A look at typological influences in the development of switch-reference systems in Highlands Papua New Guinea and Australia

Received: *November 28, 2023*

Revised: *December 16 2023*

Accepted: *December 18, 2023*

Jonathan Banks

University of Helsinki, Finland

jonathan.banks@helsinki.fi

Abstract

In this paper I describe the main characteristics of switch-reference systems of two regions, Highlands Papua New Guinea and Australia and show how certain underlying, general typological characteristics of languages of these regions inform and shape the nature of their switch-reference morpho-syntax and its paths of development through time. A core aspect of the differing characteristics of switch-reference systems of New Guinea and Australia is the predominating type of clause combining structure of each region, clause chaining in New Guinea vs clause embedding in Australia. I claim that this basic distinction is the basis through which the differing manifestations of switch-reference systems in each region develop diachronically.[1]

Keywords: Switch-reference, Typology, Diachrony, Highlands Papuan, Pama-Nyungan

[1] The following abbreviations are used: 1 – first person, 2 – second person, 3 – third person, ABS – absolutive, ALL – allative, ANTICSUBJ – anticipatory subject, COMPL – completive, CP – centripetal direction, DECL – declarative, DEP – dependent, DO – direct object, DS – different subject, ERG – ergative, FOC - focus, FPST – far past, GEN – genitive, IMPFV – imperfective, IRR – irrealis, LOC - locative, MAIN.O=DEP.S/A – main verb object is dependent verb subject, MED – medial, NEG – negation, NOM – nominative, NMLZ – nominalizer, NPST – non-past, O – patient-like argument of canonical transitive verb, PERSONAGREESUBJMEDCLAUSE – medial clause subject person agreement marker, PFV – perfective, PL – plural, PRS – present, PST – past, SEQ – sequential, SS – same subject, SUBJPERSONAGREE – subject person agreement marker

1. Introduction

1.1 Outline of the paper

This paper offers an account of the typological characteristic of switch-reference systems, henceforth SR, in languages of two regions, Highlands Papua New Guinea and Australia. The purpose of this paper is to show some specifics about SR systems of each region and the correlations of those characteristics of SR with the general typological characteristics of languages of the regions. These two regions were chosen as the areas of focus due to the fact that their SR systems differ remarkably from each other in a number of ways, both synchronically and diachronically. Examining this contrast leads to certain theoretical conclusions about the role of clause type and other typological features in the emergence of SR systems in general. The hypothesis I propose is that the form and function of SR as it exists in languages within a regional context is directly relatable to the general typological characteristics of the region in which it is found. In other words, certain typological features of the languages concerned shape the characteristics of the SR systems in those languages. This correlation between general typological features and SR specific characteristics has implications when considering the diachronic evolution of SR in individual languages, as well as the role of language contact in the diffusion of SR systems within regions. Specifically, I claim that the morpho-syntactic pathways of diachronic development are largely determined by the underlying typological characteristics specific to each region. For both Highlands Papua New Guinea and Australia, the common characteristics of SR in each region are discussed, as well as typological regional profiles. Though not all SR languages of any given region exhibit the same characteristics in their SR systems, this paper shows that, in a broad sense, a region based typological characterization of SR is possible due to the clustering of SR features according to regions.

The structure of this paper is as follows: Section 2 describes the data and methods used, Sections 3 and 4 present the details and theoretical conclusions for Papua New Guinea and Australia, respectively, with the conclusion in Section 5 offering a theoretical perspective recapping the findings and comparisons between the two regions, as well as directions for further research.

1.2 What is switch-reference?

The specific number of languages with SR around the world is yet to be determined, but according to my current estimate it appears that SR occurs in roughly 10% of the world's languages. According to Haiman and Munro (1983a), SR is canonically defined as a morpho-syntactic system of affixes, whose main function is to indicate whether or not the subject of a

dependent clause is the same participant as the subject of the main clause. However in a few languages, the continuity/discontinuity of non-subject participants, such as objects (van Gijn 2016b), and semantic agents and pragmatic topics are indicated as well (van Gijn ,2016a, Foley & Van Valin, 1984). In addition to indicating participant continuity, SR morphology is also often multi-functional, indicating a variety of inter-clausal semantic and pragmatic properties, such as adverbial dependencies, tense-aspect-modality values, event sequencing and participant topicality (van Gijn, 2016a). In addition to participant tracking, there is strong evidence (Mithun 1993) that in some languages the category ultimately being tracked across clauses is the continuity of main and dependent clause events, with event continuity often coinciding with continuity of the subjects. When occurring as affixes, SR marking almost exclusively occurs on dependent verbs as a system of suffixes or enclitics (Haiman & Munro, 1984), though instances of SR occurring as prefixes and particles also occur, with the latter particularly occurring in coordinated structures involving two independent clauses (McKenzie, 2015). The most common clause type associated with SR appears to be dependent adverbial clauses (van Gijn, 2016a), though complement, purposive, and medial/co-subordinate (dependent, non-embedded) are also relatively common, and in a few languages SR marking occurs on relative clauses as well.[2]

SR was first described in some detail by William Jacobsen (1967), who coined the term 'switch-reference' in consideration of the SR system of a language of Western North America, Washo. Since Washo appears to only have a single morpheme indicating a switch of subject between clauses (DS), in contrast to many languages where both continuing (SS) and switch of subject (DS) are marked morphologically, the term 'switch-reference' was first used relative to the system as found in Washo. The term 'switch-reference' was later applied to other languages where the same-subject category (SS) is indicated as well. The main comparative typological works on SR on a global scale are: the edited volume by Haiman and Munro (1983), where SR is considered mostly as a subject tracking device, Stirling (1993), which offers an account of switch reference as a phenomena primarily indicating the (dis)continuity of 'eventualities' across clauses and the edited volume by van Gijn (2016), in which the collection of papers offers a broad and state-of-the-art look at SR as it manifests in various languages around the

[2] A major clause type in Australian languages commonly marked for SR is the so-called adverbial/relative clause, in which the adverbial and relative functions are often indistinguishable. However, in other regions there are examples of SR appearing on dedicated relative clauses without an adverbial function. See van Gijn (2016b) for the language Yurakaré as an example.

world. In particular, the article contributed by van Gijn (2016a) in that volume offers perhaps the most comprehensive global overview of SR to date and includes such concepts as discourse, semantic and event continuity as possible categories that are tracked in some languages, not only the syntactic notion of subject. In addition to the above mentioned global cross-linguistic works, several typological studies with a regional focus have appeared in the literature as well, including Roberts (1997) for Papua New Guinea, Austin (1981) for Australia, van Gijn (2016b) for W. South America, McKenzie (2015) for North America and Treis (2012) for Ethiopia.

The main regions in which SR has been documented are certain language groups of New Guinea, Australia, North America and South America (van Gijn 2016a). More recently it has come to light that SR also exists in a few language groups of North-east Asia (Schmalz 2016, Matić 2016) and East Africa (Treis, 2012).

The following are examples of SR in which the (dis)continuity of subjects is tracked across main and dependent clauses. The participant category of subject appears to occur as the overwhelming majority of possible pivot categories in SR around the world (Haiman and Munro, 1983). As shown in (1), SR is coded on dependent clauses in Diyari, with the DS suffix disambiguating which participant is the subject of dependent clause in the second example (woman). In both examples, the TAM value 'imperfective' is included as part of the meaning-function of the suffix.

(1) **Diyari** (Pama-Nyungan; Australia; adapted from Austin 1981)

*karna wapa-yi, yathayatha-**rna***

man.ABS go-PRS speak-**SS.IMPRV**

'The man goes along talking'

*karna-li wilha nhayi-yi, kirli-**rnanhi***

man-ERG woman.ABS see-PRS dance-**DS.IMPRV**

'The man sees the woman dancing'

In Example (2), from Huichol, the SR suffixes apparently have no additional function apart from tracking the continuity/discontinuity of subject. They are also fully grammaticalized, such that SR marking appears along with the presence of independent noun and pronoun subjects, thus making the tracking function essentially redundant, though obligatorily expressed.

(2) **Huichol** (Uto-Aztecan; Central Mexico; Comrie 1983: 19)

*nee ne-nua-**ka**, paapaa ne pii ʔɨitɨ*

I 1SG-arrive-**SS**, tortilla 1SG 3SG give

'When I arrived, I gave him a tortilla.'

*ʔuuka nua-**ku**, nee ne-petia*
girl arrive-**DS**, I 1SG-leave
'When the girl arrived, I left.'

2. Data and Methods

The data for this paper comes from articles on individual languages with SR, general typological works in which SR is discussed, as well as region based surveys of SR where clusters of SR languages occur. Roberts (1997) provides by far the most comprehensive account of SR in New Guinea, in which he collected and analyzed data on 122 languages with SR systems, with the focus of his article specifically on languages of Papua New Guinea, henceforth PNG. Compared to PNG, somewhat less is known about SR or how widespread it is in the many languages of Papua Province, Indonesia, on the western side of the island. For this reason, the typological description of SR given here focuses on languages of PNG, with a specific focus on languages of the Highlands, since the diachronic development of SR in the Highlands can readily be seen from an examination of the synchronic data. According to Roberts, languages of PNG with SR stem from over 50 different families[3], however there are certain common general characteristics of SR marking and function in PNG, as well as general typological characteristics to be taken into account when considering the typology of SR of this region. Austin (1981) remains the most detailed typological work on SR in Australia, in which he examined numerous languages of the Pama-Nyungan branch, as well as a couple non-Pama-Nyungan languages, and detailed the distribution patterns of SR morphemes across the continent. These two papers remain to date the most comprehensive account of SR in these regions, respectively, and are a wealth of information on numerous SR languages and general characteristics of SR in these regions.

The main methodological approach used in this study is the correlation of regional SR characteristics with general regional typological characteristics. This is an important technique in identifying and tracing the diachronic development of SR, since the typological characteristics of these regions ultimately determine the pathways and processes of SR development.

[3] Many come from the so-called Trans-New Guinea Phylum, which is controversial as a valid genetic unit (see Foley 1986).

3. Highlands Papua New Guinea

According to Ethnologue online, out of a population of over 9 million, the total number of living languages for PNG is 840, stemming from many different families. Regarding the Highlands, broadly speaking, languages of this region display a strict SOV profile with complex verb stem morphology, consisting primarily as suffixes on a verbal root, often coding for subject agreement and in some languages object agreement as well. According to Foley (2017), in numerous languages of Highlands PNG, subject agreement suffixes are portmanteau morphemes which also indicate a range of TAM values as well. In dependent clauses, specific morphology associated with clause chaining, to be discussed below, occurs on the verb stem. In general, case marking on nouns is limited in comparison to other regions of the world such as Northern Eurasia and Australia (Foley, 2017), though in particular ergative marking does occur in numerous languages. According to Foley (2017), the expression of ergativity is complex in PNG and often is due to 'a number of semantic and pragmatic issues' such as agency, focus and topical status. It can even serve as a topicalizer in some languages (Potts & James, 1987, Donohue & Donohue, 1997).

3.1 Switch-reference in Papua New Guinea

SR is coded by a variety of morpho-syntactic patterns across languages of PNG, though most frequently in association with the so-called medial clause type as it occurs in clause chains. Clause chaining and medial clause types will be discussed in greater detail in Section 3.2, but as a basic description, clause chaining can be characterized as sequences of dependent, less-finite, non-embedded 'medial' clauses not functioning as an argument of a higher verb, and describe a sequence of actions, often in narration, culminating in a single, fully finite main clause at the end of the chain.

Example (3), from Watam, a non-SR language of Northern PNG, is given here as general illustration of clause chaining, where a sequence of events is recounted, with multiple non-finite, medial clauses in a chain, ending in a fully finite main clause.[4]

(3) **Watam** (Lower Sepik-Ramu; PNG; Foley 2017, p.921)

waut	*nakan*	*i*	*mbo*	*ga-**r***	*saŋga-**r***	*timoŋ*	*an*	*ŋg-utki-**r***
stone	big	a	LOC	climb-**DEP**	go-**DEP**	on.top	this	FOC-stand-**DEP**

*ak-**ri***

[4] It may be that in Watam the dependent morpheme -r has its origin as a reduced form of the past tense marker occurring on the final verb, underscoring the notion that in many cases medial verb morphology derives historically from reduced main verb morphology.

call out-**PST**

‘(He) climbed up a big rock and stood on top of it and called out.’

In those languages of PNG with SR systems, SR coding appears almost exclusively as suffixes on the verb stem of dependent medial clauses, often also indicating additional functions in addition to SR. The most commonly tracked (dis)continuity role, or SR pivot, is the participant category of subject. This appears to be the case in the vast majority of languages of PNG, though in some languages of PNG other SR pivot categories have been postulated, such as semantic agent or pragmatic topic (Roberts, 1997; Foley & Van Valin, 1984).

Example (4) from Gants, an SR language of the Highlands, illustrates a typical manifestation of SR in a clause chaining language of PNG. In this example, the continuity/discontinuity of the subject from one medial clause to the next is indicated. Though only the final clause in this example is fully finite, the DS marked clause does have a reduced form of person marking and overtly indicates a sequential value, which in this case apparently indicates a greater degree of temporal discontinuity between the DS clause and the final clause (’and then’), than among the SS clauses.

(4) **Gants** (South Adelbert; Papua New Guinea; Daniels 2015, p.1034)

gon	*tama-**da**,*	*aŋa-**da*** *ai-**da**,*	*wa*	*ga-**k-e***	*ma*	*ci-m-ek*
trap	put-**SS**	go-**SS** come-**SS**	go	perceive-**DS.SEQ-3SG**	NEG	stay-FPST-3SG

‘He set a trap, left, came back, checked, and (the animal) wasn’t there.’

As mentioned, in PNG SR markers are overwhelmingly suffixes appearing on medial verbs, very often occurring in clause chains. Suffix forms occur either as mono-exponential suffixes, indicating only the values of SS or DS, or as poly-exponential suffixes in which both SR and non-SR functions are indicated. It is important to note the additional functions associated with SR, since, they can be informative clues as to the historical development of these systems.

Some of the common additional functions coded along with SS and DS on medial verb suffixal forms in PNG are:

- person agreement, often manifesting as a reduced or even a unique paradigm of agreement in comparison to final, finite person agreement forms.

- Sequential vs simultaneous event distinctions[5]
- TAM distinctions, often with a reduced number of values in comparison to TAM values occurring on final, fully finite verbs.

Table 1 from Roberts (1997) shows the SR morphological coding strategies in PNG as a whole. These suffix combinations appear on medial verbs and form the morphological framework of SR in clause combining structures in PNG. As can be seen in the table, SR coding in PNG primarily consists of combinations of zero marking, invariable morphemes and person agreement. In a number of languages both invariable morphemes and person agreement may be poly-exponential and also indicate values such as TAM and sequencing, in addition to the SR values SS or DS.

Table 1 Coding strategies for switch-reference in Papua New Guinea according to Roberts (1997, p.136, with adaptation). The numbers of languages and example languages in his survey for each coding strategy type are given

1) SS = Ø, DS = invariable morpheme 18 languages - Menya, Binandere
2) SS = invariable morpheme, DS = Ø 1 language - Banaro
3) SS = invariable morpheme, DS = invariable morpheme 55 languages - Selepet, Koiari
4) SS = Ø, DS = SubjPersonAgree 10 languages - Ono, Kâte
5) SS = invariable morpheme, DS = SubjPersonAgree 36 languages - Angave, Rotokas
6) SS = SubjPersonAgree (paradigm i), DS = SubjPersonAgree (paradigm j) 6 languages - Nobonob, Kobon
7) SS = Ø + (SSAnticSubj)[6] 'bound pronominal reference of the following same subject'

[5] Sequential and simultaneous marking on medial verbs often occurs as morphology distinct from other TAM marking on the verb stem in PNG languages and is therefore considered separately here.

[6] Not all languages with anticipatory subject coding manifest the category SSAnticSubj, as some only code for DSAnticSubj

DS = (PersonAgreeSubjMedClause) + (invariable DS morpheme) + DSAnticSubj[7]
'bound pronominal reference of the following different subject'
17 languages - Fore, Gadsup

The person agreement strategies in Patterns 4), 5), and 6) above, which also code an SR value, are in some languages identical in form to final verb agreement, whereas in others it is a reduced form of person agreement or quite different altogether from the final agreement paradigm. Note that in pattern 6), SS is indicated by a specific subject agreement paradigm, whereas DS is indicated by a different agreement paradigm. Another pattern, schematized in 7) and which will be focused on in Section 3.3, is the so-called anticipatory subject agreement, occurring in a number of languages of Highlands PNG from different families. In this pattern, the medial clause is coded by means of an agreement morpheme, which codes the pronominal reference of the subject of following clause, NOT the medial clause itself to which it is suffixed. In the case of DS anticipatory agreement, person agreement with the subject of the medial clause, and/or an invariable DS morpheme also occurs in some languages with this pattern, though not all. As shown in Section 3.3, the language Kanite is an example of a language with SS and DS anticipatory agreement, as well as having an invariable DS morpheme and medial clause agreement in DS environments. The typological and diachronic significance of anticipatory marking is discussed in 3.3.

3.2 Clause chaining as major factor in the development of switch-reference in Papua New Guinea

From a diachronic perspective, clause chaining can be seen as a crucial conditioning factor in the development of SR systems in PNG, both in terms of the formal devices used to code SR, as well as the additional functional aspects of the morphology associated with SR coding, such as person agreement, TAM values and sequentiality. In fact, both the formal devices and the additional functions can be relatable to two major typological characteristics of these languages - strict SOV word order and clause chaining. Roberts' description of clause chaining is specific to PNG, but is applicable to other regions as well with chaining languages:

> "In PNG many Papuan languages have a particular feature known as clause chaining (from Longacre 1972) whereby many clauses can be linked together. The linking is normally indicated by distinctive verb morphology which differentiates dependent or

[7] Likewise not all languages manifest PersonAgreeSubjMedClause and/or an invariable DS morpheme, though all anticipatory subject languages do minimally code the category DSAnticSubj.

medial clauses, i.e. those clauses that occur within the chain, from independent or final clauses, i.e. those clauses that occur at the end of the chain. Typically the verbs in final clauses can be inflected for a fuller range of tense and mood categories than the verbs in medial clauses. Indeed in some languages the medial verbs (MV) may not be inflected at all for these categories. Also typically the medial verbs are inflected for SS/DS marking." (Roberts 1997: 104).

The following is a longer example of clause chaining from the SR language Chuave. The SR marking in Chuave is cataphoric (common for languages of PNG) in that it signals the (dis)continuity of the medial clause with the following clause, whether it is also a medial or final.

The SR coding type of Chuave appears to be a hybrid of types 3) and 5) from Table 1 above, since in the expression of DS both an invariable morpheme and person marking appear, whereas SS is coded only by an invariable morpheme.

(5) **Chuave** (Chimbu; PNG; Givon, 2001 from Thurman, 1978)

a. *meina i ne-**ro***
money get eat-**SS**
'(I) took the money …'

b. *ena tekoi u-**re***
then again come-**SS**
'… then (I) came back …'

c. *iki moi-i-**koro***
house be-1.SG-**DS**
'… and (I) stayed home …'

d. *tekoi u boi-n-**goro***
again call out-3.SG-**DS**
'… so then he sent for me again …'

e. *inako de-**ro***
return leave-**SS**
'… and so (I) came back …'

f. *fu-i-**goro***
go-1.SG-**DS**
'… and I went (there) …'

g. *tokoi numba lin-lin numba-i naro-Ø-m-e*
again number one-one number-that give-1SG.DO-PRV.3SG-DECL

'… and again he made me foreman (of the work-line).'

In the above example, only the DS clauses c, d and f and the final finite clause have person marking. Also, the form of the DS morpheme appears to incorporate the SS morpheme *-re/-ro*, with the additional element *-ko/-go*. According to Givon (2001, p.362), the SS morpheme is an erstwhile 3SG pronoun and the additional element in the DS morpheme derives from an erstwhile conjunction. The origins of these morphemes become especially significant in the context of proposals developed by Haiman (1983) and Givon (1983) as to the origin of SR systems, particularly in PNG, though their proposals are applicable to SR development in general.

Haiman (1983) proposes that switch reference reflects certain iconic principles and that these principles explain why DS is often the marked category in SR systems, whereas SS is often a zero. Haiman, points to the following quote from Bolinger (1975, p.136), "... *morphologicaldistance reflects conceptual distance*". This statement points to the concept that with a discontinuous subject across clauses (DS) there is a greater conceptual distance between the clauses than with continuing subject, since a complex clause with two subject participants is more likely to be construed as two separate events, or at least two less integrated events, than a complex clause with one and the same subject. Haiman further states that "*Gapping* (i.e. of same subject - JB) *clearly creates reduction, and reduction creates fusion. Fusion signals semantic fusion, while formal separation signals semantic separation"* (Haiman 1983, p.128)[8]. The above quote indicates that 'formal separation' can be seen as an iconic representation of the 'semantic separation' of events and participants, hence the appropriateness of DS marking in this regard, in contrast to zero marking for SS.

The following, adapted from Haiman (1983, p.126), indicates two of the possible origins of SR systems:

- Possibility 1. The index of SS is Ø (i.e. gapping/equi-deletion - JB) in contrast to person inflection for DS.[9]
- Possibility 2. The index of DS is a nominalizer or a conjunction, in contrast to Ø for SS.

[8] Haspelmath (2008c) proposes an alternative to iconicity as a motivation in the development of asymmetric morpho-syntactic systems, suggesting instead that frequency of occurrence is the central factor.

[9] In the case of Chuave, the index of SS is an erstwhile, grammaticalized 3SG pronoun, however this does not contradict Haiman's proposal, since this is actually a reduced pronominal form and appears to no longer refer to any argument pronominally, rather only indicates a continuing subject.

Haiman (1983, p.127 with adaptation) further schematizes this state of affairs in the following manner:

- Same Subject: Sentence1 Sentence2
- Different Subject: Sentence1 *X* Sentence2

and goes on to state that *"The interposition of any formal element X between coordinate clauses is sufficient to suggest a semantic disassociation between them…"* (Haiman 1983: 127)[10]

Haiman (1983) also suggests that in a number of Highland languages the origin of the DS invariable morpheme is a conjunction of the shape velar stop + vowel /KV/. As shown above, this conjunction appears to occur in the Highlands language Chuave in combination with the erstwhile 3SG pronoun to form the DS morpheme. It also occurs in the Highlands language Kanite, described in Section 3.3 below and may extend to SR languages outside of the Highlands area as well, such as some languages of Bougainville Island and other areas south of the Highlands. The morphological details of Chuave's SR system seem to correlate well with Haiman's proposal of iconicity, though in the case of Chuave at least, the SS morpheme is non-zero. It is however a smaller form in comparison to the DS form, both in actual size and composition, since the DS form has its origin in the combination of two separate morphemes, an erstwhile 3SG marker and a conjunction. Consideration of the size of a morphological form is compatible with Haiman's proposal and also figures well with the scenario Givon proposes for SR development.

Givon's proposal is more along the lines of explaining the diachronic encliticization of phonologically reduced pronouns, and connects the contrast between stressed and unstressed pronouns to the development of SS and DS markers on medial verbs in SOV languages.

He formulates his argument in following direct quote from Givon (1983, pp.77-78)

- "In all languages of type (anticipatory) the DS/SS system are strict SOV languages;
- In all those languages, the DS/SS morphology appears as *verb-suffixes,* i.e. also *clause final*

[10] Haiman formulates the above in terms of coordinate clauses, since clause chain structures are often translated as coordinate chains of clauses, though they are commonly still considered to be dependent due to their reduced finite characteristics.

- In all those languages only *medial, non-finite* clauses are involved in the SS/DS contrast
- Non-finite clauses are less likely to exhibit their own verb-inflections, such as tense-aspect-modality, mood, speech-act or pronominal agreement'
- In all these languages, the SS marker is consistently *smaller in size* than the DS marker. Often the SS marker is *zero*
- Such size-differential parallels closely the normal relation between stressed/unstressed or independent/zero pronouns
- We thus have full *functional* and *coding* parallels between the two systems" (i.e. pronouns and SR marking - JB)"

The significance of both proposals when taken together is what they have in common. Both refer to SS as a zero, or as in the case of Chuave a 'lighter' element, such as the reduced, erstwhile 3SG pronoun functioning synchronically as the SS marker. DS develops from some disjunctive, often phonologically 'heavier' element in between clauses. Independent pronouns represent a more disjunctive state of affairs than unstressed or zero pronouns, since they are more pronounced, both in terms of actual phonetics and in a pragmatic sense. They therefore iconically represent a greater conceptual distance than zero or unstressed pronouns. Likewise, a formal element such as a conjunction or disjunction is heavier than a zero element, and therefore is more likely to indicate a disjunctive state of affairs, such as DS. Since medial clauses in general have a tendency to undergo phonological and to some degree functional erosion (i.e. as in reduced finiteness of medial verbs), all elements following a medial clause, such as pronouns or conjunctions, over time and with frequent routinization, will have a tendency to become encliticized to preceding medial clauses. This follows from the strict SOV ordering of languages with extensive clause chaining. Though both Haiman's and Givon's proposals were made in the context of languages of PNG, the concepts they present are certainly applicable as well to SR development in languages of other regions apart from PNG and Australia.

When considering Haiman and Givon's theories of SR development through time, it can be seen that Chuave is a language which displays evidence of both proposals, in that both an unstressed pronoun and a conjunction are grammaticalized to indicate SS and DS,

respectively[11]**.** As will be shown in the next section, Kanite, another language of the Highlands, though from a different family, also shows evidence of both sources of SR marking. Kanite is an even more interesting case to consider and further validates both Haiman's and Givon's proposals and the notion that medial clauses in strict SOV languages can encliticize and incorporate even heavy elements, such as independent pronouns, onto preceding medial clause verb stems.

3.3 Anticipatory subjects in Highlands Papua New Guinea

As shown in Table 1 above, SR coding strategy 7 is the so-called anticipatory subject marking strategy. This is not to be confused with 'forward looking' cataphoric SR marking (sometimes also termed anticipatory subject marking in the literature), as in the Huichol example (2) above and Chuave (3) as well. In forward looking SR, a dependent clause marked with either SS or DS precedes a main clause and references the continuity of the subject across both clauses. With an anticipatory subject marking, however, a bound pronominal agreement suffix appears on the medial clause verb stem which actually indexes the subject of the following clause, not the subject of the medial clause itself, thereby coding the (dis)continuity of the subject. This is a highly unusual morpho-syntactic structure and appears to only exist in languages of Highlands PNG, though its existence in other regions would not be entirely surprising, since its diachronic origin is in fact relatively transparent and stems from clause chaining in strict SOV languages. This strategy is reproduced here from Table 1, and comes from Roberts (1997).

- SS = Ø + SSAnticSubj
- DS = PersonAgreeSubjMedClause + invariable morpheme + DSAnticSubj

It should be noted that not all languages have all the above morphology, though SS = Ø and DS = DSAnticipatorySubject occur in all languages with this type of SR anticipatory subject marking.

An example of a language with anticipatory SR coding with all of the above morphological marking is given by the Kanite example in (6) below.

(6) **Kanite** (Gorokan; PNG; Givon 2001: 359-360, from Longacre 1985, with additions for clarification)

a. *his-**u'a-ke-'ka***

[11] Technically, since the DS marker incorporates the SS marker in Chuave, it could be said that only DS is marked, though synchronically it likely is the case that the whole complex *-goro/-koro* functions as DS, not simply the first syllables.

do-**1.PL.MED-DS-2.PL.ANTICSUBJ**

'…If we do this, …'

DS is marked by medial clause person marking, a separate DS marker and anticipatory agreement. Next clause DS = 'you pl'

b. *naki a'nemo-**ka** hoya ali-**'ka***

so women-**2.PL** garden work-**2.PL.ANTICSUBJ**

'… you women work the garden …'

SS is marked by anticipatory marking only. Next clause SS = 'you pl'

c. *naki ali ha'noma hu-ne'atale-**'ka***

so work finish do-COMPL-**2.PL.ANTICSUBJ**

'… then you finish …'

Next clause SS = 'you pl'

d. *inuna kae-**'ka***

weeds burn-**2PL.ANTICSUBJ**

'… and burn the weeds …'

Next clause SS = 'you pl'

e. *popo hu-**'ka***

hoe do-**2.PL.ANTICSUBJ**

'… and hoe …'

Next clause SS = 'you pl'

f. *naki ha'no hu-talete-**ke-ta'a***

so finish do-COMPL-**DS-1.PL.ANTICSUBJ**

'… and when you finish …'

Next clause DS, final finite clause in chain, subject 'we'

g. *naki viemoka-**ta'a** keki'yamo'ma ha'noma nehis-i-ana*

so men-**1.PL** fence finish make-3.SG.DO-IRR

'… then we men will finish making the fence …'

In the above example, clause (6a) is a DS clause in which the subject of the medial verb 'we' is coded on the dependent verb stem, as well as the invariant DS marker '-*ke*' and the anticipatory suffixes referencing the subject (DS) of the following clause. Clauses (6c)-(6e) are all continuing subject clauses (SS) and the only SR related morphology is the anticipatory subject suffix indicating the continuing subject of the following clause. Clause (6f) is a DS clause, and also has the invariant DS morpheme as well as the DS anticipatory subject suffix. However, in contrast to the DS clause in (6a), the subject of medial clause (6f) itself is not

marked on the verb stem. According to Givon (2001, p.360), the subject of the DS medial clause in (6f) is not coded in this case because it is not signaling as strong of a thematic break as is indicated by the initial DS clause in (6a), therefore medial verb agreement in (6f) is superfluous. The point of this observation is that in chaining languages thematic boundaries, often DS for pragmatic reasons, tend to occur as initial and final clauses, whereas chain medial clauses, such as (6f), have a lesser degree of thematic discontinuity, even though they may have discontinuous subjects.

The interesting observation from the above example is that the form of the 1PL anticipatory subject in (6a)-(6e) *-ka* is very similar to what appears to be a pronominal ending in clause (6b) on the NP 'you women'. This also holds for the 1PL anticipatory subject suffix in clause (6f) *-ta'a*, which has the same form of the pronominal suffix on the NP in (6g), 'we men'. This seems to be a clear indication that the anticipatory subjects have their origins in pronominal forms. Also, the form of the invariant DS suffix in clauses (6a) and (6f) is *-ke*, which is in line with Haiman's proposal that in a number of SR languages of PNG, the DS morpheme originated in a conjunction of the shape 'kV'. Both Chuave and Kanite thus provide evidence that validates both Givon's and Haiman's claims on the origins of some forms of SR marking in PNG, namely reduced, encliticized pronominal forms and encliticized conjunctions, respectively. Moreover, it appears from this example that the anticipatory enclitics in Kanite may have their origins in heavier, perhaps even independent pronouns, since they appear to occur as noun suffixes as well, though this remains to be determined from the available data on the language. Of course not all languages of the Highlands show evidence in favor of both proposals simultaneously as do Chuave and Kanite, though many have traces of one or the other paths of development of SR morphology. It should be mentioned also that Givon and Haiman's proposals are likely not the only sources of SR morphology in PNG, though it appears that they can both be considered as major contributors to SR morphological development in this region, and are very likely to be applicable to SR development in other areas of the globe as well.

3.4 Typological features and diachronic development of switch-reference in Papua New Guinea

The Kanite and Chuave examples above appear to validate both Haiman's and Givon's claims on the origins of major aspects of the morphology associated with SR systems in PNG. However in the greater typological context their proposals are also significant in that the formal aspects of SR systems in PNG can be directly related to the typological properties of the languages in question. Specifically, SR in PNG appears to be strongly correlated with rather strict SOV word order and clause chaining. Givon (1983) indicates that SOV word order is

clearly a contributing factor in the morphological development of SR in these languages, and that SOV orientation also explains the overwhelming occurrence of SR morphemes as suffixes in PNG. "*This morphological contrast — either between stressed vs, unstressed pronouns or pronouns vs. zero, respectively, merely became cliticized on the preceding verb, given the strict SOV typology of these languages*" (Givon 1983, p.78). It is certainly the case that, cross-linguistically, SR morphology has a strong tendency to occur as suffixes, though not exclusively, as some SR morphology does occur as particles or even prefixes in some languages[12]. However it is specifically suffixes on medial-type verb stems in the strongly SOV languages of PNG that is of typological interest, since in many cases, as described above, the diachronic origins of these suffixes can be seen as stemming from the encliticization of inter-clausal material, such as pronouns or conjunctions, occurring in medial clause chains.

Regarding the development of anticipatory subject marking in languages such as Kanite, Givon (2001, p.360) characterizes the process of encliticization in SOV chaining languages with anticipatory subject marking as stemming from L-dislocation structures at thematic boundaries, often correlating with DS clauses, stating that "*The use of both devices (*i.e. topic-switching independent pronouns, and L-dislocation structures*) for switch reference is wide-spread cross-linguistically*" (Givon 2001, p.360). This correlation between topic/subject switching and thematic boundaries sets the stage for the generally more marked SR category, DS, to be re-analyzed from available morphological material, such as pronouns and conjunctions occurring directly after the clause representing the thematic boundary. Though SS anticipatory agreement in languages which express it appears to also have pronominal origins, it is the case that a number of languages with anticipatory agreement only express DS anticipatory agreement on medial verbs, with zero marking for SS, further indicating the more marked nature of DS coding.

In consideration of the co-occurrence of clause chaining and SR in New Guinea in general, Foley (2017, p.922) says that SR is "*missing from the left-headed West Papuan and Torricelli languages, because they completely lack clause chaining, a necessary prerequisite for switch reference."* However, in contrast to this assertion, it seems clearly to be the case that SR does exist elsewhere in languages without a strong clause chaining propensity, as in Australia, for example, discussed in Section 4. Nevertheless, I suggest that clause chaining is

[12] Lenakel, an Austronesian language of Vanuatu, is an example of a language with SR prefixes (de Sousa 2008), whereas Kiowa, a Kiowa-Tanoan language of the U.S.A, appears to be an example of a language which indicates SR by inter-clausal particles (McKenzie 2015).

possibly a natural clause combining structural development in strict SOV languages, which in turn has a natural tendency for the phonological erosion of material in between and at the boundaries of clauses to occur, leading to the encliticization of material onto preceding clauses, such as independent or unstressed pronouns, conjunctions, nominalizers and grammaticalized lexical items, all potential sources of SR marking. As for the underlying motivation for the development of reference tracking in SOV chaining languages, I suggest that SR is a natural development in languages with extensive clause chaining, for the reason that subject or topic reference tracking in long chains of clauses is an important aspect of discourse cohesion. Evidence in support of this is, as mentioned above, that thematic breaks and DS clauses tend to coincide, and therefore referential tracking at thematic break junctures becomes important for discourse cohesion, because it is here that changes of subject/topic are likely to occur.

Foley (2017, p.901) also mentions that *"Case systems are usually not very complex in Papuan languages, in comparison, say, to Eurasian or Australian languages."* Topic marking, particularly associated with pragmatic 'optional' ergative marking, appears to be more prevalent in PNG than core case marking in general. A consequence of this relative to clause chains is that topic tracking appears to be a significant feature of these languages, as it tends to be overtly coded. This may be evidence that languages of this type, specifically SOV clause chaining languages, have the need to identify and track topics for discourse pragmatic reasons and therefore may be pragmatically motivated to develop SR systems.

Regarding distinctive person paradigms indicating SR marking on medial clauses in some languages, as medial clause structures develop, pronominal material in between clauses becomes reduced, eroded and encliticized, resulting in remnants of person marking patterns on medial verb stems in some languages, such as the non-referring erstwhile 3SG pronoun of Chuave, but full paradigms and/or referring bound pronouns in others, such as in Kanite. This may be a reflection of the length of time involved in the development of the system.

To summarize, the main typological pressures mentioned here leading to the development of SR in languages of PNG with clause chaining are:

- Clause chaining involves essentially flat, semantically coordinate structures, though syntactically dependent on a main clause. A consequence of this flat structure, particularly in long chains of clauses narrating a sequence of events, is the need for thematic boundary distinctions, as implicitly suggested by Givon (2001). This in turn naturally correlates with change of topic/subject, since a thematic change often correlates with a change of topic/subject.

- Given the need to track participants at thematic boundaries, where topic changes occur, strictly SOV clause chaining languages may have a natural tendency to develop topic/subject tracking suffixes on medial clauses, particularly at thematic junctions, often involving erstwhile L-dislocation constructions.
- Topic marking may be more prevalent in PNG than core case marking, a further indication of the pragmatic pressures in these languages to identify and track topical participants in long clause chains.
- As a consequence of the natural tendency of medial clause material to become eroded through frequency and routinization, inter-clausal pronominal forms erode and encliticize onto the preceding verb stem, leading to the establishment of specific, often reduced, paradigms associated with SR, in contrast to full agreement paradigms of fully finite verbs.

4. Australia

According to Ethnologue online, out of a population of 25.5 million, the total number of living indigenous languages for Australia is 214, many of which are endangered to severely endangered, with 190 indigenous languages already having become extinct. Almost all SR languages in Australia occur in one large sub-branch, Pama-Nyungan, henceforth PN, with 2, possibly 3 exceptions in related non-PN languages that are located adjacent to PN languages, strongly suggesting transfer through contact (see Banks 2023a). The typological profile of PN languages is one of variable, pragmatically driven word order, with many languages having extensive case marking concord on elements in noun and verb phrases. Morphologically, split ergative alignment patterns appear in most Pama-Nyungan languages, with the exception of several languages in the North-West of the continent, which have predominantly nominative/accusative morphological alignment. According to Dixon (2002), the common pattern is for pronouns (both free and bound on the verb stem) to occur with nominative/accusative alignment, or in some cases, a three-way case marking distinction, and for full nouns to have ergative/absolutive alignment. Dixon (2002) also states that verbs are strictly sub-categorized as transitive or intransitive, though various derivational processes occur which alter the transitivity of a verb root. Verb roots are also typically inflected by suffixes indicating TAM distinctions.

4.1 Switch-reference in Australia

The most comprehensive look at SR in Australia is Austin (1981), who shows that specifically locative and allative case marking, or forms derived from them, occur in regional

patterns among SR languages in Australia. Moreover, SR appears primarily on two types of dependent clauses in Australia, adverbial/relative clauses (grouped together, since in many cases, though not all, adverbial functions and relative functions are indistinguishable in these languages) and purposive clauses. Overall there is more similarity among SR systems in Australia than in PNG, in that in Australia the typical pattern is for nominalized clauses to be coded by SR markers often formally very similar or identical to oblique case forms on nominals. Also, as in PNG, certain additional semantic and functional values are often indicated as well by SR markers, such as aspectual and sequentially distinctions, however overall a greater degree of morphological, semantic and functional variation is present in the SR systems of languages of Papua New Guinea than Australia. This is however not surprising, since it may be that all the languages of Australia are ultimately related (Alpher 2004, pp.93-94), though with a deep division between Pama-Nyungan and other branches, whereas even assuming that the Trans-New Guinea Phylum hypothesis is a valid genetic unit, there still remain a number of likely separate genetic groups in Papua New Guinea. This should not, however, be taken to imply that SR in the PN languages of Australia is necessarily an inherited phenomenon from Proto-PN, since SR is more likely to have spread through language contact in Australia (Austin, 1981; Dixon, 2002; Banks. 2023a). Rather, due to the closer typological similarity of these related languages, the manifestations of SR appear with more limited variation in Australia than in Papua New Guinea.

As in PNG, the grammatical category that is most consistently the pivot, or tracked argument, in Australian languages is subject. However 'object SR', where co-reference between subjects and objects are tracked across clauses, is also present in certain languages in Australia, specifically in the North-central area of the range of SR languages. An important difference between SR in Australia and PNG is that in Australia SR occurs mainly on embedded dependent clauses, such as adverbial/relative and purposive clauses, functioning as core or adjunct arguments of a main clause verb, in contrast to the non-embedded, dependent chaining of PNG described in Section 3.1. Chaining has also been described as syntactic clause combining type for some languages of Australia, such as Mpwente Arrente (Wilkins 1998), though it possibly is restricted to only certain languages, with limited or no interaction with SR systems. For instance Wilkins (1988) describes chaining structures in Mpwente as not interacting with the SR system of the language, since the chaining structures in that language only appear in continuing subject (SS) contexts, therefore the possibility of DS chaining, as in

PNG, does not occur.[13] Whether or not chaining structures as dependent, non-embedded clause strings comparable to the ubiquitous clause chains of PNG occur in Australia remains to be further investigated, though it seems clear that it is not nearly as common in Australia.

A crucial aspect of SR formally in Australia is that the overwhelming majority languages have SR markers that are similar to locative, allative and other oblique case marking on nouns as shown in Austin (1981), obviously suggesting a diachronic connection. In fact, in a number of languages, SR making is formally identical to oblique case marking, as exemplified by the Warlpiri examples below from Simpson (1988). Note that Warlpiri is one of the few languages so far identified in Australia in which the SR system indicates not only subject (dis)continuity across clauses, but also object to subject continuity across main and dependent clauses, as shown below. Other languages of Australia, such as Martuthunira and Yinggarda, appear to have object SR as well, or at least the beginning stages of developing it.[14]

(7) **Warlpiri** (Pama-Nyungan; Australia; adapted from Simpson 1988)

a. *yapa* *ka-lu* *pina-rni* *ya-ni* ***ngurra-kurra***
person PRS-3.PL back-CP go-NPST camp-ALL
'The people come back home.'

b. *ngarrka-ngku* *marlu* *pantu-rnu* ***marna*** ***nga-rninja-kurra***
man-ERG kangaroo spear-PST grass eat-NMLZ-MAIN.O=DEP.S/A
'The man speared the kangaroo while it was eating grass.'

In (7a) *-kurra* is the allative case, indicating movement towards or into a location. In (7b), the same suffix appears on the dependent clause (nominalized by the suffix *-rninja***)** and indicates that the object of the main clause 'kangaroo' is the same participant as the subject (S/A) of the dependent clause. Similar examples of case and SR marker correspondences can be found in numerous languages of Australia. Banks (2023a) provides a look at some of these correspondences in a selection of PN and 2 non-PN languages with SR. Out of 18 languages looked at, a total of 16 languages had identical or nearly identical forms for SR and oblique

[13] The actual subordinating form of the same-subject only chains in Mpwente Arrente is identical to SS marking in non-chaining clause combinations, such as adverbial dependencies, indicating a functional and historical connection. Nevertheless, the chaining structures of Mpwente diverge from the SR chains in PNG in that only same-subject chains occur.

[14] Object SR is highly elaborate in some languages of Western S. America (Overall & Vuillermet 2015), indicating various combinations of continuity tracking between subjects, objects and oblique participants across main and dependent clauses. Object SR possibly also exists in some Uto-Aztecan languages of N. America as well, though to a much more simplified degree than in the W. Amazon region.

cases, mostly locative and allative cases, but dative and ergative (often identical to instrumental) was also identified in a couple of languages.[15] As mentioned, this state of affairs in Australian languages points to a diachronic connection between case marking and SR marking, with the assumption that SR morphology developed from re-analyzed case marking, as proposed in Austin (1981) and Bickel (1999).

4.2 Case marking origins of switch-reference

When considering the diachronic connection between nominal oblique case forms and SR marking in Australia, an intriguing proposal for a diachronic scenario of development is given by Bickel (1999). In his paper, Bickel proposes that oblique case marked, adverbial dependent clauses, using the term 'absolute constructions', may be a common source of the development of SR systems cross-linguistically. The scenario he outlines is that in languages which mark nominalized dependent clauses with oblique case marking, there is a pragmatic tendency for these 'absolute' oblique case marked clause constructions to occur where the subject of the dependent clause and the main clause are different, since absolute constructions in fact function as a discourse topicalization device, likely occurring at thematic boundaries in discourse. I suggest that this pragmatic tendency can be connected back to Haiman's iconicity proposal outlined in Section 3.2, specifically that non-core cases conceptually represent a greater degree of discontinuity in the overall nature of an event, in that participants in oblique functions are conceptually on the periphery of a predication, as compared to the core functions of subject and object, which are more central and integrated into the unfolding action of a predication. According to Bickel, as this pragmatic tendency for oblique case marking of dependent clauses with a discontinuous subject becomes more frequent and routinized, there is a further tendency for the oblique case marking to become re-analyzed as coding a change of subject, i.e. DS. This would provide an explanation, in addition to Haiman's iconicity proposal above, for the well-known fact that in many SR systems around the world the DS category is more marked than SS in the overall SR system. He also suggests that this in turn sets up further pragmatic conditions for a contrasting set of SS markers to grammaticalize from available morphological resources, which in the case of Australian languages is a separate oblique marker. According to Bickel a similar process occurred in Latin and Ancient Greek, with a tendency for oblique marked absolute constructions to occur in DS environments and nominative marking to occur in SS environments on dependent participles (nominalizations),

[15] Baagandji had an SS form of the same shape as a locative case suffix in a neighboring language, but from the source (Hercus 1982) it did not appear to have a similar oblique case in the language itself.

indicating that those languages drew up on different morphological resources as its source of (emergent) SS marking than the grammaticalization of various oblique cases as both DS and SS in Australian languages.[16] Austin (1981) offers an explanation why certain oblique case markers, such as locative or allative, have tendency to grammaticalize as SS or DS, respectively, in Australia. *"In transitive clauses, a locative NP can refer to the location of the subject and/or object NP; thus a sentence which translates 'The man shot the bird tree-Loc' can mean that the man and/or the bird were in the tree. However, allative-marked NP's can also be used in transitive non-motion clauses, but they refer only to the location of the OBJECT. Thus 'The man shot the bird tree-ALL' can only mean that the bird was in the tree and the man was not. If the case markers came to be used with non-finite (nominalized) verbs in sentences like 'The man shot the bird, singing-LOC/ALL', and if locative narrowed to refer only to subject, then this relative locational system could have developed into the present-day switch-reference systems."*[17](Austin 1981, p.332). However this must be viewed as only a tendency for the pathway of grammaticalization of obliques into SR values in Australia, since not all languages grammaticalize the same oblique markers in the same SR function. As Austin mentions, his proposal only applies to certain northern SR languages of Australia, since in many southern languages the locative actually codes DS on nominalizations. Further research is required to more fully determine the choice of oblique markers for grammaticalizations into either SS or DS function in Australian languages.

Another important consideration in the development of SR marking from absolute constructions is the occurrence of case concord on elements in the nominalized clause, much in the same way that in some languages adjectives in noun phrases agree with the head noun in terms of case and other coding features (such as gender). The example below from Panyjima shows accusative case concord in the dependent clause.[18] Note that in this example, all elements of the dependent clause are case marked as accusative.

(8) **Panyjima** (Australia; Pama-Nyungan; adapted from Dench, 1991, with a possible alternate translation)

[16] The reasons for this discrepancy in the grammaticalization of SS marking between Australian and Indo-European languages is beyond the scope of this paper, but it should be stated that both Latin and Ancient Greek likely only had emerging SR systems, which were not fully grammaticalized as coding SS and DS.

[17] A possible underlying motivation for the narrowing of the locative to refer only to subject was not touched upon in Austin's paper and remains an interesting avenue of future research.

[18] Panyjima is an example of an Australian language with nominative/accusative core case marking on full nouns, however by far vast majority of Australian languages have ergative/absolutive alignment on full nouns.

ngatha	*mirta*	*mirnu*	*yukurru-ku*	*ngarna-lha-ku*	*nyinkutharntu-ku*	*mantu-yu*
1SG.NOM	NEG	know	dog-ACC	eat-PFV-ACC	2.SG.GEN-ACC	meat-ACC

'I didn't know a dog had eaten your meat.'

It is interesting to note that the accusative marker in Panyjima *-ku/-yu* is likely cognate with the dative marker in other PN languages, possibly indicating a previously oblique structure or generalized acc/oblique marking, in which case this could be considered an absolute structure[19].

Taken together, the strong tendency in Australian languages for nominalized dependent clauses to be oblique case marked, either as peripheral arguments to the main predication or in concord with a particpant of the main predication, and for the case marking to grammaticalize into indicating subject (dis)continuity across clauses, is one likely main pathway of diachronic development of SR in Australia, and is in agreement with Bickel's hypothesis.

4.3 Typological features and diachronic development of switch-reference in Australia

As discussed above, case concord among elements in NP's and de-verbal nominalizations is a common typological feature of PN languages and forms the backdrop for the development of SR systems and the ways in which it manifests, most overtly as oblique case marking grammaticalizing into SR marking. Also, PN languages are well known for having quite flexible, pragmatically conditioned word order. It seems likely that this feature of the flexibility of word order is the motivating factor for case marking concord on most or all elements of an phrase. Since dependent clauses are usually nominalizations, as opposed to finite clauses with subordinators, this propensity towards case concord is naturally extended to case marking on nominalized dependent clauses. As dependent elements of an NP can be moved around linearly, and in some case appearing quite distant from their head in terms of linear ordering, it seems languages of this type have developed a strategy of concord to help keep track of dependencies between heads and dependents in a phrase. As such it is a natural step from intra-phrasal dependency coding to inter-clausal participant reference tracking systems such as SR, particularly when dependent clauses are de-verbal nominalizations.

Another typologically conditioning factor in Australia is that clause combining tends to be bi-clausal, with dependent clauses embedded as actual core or peripheral arguments of a

[19] In fact I analyze the Panyjima suffix *-jangu* as a DS marker, which could consist of some element *-jan* plus the dative/accusative *-gu*, though this remains speculation at this stage.

main verb,[20] in contrast to the generally non-embedded, coordinate dependent clause structures of clause chains in PNG. This bi-clausal tendency, with the embedding of a dependent clause as an argument, helps to further set the stage for absolute constructions as described above, since they are treated syntactically as arguments in these languages, as opposed to chaining languages, where the clause structure is flatter and more coordinate and does not involve construing dependent clauses as arguments of a higher verb.

Almost all PN languages have ergative morphology (with the notable exception of some nominative/accusative languages of the North-West, such as Panyjima). As mentioned in Section 3.4, many languages of PNG, including SR languages, also have ergative marking, however there are different tendencies in the nature of the ergative marking of languages of Australia and PNG. In PNG, the overt expression of ergativity in general is conditioned more along the lines of the pragmatic expression of topicality, where the A argument of transitive clauses tends to only be coded overtly with ergative marking when it is highly topical. In PN languages, in contrast, ergativity is much more syntactically obligatory and conditioned by transitivity frames, where the A argument has a much stronger tendency to be ergative marked than in PNG, regardless of its topicality status in the discourse. Of course there are variations from language to language in how strongly these tendencies are realized, and this is given here only as general typological tendencies of ergative marking conditions in the two regions. Nevertheless this difference in the underlying motivations for ergative marking in these regions underscores the differing underlying syntactic frameworks in general. Core case marking in Australia is fairly routinized and syntactically conditioned, with a dependence on verb sub-categorization and transitivity, whereas in PNG it tends to be more pragmatically conditioned. This may well be a consequence of the flatter structure of clause combining strategies involving multiple, essentially coordinate clauses in PNG in comparison to the embedded, essentially bi-clausal structures in Australia. A natural way of tracking topics and subjects in multiple coordinate structures is by person agreement and conjunctions, as found in PNG, whereas with embedded bi-clausal structures, nominalization strategies of dependent predications with case concord on the nominalization serve as a natural subject/topic tracking device.

The following is a summary of the typological features in Australian languages shaping the way in which SR systems manifest in Australia:

[20] In this paper, I consider peripheral arguments as embedded, in that they function as arguments of a main verb and figure into the syntactic verb frame of the predication.

- The defining typological characteristic influencing development of SR in Australia appears to be clause embedding and a strong tendency for case concord within phrases, with case concord extended to nominalized clauses in absolute-like constructions.
- Relatively free, pragmatically oriented word order in PN languages correlates with a tendency towards clause embedding as a hierarchy is established between the clauses, as opposed to the flatter, coordinated non-embedded chaining structures in the strict SOV languages of PNG.
- As opposed to Highlands PNG, most PN languages have obligatorily expressed ergative morphology[21] conditioned by the transitivity of verb roots. Moreover, case marking in general in Australian languages appears to be an obligatory feature conditioned by the transitivity sub-categorization of verbs. As an obligatory syntactic feature, it is more likely that extensive concord within NP's and VP's will occur.
- Clause combining structures in Australia tend to be bi-clausal with the nominalization of dependent predications and their incorporation as arguments of a higher verb. A consequence of this is subjects are more likely to be tracked with morphology associated with nominals, such as case marking, than from other sources, such as person marking and conjunction in PNG.

5. Conclusion and areas for further research

Sections 3 and 4 provided the background and details of synchronic SR systems, general typological characteristics and proposals for the development of SR in Highlands PNG and Australia. Table 2 summarizes the typological differences between PNG and Australia and offers a comparison of the different diachronic origins of SR in the two regions.

Table 2 Summary of typological and switch-reference characteristics of Highlands Papua New Guinea and Australia

	Highlands Papua New Guinea	**Australia (Pama-Nyungan)**
Primary clause combination orientation	multi-clause chaining	bi-clausal embedding
Word order	strict SOV	pragmatic 'free'
Case marking	pragmatically expressed	obligatorily marked according to verb sub-categorization

[21] However many Australian languages are actually split-ergative morphologically, in which ergative alignment occurs only on certain categories of nominals, such as full NP's for example.

	Highlands Papua New Guinea	**Australia (Pama-Nyungan)**
SR forms	suffixes on medial verb stems	suffixes on nominalized verb stems
Diachrony of SR morphemes	SR morphology originates primarily from conjunctions and pronouns, and encliticized onto preceding clause. DS may be the first SR category to be established at thematic boundaries, with the subsequent development of the SS category for pragmatic contrast	Oblique case marking on nominalized dependent verbs in absolute constructions, re-analyzed as SR markers, likely first grammaticalizing DS with one type of oblique marker, then as a pragmatic contrast, a different set of oblique markers grammaticalize into SS markers

From the above comparison of the synchronic and diachronic details of SR systems in languages of Highlands PNG and Australia, I suggest that the most important, core consideration in the typological shaping of SR systems in both regions is clause combining type. As discussed above, the predominant clause combining type in Highlands PNG is clause chaining, which is essentially a flat structure of multiple coordinated, dependent clauses. Clause combining in Australia on the hand appears to be more oriented towards a syntactic hierarchy of clausal embedding. As such it appears that clause combining in Australia has a greater tendency towards being bi-clausal as opposed to multi-clausal in its basic syntactic orientation. These considerations also correlate with the semantic and functional types of clauses that appear to predominate in each region; dependent coordinated chains vs. embedded adverbial/relative and purposive clauses. This seems to point to a correlation between multi-clause chaining and coordination on the one hand, and bi-clausal structures and embedded clauses on the other. Of course languages of both regions can and do have ways of expressing a range of clause types, but I suggest that as an underlying syntactic organizing principle, the languages of Highlands PNG and Australia differ in this fundamental way; multi-clausal coordination vs bi-clausal embedding. These underlying syntactic foci have ramifications on the way in which participants are tracked and the devices available for tracking them, as described above for the SR systems of both regions.

The question then arises, why would languages of two different regions have two basic differences in their orientation to clause combining? Different clause combining strategies may well be a result of different inclinations in overall discourse organization, resulting in drastically different morpho-syntactic structures in the expression of clause combining. But what would condition different discourse organization strategies to the degree that it produces significant

differences in the morph-syntax of clause combining, which in turn influences the diachronic development of systems associated with clause combining, such as SR? I speculate that narration styles, influenced by cultural factors involved in the form of narration, ultimately lead languages of these regions to develop different orientations to discourse organization and clause combining overall. Moreover, it may not be coincidental that issues of word order and case marking similarly diverge significantly among these two regions, in that the hypothesized different orientations to discourse organization and clause combining may ultimately also influence word order tendencies, which in turn influences the way in which cases systems develop and the extent to which they form more of a syntactic or a pragmatic role in the grammar of a language. I summarize the above hypothetical points, which remain specualtion at this stage, in the following schemas, with each point leading to the development of the following point:

Australia

cultural narrative style1 > bi-clausal combinations > embedding > pragmatic word order/extensive case marking concord in phrases > SR characteristics of Australia

PNG

cultural narrative style2 > multi-clausal > chaining > fixed sov /pragmatically oriented case marking > SR characteristics of PNG

In the above schemas, since the underlying factors and characteristics of the narration styles, speculatively influenced by cultural practices, are unknown, they are simply left as Style 1 and 2. I hypothesize that for Australia, cultural narrative Style 1 leads to bi-clausal structures, which leads to embedding as a natural result of indicating a hierarchical relationship between two clauses. This interpretation of a hierarchy between clauses in turn leads to case marking as a way of coding that one (nominalized) clause is an argument of a higher clause, whether core or peripheral. Once case marking becomes more routinized and obligatory, based on transitivity frames and appearing on many elements in either a noun phrase or verb phrase, pragmatically free word can then development, since case marking indicates which elements are members of a phrase even when not appearing in linear order.

In PNG, for narration style 2, multi-clausal narrative structures naturally lead to the chaining of temporally sequenced coordinate clauses, though they are morpho-syntactically dependent on a final clause. This in turn naturally leads to fixed SOV order, since sequencing of events is a verbal category and therefore verbs are the most appropriate category to appear at the end of a clause boundary, hence SOV order. Since the order is fixed and clauses are not

embedded, case marking is less important from a syntactic functional perspective, but yet appears on pragmatically marked participants, such as topical nouns or pronouns.

I want to emphasize that the above proposals and schemas are speculative at this stage and need to be verified with empirical data to determine their validity. They are given here as preliminary possible explanations of the issues described in this paper and as avenues of future research into clause combining and SR. Whether or not they ultimately prove to be valid explanations for the observed differences between the two regions, these proposals offer interesting directions of future research, which undoubtable would yield interesting results.

Other directions of future research would be a diachronic account of the choice of oblique markers for SR markers in Australia. Their distribution patterns are well described in Austin (1981), though an account of what motivates the difference in distribution has to my knowledge not yet be developed. Also, relative to PNG, the focus of this paper was languages of the Highlands region, however SR exists in various other forms in languages of other regions of PNG, with certain differing characteristics, though the majority even outside of the Highlands seem to follow clause chaining patterns. An interesting example is the case of Dami, an Austronesian language of Madang province. Dami is the only known Austronesian language of PNG to have SR and it was convincingly argued by Roberts (1997) that Dami's SR system developed due to contact with surrounding Papuan SR languages[22]. Interestingly, in addition to SR, Dami appears to have developed both SOV word order and clause chaining, neither of which are known for being typical for Austronesian languages. The details of Dami serve as another indication that the properties of SOV word order, clause chaining and SR tend to coincide, at least in the regional context of PNG. As the location of Dami is outside of the Highlands region, though it has similar SR features to Highlands languages, it would be very elucidating to look at multiple areas of PNG and see in what ways both similar and differing typological characteristics to those of the Highlands shape the forms and functions of SR systems in other regions of PNG.

References

Alpher, Barry. (2004). "Pama-Nyungan: Phonological reconstruction and status as a phylogenetic group", In C. Bowern, H. Koch (eds.), 93-126, 387-574, 681-687.

Austin, Peter. (1981). "Switch-reference in Australia". *Language*, *57*(2), 309-344.

[22] See also Banks (2023a) for a discussion of SR pattern borrowing in Dami in comparison to other SR contact situations of other regions.

Austin, Peter. (ed.), (1988). *Complex clauses in Australian Languages.* v.15, Typological Studies in Language. John Benjamins.

Banks, Jonathan. (2023a) (in press). "The diffusion of switch-reference systems through language contact: SR system transfer versus morpheme borrowing." *Revue Romaine de Linguistique*, v. 68, Bucharest.

Bickel, Balthasar. (1999). "From ergativus absolutus to topic marking in Kiranti: A typological perspective." Proceedings of the 25th Annual Meeting of the Berkeley Linguistics Society (1999), General Session, 38-49.

Bolinger, Dwight. (1975). *Aspects of Language.* New York: Harcourt & Brace Jovanovich.

Bowern, C., H. Koch. (eds.) (2004. *Australian Languages: Classification and the Comparative Method* (Current Issues in Linguistic Theory 249), Amsterdam/Philadelphia, John Benjamins.

Comrie, Bernard. (1983). "Switch-reference in Huichol: a typological study." In Haiman & Munro (eds.), 17-37.

Daniels, Don. (2015). *A reconstruction of Proto-Sogeram phonology, lexicon and morphosyntax.* PhD thesis: University of California at Santa Barbara.

Dench, A. (1991). "Panyjima". In Dixon, R.M.W. and Blake, B. (eds.), 124-243.

Dixon, R. M. W., B. J. Blake (eds.), (1991). *Handbook of Australian languages*, Vol 4, Melbourne, Oxford University Press.

Dixon, R. M.W. (2002). *Australian Languages: Their Nature and Development*. Cambridge, Cambridge University Press.

Donohue, Mark P. & Catherin Donohue. (1997). "Fore case marking". *Language and Linguistics in Melanesia, 28*, 69-98.

Foley, W. & Van Valin, Robert D. (1984. *Functional Syntax and Universal Grammar*. Cambridge University Press.

Foley, W. (1986). *The languages of New Guinea.* Cambridge University Press.

Foley, W. (2017). "The morphosyntactic typology of Papuan languages." In Palmer (ed.), 2007.

van Gijn, Rik & Hammond, Jeremy (eds.), (2016). *Switch-reference 2.0.* Amsterdam/Philadelphia: John Benjamins.

van Gijn, Rik. (2016a). "Switch-reference: An overview." In van Gijn & Hammond (eds.), 1-53.

van Gijn, Rik. (2016b). "Switch reference in Western South America". In van Gijn & Hammond (eds.), 1-53.

Givón, T. (1983). "Topic continuity in discourse: The functional domain of switch-reference". In Haiman & Munro (eds.), 51-82.

Givón, T. (2001). *Syntax: An Introduction.* Volume I, John Benjamins, 2001.

Haiman, John. (1983). "On some origins of switch-reference marking". In Haiman & Munro (eds), 105-128.

Haiman, John & Munro, Pamela (eds.), (1983). *Switch-reference and Universal Grammar.* (Typological Studies in Language 2). Amsterdam: John Benjamins.

Haspelmath, Martin (2008c). "Frequency vs. iconicity in explaining grammatical asymmetries". *Cognitive Linguistics, 19*(1), 1-33.

Hercus, L. A. (1982). *The Baagandji language*. Pacific Linguistics, Series B-67.

Hymes, Dell & William Bittle, (eds.), (1967). *Studies in Southwestern ethnolinguistics*. The Hague: Mouton.

Jacobsen, William. (1967). "Switch-reference in Hokan-Coahuiltecan". In *Studies in Southwestern ethnolinguistics*. In Hymes & Bittle, (eds.), 238-263.

Longacre, R. (1985). "Sentences as combinations of clauses". In T. Shopen (ed., 1985, Vol. II).

Matić, Dejan. (2016). "Even converbs and the syntax of switch-reference". In van Gijn & Hammond (eds.), 335-376.

McKenzie, Andrew. (2015). A survey of switch-reference in North America. *International Journal of American Linguistics, 81*(3), 409-448.

Mithun, Marianne. (1993). "Switch-reference: Clause combining in Central Pomo." *International Journal of American Linguistics, 59*(2), 119-136.

Overall, Simon & Marine Vuillermet. (2015). "The Eastern foothills as a contact zone: Evidence from non-canonical switch-reference". Talk given at the European Network for the Study of Andean Languages (REELA), Centre for Linguistics, Leiden University, 7 September 2015.

Palmer, Bill. (ed.), (2017). *The Languages and Linguistics of the New Guinea Area: A Comprehensive Guide.* De Gruyter.

Potts, Denise M. and Dorothy James. (1987). "Split ergativity in Siane: a study in markedness." *Language and Linguistics in Melanesia 18*, 71-108.

Roberts, John R. (1997). "Switch-reference in Papua New Guinea: A preliminary survey". In *Papers in Papuan Linguistics, No.3.* (Pacific Linguistics). Canberra: ANU.

Schmalz, Mark. (2016). "Typologically relevant peculiarities of the switch reference system in Yukaghir". In van Gijn & Hammond (eds.), 301-334.

Shopen, T. (ed.), (1985). *Language Typology and Syntactic Description*, 3 volumes (Vol.II: *Complex Constructions*), Cambridge: Cambridge University Press.

de Sousa, H. (2008). “The development of echo-subject markers in Southern Vanuatu”. in T.J. Curnow (ed.), *Selected Papers from the 2007 Conference of the Australian Linguistic Society*.

Stirling, Lesley. (1993). *Switch-Reference and Logophoricity in Discourse Representation Theory*. Edinburgh: EUP.

Treis, Yvonne. (2012). “Switch-reference and Omotic-Cushitic language contact in Southwest Ethiopia”. *Journal of Language Contact, 5*(1), 80-116.

Thurman, R. (1978). *Interclausal Relations in Chuave*, M.A. Thesis, University of California, Los Angeles.

Wilkins, David. (1988). “Switch-reference in Mparntwe Arrernte (Aranda): form, function, and problems of identity”. In Austin, P (ed.), 1988.