikkoinan
REAL-3SG-bite-PRF 1SG crocodile
'A crocodile bit me.' [VOS]

c. aku a-i-sotn-an sikkoinan
1SG REAL-3SG-bite-PRF crocodile
'I was bitten by a crocodile.' [OVS]

d. *sikkoinan aku a-i-sotn-an [*SOV]

e. *a-i-sotn-an sikkoinan aku [*VSO]

f. *aku sikkoinan a-i-sotn-an [*OSV]

The grammaticality of both SVO and OVS orderings for sentences with agreement-marked (or more precisely, non-voice-marked) verbs creates an interesting situation when both SUBJ and OBJ DPs in such a sentence have PERS and NUM values which are compatible with those of the agreement morpheme: the same sentence can be interpreted as either SVO *or* OVS. Crucially, the OBJ-initial ordering is pragmatically marked, and may only be used when a) the SUBJ is well-established in discourse; and b) the OBJ is contrastive, or in some way unexpected/novel.

(30) *Yosep aikukru jo'jo' nera*

a. [SUBJ Yosep] a-i-kukru [OBJ jo'jo' nera]
Yosep REAL-3SG-chase dog that
'Yosep chased the dog.' (preferred) [SVO]

b. [OBJ Yosep] a-i-kukru [SUBJ jo'jo' nera]
Yosep REAL-3SG-chase dog that
'Yosep was chased by the dog.' (requires discourse context) [OVS]

If the SUBJ does not have the required PERS/NUM values to agree with the agreement marking, for instance if it is a question word, then only the OBJ-initial order is a licit interpretation.

(31) kasei a-i-kukru jo'jo' nera
who REAL-3SG-chase dog that
'Who was chased by the dog?' (NOT 'Who chased the dog?') [OVS]

This demonstrates that even in non-AV-bearing sentences where the 'undergoer' of the action appears pre-verbally, it remains the syntactic OBJ, since the agreement goes with the (post-verbal) SUBJ. Thus, SUBJ in main clauses (though not in subordinate clauses) is always coextensive with the 'actor' of the sentence, and OBJ with the 'undergoer'. The superficially 'passive-like' OVS ordering, although OBJ-initial and pragmatically marked, is not actually a passive or UV construction where the OBJ is actually ACTOR, but rather a fronting transformation more closely resembling that of Kambera (§2.4) and other MP languages where the voice morphology has eroded, either in whole or in part.

3.6 Summary of voice and agreement properties

This brief subsection summarises the functions of Actor Voice *m-* and its interactions with other elements of the morphology, particularly agreement, for reference in later theoretical sections. These are:

1. AV *m-* can never appear on the same verb as agreement morphology, but in main clauses either voice or agreement morphology *must* be present.
2. In REL clauses, AV *m- cannot* be present; relatedly, which argument is being controlled by the main clause is ambiguous when the REL clause is transitive and the SUBJ is not provided by agreement morphology.
3. In XCOMP clauses, if the main clause which provides the SUBJ does not obligatorily subcategorise for XCOMP SUBJ, then the verb of the XCOMP clause must have AV *m-*.
4. AV *m-* restricts possible constituent orderings to those where SUBJ precedes OBJ; agreement marking is more free through disallows VSO; V-final orderings are never permitted.
5. AV *m-* cannot appear on unaccusative or reciprocal verbs, but agreement morphology can.
6. All verbs which host AV *m-* have REAL mood.

4 LFG analysis

4.1 The verb template

As a preliminary to the following subsections, which discuss pronominal agreement and voice in the context of their position within the verb, it is pertinent to provide the templatic structure of the verb in SM, which is given in (32).

(32) V → (MOOD.PREF) + (AGR.PREF) + (VOICE.PREF) + AKT.PREF + (ASP.REDUP)
STEM + (AGR.SUFF) + (ASP.SUFF)

The slots STEM and AKT.PREF are the only slots which are used on all verbs.[14] Not all slots are can be used on the one verb, since some assign conflicting values to the same attribute. Namely, ASP.REDUP and ASP.SUFF both assign to the attribute ASP, and PRON.PREF and PRON.SUFF both assign to SUBJ. The slots MOOD.PREF, ASP.REDUP, and ASP.SUFF are all single-member slots which assign values to ASP or MOOD:

(33) a. *a-* MOOD.PREF (↑MOOD)=REAL
b. *CV-* ASP.REDUP (↑ASP)=PROG
c. *-an* ASP.SUFF (↑ASP)=PRF

[14]There are a few exceptions to the requirement for AKT.PREF, including (↑ADJUNCT TYPE)=REL and the aforementioned lexically-specified list of exceptions including *ei* 'go' and *oi* 'come'; these are not especially relevant to the points at hand and will not be modelled in detail here.

As is convention in LFG since Bresnan & Mchombo (1987), we can represent multi-functionality between grammatical and anaphoric agreement by adding the optional ((↓PRED)='PRO') to the lexical entries of agreement markers, which is mapped to SUBJ by (↑SUBJ)=↓: if a PRED value for SUBJ is provided by an external DP, then the pronominal PRED is simply not used. A sample lexical entry for 3SG agreement prefix *i-* is given in (34); others in the PRON.PREF slot are assumed to differ only in terms of their PERS and NUM values.

(34)	*i-*	PRON.PREF	(↑SUBJ)=↓ (↑PERS)=3 (↑NUM)=SG (↑ADJUNCT TYPE)≠REL	((↑PRED)='PRO')

Since the suffixal agreement set can appear in clauses which are either relative or have MOOD value PRF, members of this set can be modelled as requiring a disjunct between the ADJUNCT TYPE of REL and the morphological form indicating PRF at m-str (Butt et al. 1997; Findlay et al. 2023). The constraining equation (↑ADJUNCT TYPE)≠REL restricts *i-*, as a member of PRON.PREF, to appearing only in non-relative clauses; the distribution for members of PRON.SUFF is managed by the disjunct {(↑ADJUNCT TYPE)=REL|(↑μASP)=PRF} - i.e., either the clause must be relative, *or* the ASP value PRF must be projected to the m-str (Butt et al. 1996). A sample PRON.SUFF entry is given for *-na* in (35).

(35)	*-na*	PRON.SUFF	(↑SUBJ)=↓ (↑PERS)=3 (↑NUM)=SG {(↑ADJUNCT TYPE)=REL|(↑μASP)=PRF}	((↑PRED)='PRO')

The prohibition on the co-ocurrence of VOICE.PREF and either PRON.PREF or PRON.SUFF relates to the function of AV *m-*, the only member of the VOICE.PREF slot, and will be covered in §4.3.

4.2 C-structure

Key features of SM syntax as presented here can be described with the following Phrase Structure Rules. We posit S as a small clause unit that is used to deal with non-finite clauses.

(36)

CP → XP (↑DF-C)=↓ C'

C' → C IP

IP → XP (↑TOP)=↓ I'

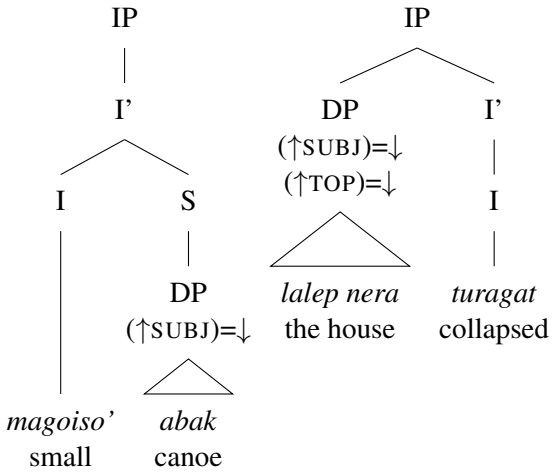
I' → I S

S → DP (↑GF)=↓ V DP (↑GF)=↓

These are capable of generating all the key constituent orders which we saw in §3.5. The

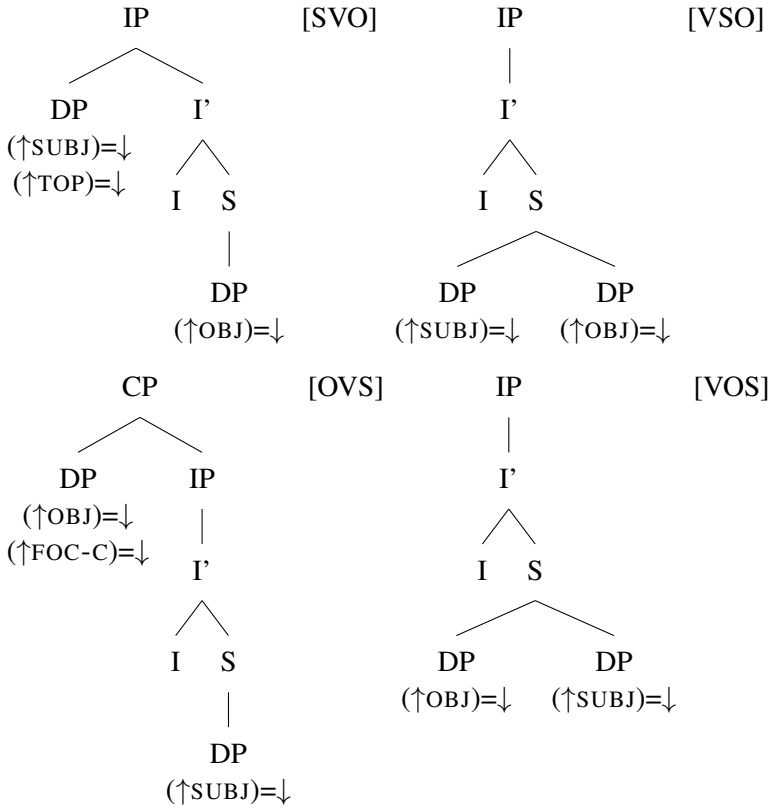
two possible orderings of intransitives are generated below (diagramming (12) and (13)) - in both cases, the single GF is assigned SUBJ; in the SUBJ-first ordering this receives default TOP as its DF; in the verb-first ordering it is backgrounded. We can understand this in information-structure terms (Arka & Sedeng 2018) as the SUBJ in Spec of IP being [+given], while the SUBJ that sits within S, which is not associated with a specific DF, is [-given].

(37)



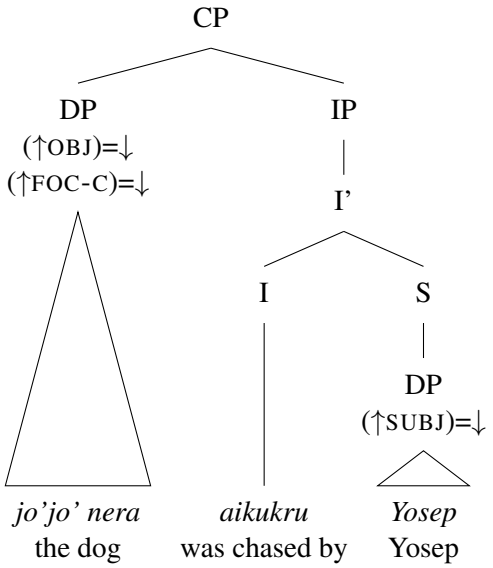
In transitive sentences, these rules generate all possible orderings other than those which are word-final (i.e., OSV and SOV), since I is always to the left of any DPs that sit within S. The selection of which of the generated orders (SVO, VSO, OSV, and VOS) is grammatical with which verb type (voice-marked or agreement-marked) is determined by the morphology (§4.3). Note that Spec position of IP carries with it the assignment of a contrastive Discourse Function DF-C (i.e., a DF which has the value [+contrast], following Arka 2021) - either contrastive topic or focus, depending on information-structure factors provided in discourse.

(38)



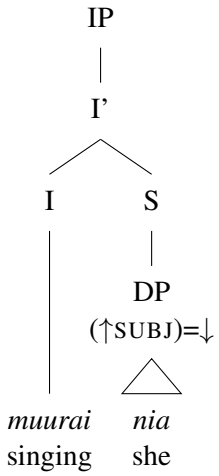
OVS is the only ordering to make use of Spec of CP, which assigns the relevant argument contrastive focus – this is the reason that the OBJ-fronted order is so pragmatically marked, as demonstrated by (30-b), diagrammed below. This also requires the backgrounding of SUBJ in the discourse – instead of being in a TOP-marked position (Spec of IP), it sits within S and receives no special DF.

(39)



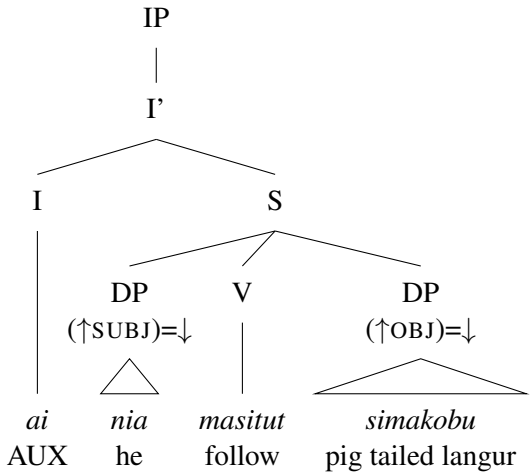
The head of IP can be either an AUX element I (below, diagramming (27-a)) or a lexical V (in (41)), since both share the same categorical features of [+predicative, +transitive] (Bresnan et al. 2015: 103). This allows us to generate both AUX-initial and V-initial orderings in intransitives, since both instantiate the same category I.

(40)



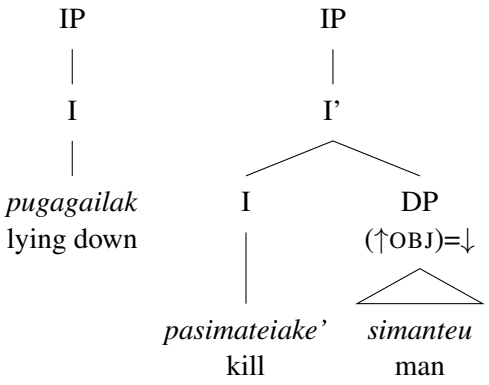
The structure of S also explains the prohibition on VSO ordering (where V is the lexical verb) in sentences which have an AUX verb: if the I slot is filled by AUX, then its sister S must not have two DPs following a verb, but rather place the verb in between the two DPs (this still allows in principle for OVS, which will be dealt with in §4.3). DPs within S are 'unmarked', in the sense that they are not associated with any particular DF, though DPs which are further left are still more prominent.

(41)



RELative clauses are also IP, as shown by their ability to take inflection (e.g., *pugagailak* 'lying down' below, from (21), which must be I since it takes (inflectional) aspect *CV-* PROG). Since every IP must contain a SUBJ, and since all transitive REL clauses obligatorily have one of their arguments controlled from the matrix clause, the postverbal DP in a REL clause is always a local OBJ - as below, with *simanteu* 'man', though note that its assignment to ACTOR or UNDERGOER is not clear, since neither the matrix clause nor the relative clause in (23), reproduced below, contains AV *m-*. which would clarify the mapping of OBJ *simanteu* 'man' to the ACTOR (§4.3).

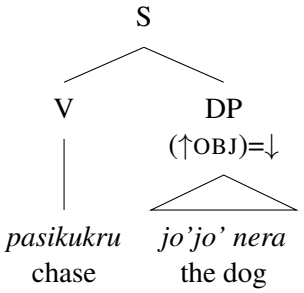
(42)



(43) nera sikkoinan [$_{REL}$ si=pasi-matei-ake' [$_{OBJ/A?U?}$ simanteu]
that crocodile REL=TR-dead-CAUS male
'That's the crocodile that killed the man.' (OBJ=U) OR
'That's the crocodile that the man killed.' (OBJ=A)

XCOMP clauses, meanwhile, are represented with unit S, since they always have an OBJ DP and are non-finite (contain a V, not I). Although the c-structure permits two DPs within S, the local SUBJ is always controlled (gapped) by an argument within the matrix clause, so the DP within XCOMP is always an OBJ – as below, diagramming (25-b).

(44)



4.3 The function of Actor Voice *m-*

This subsection will outline the behaviour of AV *m-*, the remnant of the voice alternation in SM. AV *m-* can be understood as a morpheme which maps SUBJ to ACTOR; requires linear precedence of SUBJ over OBJ; and assigns (↑MOOD)=REAL to any verb it appears on, overruling the default optional ((↑MOOD)=IRR). We can thus represent the lexical entry of *m-* as:

(45) *m-* VOICE.PREF $(\uparrow \text{SUBJ})\sigma = \uparrow\sigma \text{ACTOR}$
$(\uparrow \text{MOOD}) = \text{REAL}$

The first component of AV *m-*, which is not represented in the lexical entry, is a requirement for linear precedence and information-structural prominence of SUBJ>OBJ.

This requirement for SUBJ prominence is the reason that, out of all the possible orderings generated by the PSRs, only SVO and VSO are permitted when *m-* appears on the verb: these are the only orderings where SUBJ is more prominent than OBJ. This is true also in sentences with AUX *ai*, where OSV is disallowed (VSO and VOS being ruled out by the c-structure). The only ordering specifically prohibited for agreement-carrying verbs that is not ruled out by the c-structure (§4.2) is VSO; this can be interpreted as an information-structural dispreference for fronting SUBJ when it is not prominent.

AV *m-*	Agreement marking
SVO	SVO
VSO	OVS
	VOS

Table 5: Constituent orders permitted with particular verb morphology.

The inability for AV *m-* and agreement morphology to co-occur on the same verb is a result of this same entry: if the SUBJ is pronominal, and provided by the verb, it is not prominent in any information-structure sense; the requirement for linear precedence also stipulates that it must be a SUBJ DP which precedes OBJ.

This requirement for SUBJ prominence is also why AV *m-* cannot appear on unaccusative verbs: the SUBJ of these verbs is not prominent, and therefore cannot be 'elevated' on the prominence hierarchy. The same principle is responsible for the prohibition on AV *m-* on reciprocal/middle voice verbs: these are low in transitivity and do not have an information structurally-prominent SUBJ. Thus, SM is a 'split-intransitive' language, i.e., one that groups the SUBJ of intransitive verbs as sometimes more like transitive SUBJ and sometimes more like OBJ according to semantic properties, here unergativity/unaccusativity – similar to geographically nearby Acehnese (Durie 1987), but a historically independent development.

The second component, the identification of local SUBJ with macrosemantic ACTOR, explains the behaviour of bare verbs in subordinate clauses and the prohibition against co-occurence on the same stem as agreement morphology.

An XCOMP clause, like a REL clause, will always need one of its arguments to be controlled from a matrix clause. For some matrix clause verbs, such as *oba'* 'want', an XCOMP SUBJ argument will be non-core. In these cases, the XCOMP clause will require the explicit identification of a SUBJ. The entry $(\uparrow \text{SUBJ})\sigma = \uparrow\sigma \text{ACTOR}$ tells the XCOMP verb that the ACTOR found in its own argument structure is coreferential with some SUBJ, and thus the XCOMP that it needs to look for a SUBJ in the matrix clause. In cases where

the verb of a matrix clause, such as *gugluake'* 'command', imposes an obligatory XCOMP SUBJ, this will always be identified some argument in the matrix clause by the argument frame of the matrix clause verb (i.e., imposes (↑XCOMP SUBJ)=(↑SUBJ)). Hence, the assignment of the local SUBJ, i.e., XCOMP SUBJ, to ACTOR of the local verb is redundant. This explains why, with such verbs, AV *m-* can be freely deleted in the XCOMP clause with no change to meaning or ambiguity.

In REL clauses, AV *m-* is never present - this is possibly the result of the nominalising property of the Determiner *si=* REL making the contents of the REL clause opaque to aspects of the verb morphology. Without AV *m-*, however, there is no language-internal requirement in SM for SUBJ to correspond to either ACTOR or UNDERGOER. Although the c-structure will only ever provide a SUBJ as the DP within a REL clause (§4.2), the mapping to macrosemantic role is underspecified. This also further demonstrates that non-voice-marked voice are *not* carrying a UV morpheme that simply has a zero form, as in Balinese - if that were the case, the mapping would always be to UNDERGOER. In practice, the controller of SUBJ is often identified through agreement.

The prohibition against *neither* agreement morphology or AV *m-* occurring, at least aside from on unaccusative verbs, stems from the fact that these verbs demand an identification of SUBJ, either a local pronominal PRED or through the identification of SUBJ with a macrosemantic role, of which the only option is ACTOR.

The final component of the lexical entry, (↑MOOD)=REAL, is used to deal with the fact that verbs which bear *m-* are reliably REALis in mood. All verbal stems are assigned an optional (↑MOOD)=IRR, which is overruled by REAL where it appears.

5 Conclusion

This paper has provided an account of the semi-alternating voice system of Sipora Mentawai, an underdescribed language variety of Western Indonesia. It has outlined how, in place of a reduced symmetrical voice alternation between Actor and Undergoer voices, the relic Actor Voice prefix *m-* persists and assigns SUBJ to ACTOR macrosemantic role, but with no corresponding Undergoer Voice. Instead, a system of agreement morphology on the verb permits the fronting of OBJ through a 'passive-like' dislocation. This is demonstrated through agreement on the verb, which always identifies SUBJ as being ACTOR. The distinction between syntactic and macrosemantic roles language-internally is justified by a number of patterns involving subordinate clauses, where AV *m-* reliably associates SUBJ with ACTOR but its absence creates an ambiguity in the relation between syntactic and macrosemantic roles. This paper thus adds to the existing literature on voice systems of the Western Region of MP by showing how a voice system can be 'semi-alternating', preserving TAM and information-structure correlates of a former voice alternation while having lost its symmetricality.

References

Akoli, Marcelinus Y F, I Wayan Arka & Yuchen Li. 2025. Conservative but not frozen: Sigulai and the diverging Austronesian voice systems of the Barrier Islands languages. Australian Linguistics Society Conference, Gold Coast, Australia. https://www.indonesianlanguages.org/wp-content/uploads/2025/08/Sigulai-Abstract-for-the-ALS-finalversion.pdf.

Arka, I Wayan. 2003a. *Balinese Morphosyntax: A Lexical-Functional Approach* (Pacific Linguistics 547). Canberra: Australian National University.

Arka, I Wayan. 2003b. Voice systems in the Austronesian languages of Nusantara: Typology, symmetricality and Undergoer orientation. *Linguistik Indonesia* 21(1). 113–139.

Arka, I Wayan. 2017. The Core-Oblique distinction in some languages of Indonesia and beyond. *Linguistik Indonesia* 35(2). 101–144. https://doi.org/10.26499/li.v35i2.66.

Arka, I Wayan. 2021. Pivot and puzzling relativisation in Indonesian. In I Wayan Arka, Ash Asudeh & Tracy Holloway King (eds.), *Modular design of grammar: Linguistics on the edge*, 181–202. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780192844842.003.0012.

Arka, I Wayan. 2024. Attrition of voice symmetricality in the Austronesian voice system: Insights from Indonesia's Barrier Island languages and beyond. 15th International Conference of the Association for Linguistic Typology, Singapore. https://enggano.ling-phil.ox.ac.uk/static/papers/Arka-ALT2024%20final2.pdf.

Arka, I Wayan & Christopher Manning. 1998. Voice and grammatical relations in Indonesian: A new perspective. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'98 Conferences*, Stanford: CSLI Publications. https://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG3-1998/lfg98arkamanning.pdf.

Arka, I Wayan & I Nyoman Sedeng. 2018. Information structure in Sembiran Balinese. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages* (Studies in Diversity Linguistics 21), 139–175. Berlin: Language Sciences Press. https://doi.org/10.5281/zenodo.1402543.

Artawa, I Ketut. 1998. *Ergativity and Balinese syntax*, vol. 42-44 (NUSA Linguistic Studies of Indonesian and Other Languages of Indonesia). Jakarta: Badan Penyelenggara Seri NUSA. http://sealang.net/archives/nusa/pdf/nusa-v42.pdf.

Billings, Blaine & Bradley McDonnell. 2024. Sumatran. *Oceanic Linguistics* 63(1). 112–174. https://doi.org/10.1353/ol.2024.a928205.

Blust, Robert A. 2004. Austronesian nasal substitution: A survey. *Oceanic Linguistics* 43(1). 73–148. https://doi.org/10.1353/ol.2004.0004.

Blust, Robert A. 2013. *The Austronesian Languages* (Asia-Pacific Linguistics 8). Canberra: Australian National University.

Blust, Robert A. 2023. Repair and drift in Austronesian languages: Avoidance of dissimilar labials as the onsets of successive syllables. *Oceanic Linguistics* 62(2). 304–323. https://doi.org/10.1353/ol.2023.a913563.

Bresnan, Joan, Ash Asudeh, Ida Toivonen & Steven Wechsler. 2015. *Lexical-Functional Syntax* (2nd edn.). Oxford: Wiley-Blackwell. https://doi.org/10.1002/9781119105664.

Bresnan, Joan & Sam A Mchombo. 1987. Topic, pronoun, and agreement in Chicheŵa. *Language* 63(4). 741–782. https://doi.org/10.2307/415717.

Brown, Lea. 2001. *A Grammar of Nias Selatan*. PhD thesis, Unversity of Sydney.

Butt, Miriam, Mary Dalrymple & Anette Frank. 1997. An architecture for linking theory in LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'97 Conference*, Stanford: CSLI Publications. https://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG2-1997/lfg97butt-dalrymple-frank.pdf.

Butt, Miriam, María-Eugenia Niño & Frédérique Segond. 1996. Multilingual processing of auxiliaries within LFG. In Dafydd Gibbons (ed.), *Natural Language Processing and speech technology: Results of the 3rd KOVENS Conference, Bielefeld*, 111–122. Berlin: de Gruyter. https://doi.org/10.1515/9783110821895-013.

Chen, Victoria & Bradley McDonnell. 2019. Western Austronesian voice. *Annual Review of Linguistics* 5(18). 173–195. https://doi.org/10.1146/annurev-linguistics-011718-011731.

Durie, Mark. 1987. Grammatical relations in Acehnese. *Studies in Language* 11. 365–399. https://doi.org/10.1075/sl.11.2.05dur.

Findlay, Jamie Y, Roxanne Taylor & Anna Kibort. 2023. Argument structure and mapping theory. In Mary Dalrymple (ed.), *Handbook of LFG* (Empirically Oriented Theoretical Morphology and Syntax 13), 699–778. Berlin: Language Sciences Press. https://doi.org/10.5281/zenodo.10185966.

Foley, William A. 1998. Symmetrical voice systems and precategoriality in Philippine languages. Workshop on voice and grammatical functions in Austronesian languages, LFG98, Sydney.

Foley, William A. 2008. The place of Philippine languages in a typology of voice systems. In Peter K Austin & Simon Musgrave (eds.), *Voice and grammatical relations in Austronesian languages*, 22–44. Stanford: CSLI Publications.

Hemmings, Charlotte. To appear. Verbal morphosyntax in Enggano. In I Wayan Arka, Charlotte Hemmings & Mary Dalrymple (eds.), *Enggano: Historical and Contemporary Perspectives* (Asia-Pacific Linguistics), Canberra: Australian National University.

Hemmings, Charlotte & Mary Dalrymple. To appear. Relative clauses in Enggano. In I Wayan Arka, Charlotte Hemmings & Mary Dalrymple (eds.), *Enggano: Historical and Contemporary Perspectives* (Asia-Pacific Linguistics), Canberra: Australian National University.

Kaufman, Daniel. 2009a. Austronesian typology and the nominalist hypothesis. In Alexander Adelaar & Andrew Pawley (eds.), *Austronesian historical linguistics and culture history: A festschrift for Bob Blust* (Pacific Linguistics 601), 187–215. Canberra: Australian National University.

Kaufman, Daniel. 2009b. Inner causatives and the reconstruction of the Proto-Austronesian affixes *pa-, *pa<R> and *pa<N>-. International Conference on Austronesian Linguistics XI, Aussois, France. https://bahasawan.com/wp-content/uploads/2020/01/On-Austronesian-pa-paR-and-paN-H.pdf.

Klamer, Marian. 1996. Kambera has no passive. In Marian Klamer (ed.), *Voice in Austronesian*, vol. 39 (NUSA Linguistic Studies of Indonesian and Other Languages of Indonesia), 12–30. Jakarta: Badan Penyelenggara Seri NUSA. http://sealang.net/archives/nusa/pdf/nusa-v39-p12-30.pdf.

Mullan, Keira. 2025. Grammatical relations and person marking in Barrier Islands languages. Australian Linguistics Society Conference, Gold Coast, Australia. https://www.indonesianlanguages.org/wp-content/uploads/2025/08/Keira-ALS-Person-marking-in-Barrier-Islands-languages.pdf.

Riesberg, Sonja. 2014. *Symmetrical Voice and Linking in Western Austronesian Languages* (Pacific Linguistics 646). Berlin: de Gruyter Mouton. https://doi.org/10.1515/9781614518716.

Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of western Austronesian voice systems* (Pacific Linguistics 518), 17–62. Australian National University. https://doi.org/10.15144/PL-518.17.

Ross, Malcolm. 2006. Reconstructing the case-marking and personal pronoun systems of Proto-Austronesian. In Henry Y Chang, Lilian M Huang & Dah-an Ho (eds.), *Streams Converging Into an Ocean: Festschrift in Honor of Professor Jen-Kuei Li on his 70th Birthday* (Language and Linguistics Monograph Series W5), 521–563. Taipei: Academia Sinica.

Smith, Alexander D. 2025. Late Malayo-Polynesian: A new model of Austronesian linguistic relations. *Diachronica* 42(5). 646–698. https://doi.org/10.1075/dia.23062.smi.

Woollams, Geoff. 1996. *A Grammar of Karo Batak, Sumatra* (Pacific Linguistics C130). Canberra: Australian National University. https://doi.org/10.15144/PL-C130.