The Diachrony of Preverbal Subject Marking across *NTT*

Overview The presence of preverbal agreement and/or subject-marking is a well-known areal feature of South-East Indonesia (Klamer 2002), and is especially prevalent across the languages on and around Timor. This is exemplified in Uab Meto [Amarasi dialect] below:

(1) a. Ho **m**-kius kau
2SG.NOM 2SG-see 1SG.ACC
'You see me.'

b. Au **ʔu**-skora **ʔ**-bi Kopan.
1SG.NOM 1SG-school 1SG-at Kupang
'I go to school in Kupang.'

The diachronic origin of these preverbal elements has remained a long-running puzzle among scholars, most of whom have focused only on the Celebic and South Sulawesi languages (Wolff 1996; Zobel 2002; Billings & Kaufman 2004). This paper addresses an empirical gap by surveying the range of subject-marking systems found across *Nusa Tenggara Timur* (NTT), focusing on innovations within the Timoric, Flores-Lembata, and Sumba-Flores subgroups. From a theoretical perspective, this paper posits a diachronic pathway for these developments based on i) the cross-linguistically common reanalysis of Topics to Subjects causing the loss of VOS word order and Austronesian Focus; and ii) the role of agentive Serial Verb Constructions in facilitating grammaticalisation of the PMP genitive enclitics as preverbal agreement markers.

Cross-NTT Patterns An empirical survey of the strategies various NTT languages use to morphologically mark verbal arguments is schematised in (2), distinguishing between S/A/P and Possessor arguments as well as the S_{Agt} of an unergative and S_{Pat} of an unaccusative verb.

(2)

	Enclitic	Suffix	Proclitic	Prefix
PMP	Non-Pivot A, Poss	-	-	-
Manggarai	S/A, (3SG P), Poss	-	-	-
Komodo	S/A, Poss	-	-	-
Pulau Helong	-	S_{Agt}, Poss	-	-
Funai Helong	-	S_{Pat}, Poss	-	S_{Agt}/A
Dhao	P, Poss	(S_{Pat})	S/A	S_{Agt}/A
Wejewa	S_{Pat}/P, Poss	-	S_{Agt}/A	-
Idate	Poss	-	S/A	-
Uab Meto	P	Poss	-	S/A

In each language, the morphemes used to mark S/A arguments are clearly derived from the PMP genitive enclitics (3a), which originally introduced POSSESSORS and non-pivot AGENTS. Compare for instance the Wejewa paradigm in (3b) used to mark S_{Agt}/A arguments.

(3) a. Proto-MP Genitive Enclitics

	SG	PL	
1	*=ku	*=ta	(INCL.)
		*=mi	(EXCL.)
2	*=mu	*=(mu)-ihu	
3	*=ia	*=da	

b. Wejewa Proclitics (S_{Agt}/A)

	SG	PL	
1	ku=	ta=	(INCL.)
		ma=	(EXCL.)
2	mu=	mi=	
3	na=	a=	

One novel finding of this survey is that many languages have developed incipient Split-S or Nom-Acc alignment systems (contra the Erg-Abs systems of Sulawesi), and have reconfigured PMP pronominal morphology to these ends. Thus, Wejewa has retained **enclitic** pronouns for introducing the unaccusative S_{Pat} arguments of verbs like *etti* 'go' and *ammi* 'come':

(4) Yow'wa nda ammi =ki =**ngga** mana oro na mangeda pawili
1SG NEG come =NEG =1SG yesterday because DEM a.lot work
'I didn't come yesterday because I was busy.' (Kasni 2016: 105, ex. 14)

This survey not only adds to Klamer's (2008) typological review of Split-S systems in Indonesia, but also sheds light on the intermediate steps involved in the diachronic restructuring of these languages' voice and agreement systems. The following trends can be observed:

(5) a. Genitive enclitics first shifted from marking **non-pivot** A arguments to **all** A and S arguments, as in Komodo and Manggarai.

b. As Genitive markers shifted from post- to preverbal, languages first distinguished the marking of intransitive S_{Agt} arguments from S_{Pat} arguments, resulting in an incipient **Split-S** alignment system as in Helong, Dhao (affixes), and Wejewa. In languages which overtly distinguish the two, preverbal marking is used for S_{Agt}/A arguments, while postverbal marking is preserved for S_{Pat}/P arguments.

c. Languages like Idate, Dhao (clitics), and Uab Meto later developed **Nom-Acc** alignment, distinguishing all S & A arguments from transitive P arguments.

Stage 1. Losing Focus I argue that Komodo and Manggari lost Austronesian Focus when clause-final pivots were reanalysed as subjects, causing Object Focus constructions to be reinterpreted as (Nom-Acc) passives, i.e. a shift from PMP V=$_{\text{AGT}}$Pivot$_{\text{PAT}}$ $\rightarrow$ V=O$_{\text{AGT}}$S$_{\text{Pat}}$ order. Formally, this involved the innovation of VoiceP's ability to existentially-bind AGENTS (Legate 2014). Per Erlewine et al.'s (2020) analysis of genitive case on AGENTS like (3a) as Last Resort licensing for non-pivot arguments in Austronesian, I argue that these enclitics, already marked 'oblique' via genitive case, were reinterpreted as *by*-phrases reintroducing demoted AGENT arguments. These languages then lost the θ-role restrictions, resulting in active VOS order.

Stage 2. VOS to SVO. Languages in (2) then reanalysed the left-peripheral topic as a true subject, and the postverbal AGENT as an agreement marker, a well-documented cyclical change cross-linguistically (van Gelderen 2011). The shift from Topic VX=S to S VX=AGR produced languages like Komodo, Manggarai (6a), and Pulau Helong (6b) with postverbal marking:

(6) a. Aku cero latung =**k**
1SG fry corn =1SG
'I fry/am frying corn.'
(Arka & Kosmas 2002: 2, ex. 1a)

b. Oen tilu-**s** sii-**s** lako-**s** tasi-**s**
3PL three-3PL alone-3PL go-3PL fish-3PL
'Only the three alone went fishing.'
(Balle 2017: 16, ex. 10)

This shift is supported by Arka & Kosmas's (2002) observation that Topics in Manggarai can only co-index structural subjects; this obligatory co-reference would easily faciliate reanalysis.

Stage 3. Post- to Preverbal Innovation of preverbal $S_{(Agt)}$/A marking was then fed by Serial Verb Constructions headed by (aspectual) light verbs. Every language with preverbal marking in (2) also productively introduces SVCs with the descendant of PMP **lakaw* 'go', exemplified with Uab Meto (7a) below. Taking as archaic a system in which SVCs took an AGENT enclitic on only the **first** verb in a series, the resulting phonological string would have allowed prosodic reanalysis of this enclitic as **proclitic** on the second verb. This is schematised in (7b):

(7) a. Au **ʔ-nao** ʔ-meo ʔ-aan bi Uli
1SG 1SG-go 1SG-see 1SG-RES Ms. Uli
'I'll go and see Ms. Uli.'
(Edwards 2020: 281, ex. 204)

b. **Stage 1:** go=AGT V O (Pulau Helong)
Stage 2: go AGT=V O (Dhao clitics)
Stage 3a: (AGR)-go AGR-V O (Uab Meto)
Stage 3b: AGR=go V O (Wejewa)

All stages of the proposed shift are attested in NTT. This analysis is supported by i) the exceptional treatment of 'go' in Dhao (and Funai Helong) as the sole verb which is suffixed rather than prefixed (8); and ii) 'low agreement' in Uab Meto where prefixation skips over TAM auxiliaries (Lemon 2020), which also introduce SVCs and arguably used to take enclitics (and still do in Wejewa). This also accounts for Split-S in Helong and Wejewa: as 'go' preferentially combines with agentive/volitional verbs, reanalysis only occurred with S_{Agt}/A arguments.

(8) **La-mi** pare ku aj'u
go-2PL slaughter PRT wood
'You go to cut wood.'
(Balukh 2020: 92, ex. 81)

Selected References Arka & Kosmas (2002). Passive without passive morphology? Evidence from Manggarai. Balle (2017). Types of reduplication in Helong. Balukh (2020). A grammar of Dhao. Edwards (2020). Metathesis and unmetathesis in Amarasi. Erlewine et al. (2020). The typology of nominal licensing in Austronesian voice system languages. Kasni (2016). Klausa keterangan dalam Bahasa Sumba Dialek Waijewa. Klamer (2008). Split S in the Indonesian area.