The prefix *si-
in Western Indonesian, Sulawesi, and Philippine languages

Inaugural dissertation
to complete the doctorate from the Faculty of Arts and Humanities
of the University of Cologne
in General Linguistics

presented by
Yuko Kitada

born on 26 July 1983
in Tokorozawa
(Saitama Prefecture, Japan)

Cologne, April 2021

Table of Contents

Abbreviations

1	first person
2	second person
3	third person
ABS	absolutive
AC	actor
AD	action-denoting
APPL	applicative
ART	article
AV	actor voice
DU	dual
DEF	definite
EMPH	emphatic
EP	event plural
ERG	ergative
GCP	Greater Central Philippine
GEN	genitive
INDEF	indefinite
INV	inversion
KIN	converse kinship pair/group
LK	linker
LOC	locative
LV	locative voice
NEG	negation
NOM	nominative
OBL	oblique
OD	object-denoting
PAn	Proto-Austronesian
PARI-	reflex of Proto-Oceanic *paRi-
PMP	Proto-Malayo-Polynesian
PE	plural exclusive
PI	plural inclusive
PN	personal name
PL	plural
POc	Proto-Oceanic
POSS	possessive
PROG	progressive
PV	patient voice
QD	quality-denoting
RDP	full reduplication
REC	reciprocal
REFL	reflexive
REL	relativizer
SG	singular
TV	thematic voice
UG	undergoer
UV	undergoer voice

1 Introduction

This study explores the semantic and morphological properties of reflexes of the prefix *si- in Malayo-Polynesian languages in Western Indonesia, Central and South Sulawesi, and the Philippines (abbreviated as 'WISP languages').[1] Employing data from approximately 30 languages in these areas, the current study discusses various formations with *si- from comparative, historical, and typological points of view.

Previous comparative and historical studies on Austronesian verbal affixes have been primarily concerned with those indicating voice, mood, and aspect. By contrast, little is known about the remaining verbal affixes from a comparative and historical perspective, despite the fact that Austronesianists have benefited significantly from the recent increase in descriptive works on individual Austronesian languages. The current study concerning formations with *si- and their meanings will contribute to a more comprehensive understanding of the morphological history of Malayo-Polynesian languages.

One of the interesting points with regard to the *si- formations is that they cover a broad range of meanings related to situations with plural participants. For example, when they appear with action-denoting (i.e., verbal) roots in actor voice constructions, they generally encode situations with plural actors in Philippine languages and reciprocal and reflexive situations in Indonesian languages:

Actor plural in actor voice: Tagalog *mag-si-*

(1) **mag-si-**kanta ʻsing (plural)' — k<um>anta ʻsing' (Schachter and Otanes 1972:335)

Reciprocal: Toba Batak *mar-si-*

(2) **mar-si-**gulut 'to quarrel with each other' — gulut [ʻquarrel'] (van der Tuuk 1971:140)

Reflexive: Toba Batak *mar-si-*

(3) **mar-si-**ájar 'study [teach oneself]' — ájar 'teach ([transitive])' (Nababan 1981:99)

Furthermore, formations with possible reflexes of *si- are also employed to indicate, for example, the equal state of two participants in subject function with quality-denoting (i.e., adjective) roots, such as Tagalog *mag-ka-sing-bago* '(two items) are equally new' (< *bago* 'new'), the 'associate' meaning with object-denoting (i.e., nominal) roots, such as Cebuano *i-sig-ka-tawo* 'fellow being' (< *tawo* 'person'; lit. 'an associate who is also a person'), and the converse kinship pair meaning with kinship terms, such as Cebuano *ma-sig-ka-asawa* 'husband and wife' (< *asawa* 'wife').

These forms seem formally similar because they share the common segment *si-; however, their functions are heterogeneous. We may attribute some of the semantic differences to the formal distinctions and the semantic types of the roots. For example, some *si- formations contain the segment *ka- when they occur with quality- and object-denoting roots (including kinship terms); on the other hand, they do not when they appear with action-denoting roots. However, the same affix combinations do not always signify the same meaning. For example, the form *maR-si- (> Tagalog *mag-si-* and Toba Batak *mar-si-*) indicates an actor plural construction in one language, but a reciprocal construction in another. In addition, the forms display some other segments, such as /g/

1 A more comprehensive definition of the term is provided in Section 1.3.

after /si/ in the Aklanon form *mag-si(g)-*.

Our research question here is whether and how the forms with *si- that display a variety of meanings are related. To answer this question, this study will identify and analyze the relevant formations and their functions from phonological, morphological, syntactic, and semantic perspectives.

1.1 Methodology

WISP languages have a number of derivational affixes. Some of these affixes are regarded as 'multifunctional'. For example, Blust (2003a) assigns seven functions to the Proto-Austronesian prefix *ka-.

However, in view of the fact that WISP languages generally have simple phonological structures, it is plausible to assume that a prefix may become formally indistinguishable from other affixes and may become semantically perceived as being related to them in some languages. For example, the Malay prefix *si-* (allomorph *se-* /sə-/) had not been recognized as being a different prefix from the prefixed numeral *se-* /sə-/ 'one' in Malay and Indonesian grammars until Adelaar (2005a:132) differentiated between the two because both morphemes appear to be homophonous (Adelaar ibid.) and possibly because it is easy for speakers and researchers to relate the two meanings. This example implies that the formal and semantic resemblance of the *si- forms may stem from a formal merger of different sources. Such a merger cannot be detected if we only consider the phonological correspondences among languages or phonological similarities between two items within a language.

Historical linguistic studies typically employ a technique termed the 'comparative method' to explore relationships among languages or those among word forms. This method primarily involves comparing phonological correspondences of a given item in various languages on the basis of the Neogrammarian hypothesis that 'sound changes are regular'.

However, some morphological items cannot be easily identified via such regular sound correspondences. Ringe and Eska (2013:253) claim that '[t]here is no method for reconstructing morphology or syntax comparable to the comparative method for phonology [… because m]orphemes (the units of morphology) and the rules of syntax are not meaningless items distributed arbitrarily through the utterances of a language'. In other words, comparisons of morphological and syntactic items do not succeed because 'morphological and syntactic changes cannot and do not exhibit *mathematically exploitable recurrent* regularity in the way that phonological changes do' (Ringe and Eska 2013:253–254, emphasis in the original).

Along the same lines, Ringe and Eska also argue for the necessity of more insights than mere regular sound correspondences for the reconstruction of meanings (i.e., semantic representations, of linguistic items). This claim is particularly true when a semantic shift occurs because the meaning of a lexical item may become different from that of its phonological correspondences in other languages after the shift. Ringe and Eska (2013:254) claim that such semantic changes are caused by extra-linguistic factors or linguistic components that are yet to be established. For this reason, the reconstruction of the meaning of a proto-lexeme (i.e., proto-lexical item) is 'still very much a matter of guesswork informed by experience' (Ringe and Eska ibid.).

For resolving such formal and functional mismatches in both lexical and functional items upon reconstruction, Koch (2003:271) proposes what he terms the 'etymological method'. This method refers to a comprehensive and systematic technique for detecting cognates by using the comparative method applied to phonology, morphology, and semantics, with some flexibility with regard to

phonological, morphological, and semantic changes.

Phonological comparisons do not always produce good results in reconstructing the meanings of lexical items because 'many of the true cognates that exist in the related languages are not to be found as exactly matching translinguistic synonyms' (Koch 2003:273). For example, in comparisons between English and German words, while some English lexical items have phonologically corresponding German forms with the same meanings, such as the English *cow* and German *Kuh* pair, other lexical items, such as 'dog', 'animal', and 'head', have phonologically corresponding German forms with different but still associated meanings (Koch ibid.):

Table 1: English-German comparative wordlist (cf. Koch 2003:274)

Gloss	English	German	Cognate?	other English	other German
'cow'	cow	Kuh	yes		
'dog'	dog	Hund	no	hound	
'animal'	animal	Tier	no	deer	
'head'	head	Kopf	no		Haupt[-] 'main'

According to Koch (2003:273), these imperfect synonymic pairs 'yield nothing of use for reconstruction' if we only compare their forms. However, using the etymological method that enables us to consider the possibility of semantic change, such as generalization and specification, they can also be included for consideration in reconstructing the form and meaning of an item (Koch ibid.).

Similarly, morphological reconstruction is also impossible if we simply rely on phonological correspondences. Consider the following comparative wordlist of the Arandic family in Australia:

Table 2: Arandic comparative wordlist (cf. Koch 2003:274)

	'forehead'	'eye'	'flame'
North Alyawarr	urle	atnnge	lherrme
Akarre	urle	alknge	lherrme
Eastern Arrernte	urle	alknge	alkngenthe
Anmatyerr	urle(ange)	anng(ange)	anngenthe
Southern Arrernte	rletne	alknge	alkngenthe
Kaytetye	rlunperre	urle	urlenthe

Again, we can find imperfectly matching items. This time, most of them are items with a perfect semantic match but with only a partial morphological match. By using the etymological method, Koch (2003:274–275) analyzes the connections between different formal representations and provides the following explanations:

- The proto-forms *urle 'forehead' and *atknge 'eye' can be reconstructed.
- Anmatyerr *-ange* is an increment that carries the stress in the citation nominative case. In other case forms, such as the locative form *anngele* and the dative form *anngeke*, a regular reflex of *atknge 'eye' appears as the expected form for this increment.

- Southern Arrernte *rletne* consists of *rle* (< *urle ‘eye’) and the suffix *-tne* of uncertain origin.
- The Kaytetye word *rlunperre* consists of *rlu* (< *urle ‘eye’) and *-(t)nperre* of uncertain origin.
- The lexical item ‘flame’ in Eastern Arrernte, Anmatyerr, and Southern Arrernte comprises *alknge ‘eye’ and the segment *nthe*, which is assumed to indicate ‘flame’ because it is attested as a component of the Eastern and Central Arrernte word *nthile* ‘light a fire, make flames’ (< *nthe* ‘flame’ and *-ile-* ‘make’).
- The forms *alkngenthe* ‘flame’ and *anngenthe* ‘flame’ suggest that in this language group, the lexical item ‘flame’ consists of an item that means ‘eye’ and the segment *nthe*. Thus, the Kaytetye *urlenthe* ‘flame’ is a partial cognate of *urle ‘eye’.
- Subgroup-external evidence suggests that *urle originally meant ‘forehead’ and has come to indicate an ‘eye’ in Kaytetye.

As such, the etymological method (i.e., a comprehensive approach that takes phonological, morphological, and semantic correspondences into account) yields more plausible explanations for phonological irregularities that cannot be explained by phonological reconstruction. Furthermore, the etymological method is also useful for the investigation of functional items, as Koch (2003, 2009, and 2013) demonstrates in reconstructing some proto-forms of functional items, such as pronouns, inflectional affixes, and case markers.

Our scope, the prefix *si-, is more lexical (contentful) in nature than are functional items proper, such as tense markers. At the same time, it is less lexical than are lexical items proper, such as ‘dog’. In other words, the prefix is more subject to semantic/functional changes than are functional items are and is more subject to formal changes than are lexical items. Therefore, the etymological method is ideal for tracing the formal and functional changes that are not captured by phonological comparisons. Furthermore, we consider syntactic properties in this study in order to form a more comprehensive account of the functions indicated by reflexes of *si-, in addition to phonological, morphological, and semantic correspondences that are taken into account within the etymological method.

As indicated above, we need to define our scope concerning the morpheme *si- more precisely because, due to the relatively simple phonological structures in WISP languages, the comparisons of phonological and morphological representations of a functional item may result in the inclusion of some non-cognates. To avoid such irrelevant forms, the current study limits its scope to first addressing the WISP forms that fulfill the following three criteria: 1) They contain the key segment, *si or *sV, 2) take action-denoting (verbal) roots, and 3) encode actor plural, reciprocal, and reflexive situations. Following the identification of the forms and functions with action-denoting roots, we explore forms that 1) include the segment *si or *sV, 2) take other types of roots, and 3) indicate the meanings semantically associated with the three categories. These relevant forms are termed ‘*si- forms.’ In this process, following and expanding the argument by Adelaar (2005a), we exclude the reflexes of the numeral form *sV(N)- ‘one’, such as Malay *se-,* as in *se-ratus* ‘a hundred’[2] (*ratus* ‘hundred’).

Having defined the method and target of the data collection, what we should consider at this stage is the method of handling data from at least 300 WISP languages with few established internal subgroups. Ideally, any research that covers all the languages in a given family must conduct detailed studies of the individual languages. However, such a method is impossible at present because most of the languages that are relevant to us have not been sufficiently described.

2 Data from Sneddon (1996:56).

Moreover, even if we had all the data, it would be unrealistic to scrutinize them because they involve hundreds of languages. Nevertheless, in a similar manner to other authors, the utilization of our available data enables us to conduct a comparative and historical study. Research practices in typological works serve as a good model for the comparative works considering these under-classified but related languages because typology examines more than 6,000 languages belonging to different language families.

According to Nichols (2007:233), typological studies should ideally consider a large number of well-sampled languages across language families and linguistic areas because typology is statistical in nature. However, as Nichols also states, such large surveys in typology are not particularly common. They are typically used when the relevant information for data classification is readily available. Such classificatory criteria as those employed in typological studies emerge from hypothesis-generating works with smaller samples. Taking the hypothesis raising of typological studies as a first step, as Polinsky and Kluender (2007:280) propose, requires a method involving 'comparison of a relatively small, intelligently selected sample of languages, using deeper structural analyses than is (*sic*) currently possible with large-scale sampling methods'. This smaller-sampling strategy would also involve a 'comparative investigation of closely related languages' (Polinsky and Kluender ibid.). A good example of such a smaller-sampling approach is the series of comparative Romance works conducted by Richard Kayne and his students. It was based on the comparisons among French, Italian, and Spanish; however, it helped to create a key testing ground that later provided enormous benefits for in-depth research regarding closely related languages and dialects (Polinsky and Kluender ibid.).

Similarly, Nichols began her study concerning head- and dependent-marking languages with the comparison of two Caucasian languages, Abkhaz and Ingush (Nichols 2007:234). The pilot study was expanded to a convenience sample of about a dozen languages, the descriptions of which were familiar to her. Her idea was later demonstrated using a 60-language sample (in Nichols 1986) and, finally, in a metric study with as many as 174 languages (in Nichols 1992). In addition, her intransitivizing and detransitivizing lexical metric (i.e., Nichols 1982) developed from the comparison between Ingush and Russian verbal lexicalization and expanded in a similar manner to her study on head- and dependent-marking languages (Nichols 2007:234). Ingush and Russian are not genetically related; however, Nichols's success in the comparison between the two languages is assumed to be attributable to her comprehension of Ingush as a language in her research and that of Russian as a medium for her Caucasian linguistic studies.

Similar small-sampling methods have been practiced in Austronesian comparative and historical linguistic studies. For example, Ross's reconstruction of a Proto-Malayo-Polynesian voice, mood, and aspect paradigm is based on his reconstruction of a Proto-Austronesian voice, mood, and aspect paradigm and his examination of three languages and one cluster of dialects: Ilocano, Tagalog, Binukid, and the Bisayan dialects (Ross 2002:49 fn.). A hypothesis-testing large-scale work with regard to this reconstruction has not been conducted; however, despite slight variations, the paradigm reconstructed by Ross can also be observed in languages that he does not employ in his reconstruction, such as Buol (Gorontalo-Mongondowic, northern Sulawesi) (cf. Zobel 2005) and Kimaragang (Dusunic, northern Borneo) (cf. Kroeger 2005). This provides evidence for some reliability of the small-sampling method for linguistic reconstruction in WISP languages.

For reconstructing proto-forms based on WISP data, it is ideal to begin with bottom-up reconstructions with data in microgroups that are relatively well established (cf. Adelaar 1984; Reid 2009). In reality, however, this small-scale reconstruction method is not possible in most cases for want of sufficient data. For example, in our database, very few languages have *si- forms or good descriptions of these forms, and some microgroups do not even have a language with a *si- form. Even if relevant data are available in a language, it is difficult to find other instances in another language in the same microgroup in most cases. For these reasons, we directly posit the ancestral

forms for WISP languages, based on a few sample languages.

The primary sample languages were determined based on Ross's method for the reconstruction of the Proto-Malayo-Polynesian inflectional morphemes. They comprise two Philippine languages, namely, Tagalog and Cotabato Manobo, and one non-Philippine, but well-studied language, namely, Malay. These Philippine languages and the Malay language were selected to serve as the representatives of Philippine and non-Philippine languages in an initial hypothesis-generating process, respectively. In a similar manner to Nichols, who conducted successful hypothesis-generating studies, the current author is familiar with all the descriptions of these languages.

As stated above, we employed our large-scale language sample consisting of approximately 30 WISP languages to test the hypothesis based on the small-scale sample that comprises the four languages. In fact, we also consulted other grammars and dictionaries in that category; however, it was decided that they should be excluded here because they either did not list *si-forms or did not provide sufficient information about them.

We have consulted as many grammars and dictionaries as possible for the current study. The linguistic studies consulted include several literature concerning the Philippine languages, which was published by the Summer Institute of Linguistics (SIL) and several Dutch literature concerning Sulawesi languages. However, some of the literature could not be employed in the current study. For example, we were not able to obtain sufficient access to literature concerning languages in Borneo. Similarly, due to the insufficiency of the information with regard to the relevant structures, we could not utilize the collections of the grammar sketches and dictionaries of local languages published by the Pusat Pembinaan dan Pengembangan Bahasa (The Center for Language Learning and Research) in Indonesia in the late twentieth century. It is hoped that such literature will be employed in our future larger-sample research.

The WISP data employed in the current thesis were processed as follows. Initially, we classified our WISP data according to the meanings of words derived with *si- and the lexical types of the roots or bases with which the prefix *si- co-occurred. We identified five lexical types: 1) action-denoting roots, 2) quality-denoting roots, 3) object-denoting roots, 4) kinship terms, and 5) location words. Among these lexical classes and types, the action-denoting roots are of primary concern. For this reason, *si- formations with the action-denoting roots are addressed extensively in Chapter 3. Formations with *si- and other lexical items will be discussed as possible cognates with the *si- formations for action-denoting roots in Chapter 4.

Chapter 3 begins with the formal characteristics of *si- formations and their functions with action-denoting roots based on the studies by Adelaar (1992a, 2005a), with additional data from Tagalog, Cotabato Manobo, and a few other Manoboic languages. In this process, it is assumed that *si- forms are inflected for voice. Based on this assumption, we identify the relevant forms in actor voice constructions in other languages and present their formal and functional variations. Following the discussion of the *si- forms in actor voice constructions, we present the formal and functional variations of *si- forms in undergoer voice constructions, which are identified based on the initial assumption.

Chapter 4 explores forms with the possible reflexes of *si- with other lexical types. The lexical types addressed are 1) quality-denoting roots, 2) object-denoting roots, 3) kinship terms, and 4) location words.

Note that reconstructed forms based on our WISP data are labeled as Proto-Malayo-Polynesian (abbreviated as PMP), despite the fact that Proto-Malayo-Polynesian forms should also be based on evidence from Central and Eastern Malayo-Polynesian languages.

1.2 Previous studies

No literature has yet exclusively addressed formations with reflexes of the proto-form *si- as a Proto-Malayo-Polynesian prefix. However, there have been a few attempts at reconstructing the prefix. For example, besides discussing other issues, Adelaar (1992a, 2005a) posits the proto-form *si- as a reciprocal and reflexive marker and reconstructs it at the Proto-(West-)Malayo-Polynesian level (Adelaar 1992a) and the Proto-Austronesian level (Adelaar 2005a). Blust (2003a) lists *si- as a 'distributive' marker at the Proto-Malayo-Polynesian level; however, its use is less clear because no example in a sentence context is provided, and cross-linguistically, various situations are assigned to the term 'distributive' (cf. Kemmer 1997; Lichtenberk 2000; Nedjalkov 2007b; Cabredo Hofherr and Laca eds. 2012). In addition, although it is a lower node than Proto-Malayo-Polynesian, Zorc (1977) suggests the (Proto-Bisayan) prefix *si(g)- 'individual action' for Bisayan languages of the Central Philippine subgroup, stating that 'the prefix *si(g)-* is added to verb stems to show that the action is done individually' (Zorc 1977:143).

For Austronesian languages other than WISP languages, several attempts have been conducted for the reconstruction of reciprocal, actor plural, and reflexive markers. The most well-known instance of such affixes would be the Proto-Oceanic 'reciprocal and collective marker' *paRi-, which is reconstructed by Pawley (1973) and developed by Lichtenberk (1985, 2000). The formal and semantic features of this prefix suggest its cognacy with the 'social prefix' *baɣi- reconstructed by Dempwolff (1920:15–16) for Proto-Malayo-Polynesian. On the other hand, based on Formosan data, Zeitoun (2002) suggests the Proto-Austronesian forms *ma-Ca- ~ *pa-Ca- 'dynamic reciprocal [i.e., reciprocal form for action-denoting roots, YK], and *maR- ~ *paR- 'stative reciprocal [i.e., plural forms for object-denoting roots and converse kinship pair forms for kinship terms, YK]'.

However, few comparative studies have exclusively dealt with morphemes that encode reciprocal, actor plural, or reflexive situations, respectively. The scarcity of the studies may be attributed to the existence of various competing forms encoding such situations. For example, reciprocal formations are encoded in various formations in the Malayo-Polynesian languages, such as Tagalog *mag-* 'reciprocal' and *mag-RDP-* 'reciprocal actions of explicitly more than two actors' (Bloomfield 1917:239), Malay *ber- (-an)* (Sneddon 1996), and Cebuano *mag- -ay* (cf. Wolff 2012 [1972]). Reid and Liao (2004:457) claim that Proto-Malayo-Polynesian *maR- is a reciprocal and reflexive marker; however, they do not mention reciprocal formations that reflect other morphemes along with *maR-. It is also known that morphological items encoding actor or nominal plurality are not rare in Austronesian languages. For instance, the reflexes of the Proto-Oceanic *paRi- encode actor plurality in some Oceanic languages (Lichtenberk 2000), and full reduplication, as in Malay *anak-anak* 'children'[3] (< *anak* 'child'), is employed to indicate nominal plurality in some WISP languages (Blust 2003 and 2013). Similarly, it is known that reflexive situations are formed in various ways in the languages at issue. They are encoded morphologically by reflexes of *maR- in some WISP languages (Reid and Liao 2004:457). Furthermore, they can also be expressed periphrastically with lexical items, such as Malay *diri*, as in *Ia mem-bunuh diri* 'She killed herself'[4] (3.SG AV-kill oneself).

In contrast to the paucity of comparative and historical studies concerning the three categories in WISP languages, cross-linguistic studies with regard to reciprocals, reflexives, and their polysemies are many. For example, Frajzyngier and Curl eds. (2000a) (including Lichtenberk 2000), Nedjalkov ed. (2007), König and Gast eds. (2008), and Evans et al. eds. (2011) deal with reciprocal constructions and their polysemies. On the other hand, for instance, Geniušienė (1987), Kemmer (1993), Frajzyngier and Curl eds. (2000b), and König and Gast eds. (2008) discuss reflexive

3 Data from Blust (2013:419).

4 Data from Echols and Shadily (1994:145).

constructions and their polysemies.

Most of these studies deal with these categories in individual languages. Hence, generalizations with regard to these polysemies and their developments are language-specific. However, some studies, such as Kemmer (1993), Lichtenberk (2000), and König and Gast (2008), offer some suggestions about relations between categories in reciprocal polysemies. These suggestions assist us in identifying WISP forms with *si- and analyzing their functions.

1.3 Languages considered in this study

This study is concerned with some Malayo-Polynesian languages spoken in insular Southeast Asia, in particular, Indonesia and the Philippines. For convenience and accuracy, these languages are referred to as 'Western Indonesian, Sulawesi, and Philippine languages', abbreviated as 'WISP languages'.[5]

The label 'WISP' refers to three areal groups of languages, namely, Western Indonesian, Sulawesi, and Philippine. These language groups include five genetic higher-order subgroups, which are Smith's Western Indonesian, Sumatran, Celebic, and South Sulawesi subgroups, and the Philippine linkage or subgroup. While the former four subgroups are part of the primary Malayo-Polynesian branches posited by Smith (2017)[6], the Philippine linkage or subgroup is proposed by Zorc (1986). The use of the label 'WISP' is only to facilitate reference to the languages that are the main focus of the current research. Accordingly, we do not claim that the WISP languages form a genetic subgroup.

The following paragraphs will provide details of the languages that are referred to by the three terms of the areal groups. For the sake of explanation, the areal groups designated as WISP will be presented in reverse order because the Western Indonesian areal group can be defined more conveniently by first eliminating the Philippine and Sulawesi areal groups. Unless otherwise noted, the genetic classification will follow the one employed by Smith (2017) in the following explanations of the three areal groups.

The term 'Philippine languages' refers to the languages in the Philippines as well as in Northern Sulawesi. The microgroups that belong to this subgroup are Batanic, Cordilleran [Northern Luzon], Central Luzon, Kalamian, Southern Mindanao, Greater Central Philippine (abbreviated as GCP), Sangiric, and Minahasan subgroups, as well as the Inati language.

The term 'Sulawesi languages' refers to the Celebic and South Sulawesi subgroups. The Celebic subgroup consists of six microgroups, namely, the Tomini-Tolitoli, Kaili-Pamona, Saluan-Banggai, Bungku-Tolaki, Muna-Buton, and Wotu-Wolio microgroups.

The term 'Western Indonesian languages' refers to the remaining languages in the western area of the Malayo-Polynesian languages that are considered in this study and correspond to the Western Indonesian and Sumatran languages in Smith's genetic classification. In his classification, the 'Western Indonesian group' consists of 'all indigenous languages of Borneo, plus the Austronesian language of Sumatra (excluding Batak, Barrier Islands languages, and Nasal), Javanese, Madurese, Balinese, Sasak, [...] Sumbawa', and Sundanese in West Java (cf. Smith 2017:443, 480). It should be noted that unless otherwise noted, the term 'Western Indonesian' refers to the languages according to the areal classification in the current study: Smith's use of the term is referred to as

5 The current author is grateful to Professor Sander Adelaar for his suggestion of the term (personal communication, May 2018).

6 Smith (2017) disapproves of a 'Western Malayo-Polynesian linkage/subgroup', proposing instead seven primary branches (Western Indonesian, Sumatran, Celebic, South Sulawesi, Moken, Palauan, and Chamorro), which occur alongside Philippine and Central-East Malayo-Polynesian languages within Malayo-Polynesian.

'Smith's Western Indonesian languages'.

Among such microgroups and languages, those that are considered in the current study concerning the prefix *si- are Northern Luzon, Greater Central Philippine[7] (Central Philippine, Manoboic[8], and Gorontalo-Mongondowic[9]), Sangiric, Tomini-Tolitoli, Kaili-Pamona, Muna-Buton, South Sulawesi, Sama-Bajaw, Berawan-Lower Baram, Malayic, Sundanese, Old Javanese, Lampung, and Sumatran. In addition to these, data in languages belonging to the Bungku-Tolaki and Wotu-Wolio microgroups are also employed to discuss structures of WISP languages in Section 1.4. The following paragraphs will provide their locations and detailed clarifications regarding each subgroup, proceeding from north to south.

The Northern Luzon[10] subgroup refers to 'most languages of the Cordillera Central in northern Luzon, and a few lowland languages such as Ilokano [Ilocano, YK] and Ibanag' (Blust 2013:740).

The Greater Central Philippine subgroup suggested by Blust (1991) includes six microgroups in the central and southern Philippines, namely, the Central Philippine, South Mindoro[11], Palawanic, Manoboic, Danaw, and Subanun microgroups, and the Gorontalo-Mongondowic microgroup in northern Sulawesi (Blust ibid.).

Sangiric languages are spoken in the Sangihe (Sangir) and Talaud Islands, the northernmost part of mainland Sulawesi, and the southern tip of Mindanao in the Philippines (Blust 1991:84).

Tomini-Tolitoli languages are spoken in 'northern Central Sulawesi' (Himmelmann 2001:xxi), the area that is 'bordered on the West Coast by the Makassar Strait, and on the East Coast by the Tomini Bay' (Himmelmann 2001:45). Kaili-Pamona languages 'occupy the whole of the central part of Central Sulawesi and parts of northern South Sulawesi' (Sneddon 1993:16). Languages of the Bungku-Tolaki subgroup refer to those on the southeast coast of Sulawesi (Sneddon 1993:18). Muna-Buton and Wotu-Wolio languages are situated on islands off the southeastern part of the Sulawesi island (Donohue 2005:21–22).

The South Sulawesi subgroup includes Buginese and Makassarese in southern Sulawesi. It also includes the Tamanic languages, such as Embaloh, Taman, and Kalis, at the upper reaches of the Kapuas River in the northeastern part of West Kalimantan on the island of Borneo (Adelaar 2005b:24).

Sama-Bajaw languages are 'spoken in several spots along the coasts of the southern Philippines, Sabah, eastern Borneo[,] and eastern Indonesia' (Adelaar 2005b:16). Blust (2007, 2013) includes them in the eastern branch of the Barito language group (or linkage, cf. Smith 2017), which is a sub-division of Western Indonesian languages.

Languages of the Berawan-Lower Baram microgroup are spoken in northern Sarawak in Borneo (Blust 2010).

The Malayic varieties are spoken in coastal areas of the Malay Peninsula, in parts of Borneo, in South and Southeast Sumatra, and in nearly all major trade centers of the Indonesian Archipelago. They belong to the Malayic microgroup (Adelaar 1992b:1).

7 Blust (1991) terms this subgroup 'Greater Central Philippines'; however, we employ the term 'Greater Central Philippine', in accordance with the term 'Central Philippine' languages, which is employed in Blust (ibid.).

8 While Blust (1991) terms this subgroup 'Manobo', we term it 'Manoboic' because the term refers to a subgroup and not a language and the subgroup has some languages termed with 'Manobo', such as 'the Cotabato Manobo language'.

9 Blust (1991) terms this subgroup 'Gorontalo-Mongondow'; however, we term it 'Gorontalo-Mongondowic' because it refers to the subgroup that comprises not only the Gorontalo and Mongondow languages but also other languages (e.g., Suwawa). It should be noted that for simplicity, we do not make such terminological changes in terms that refer to Sulawesi and Western Indonesian languages, such as the Kaili-Pamona subgroup.

10 It is also termed 'Cordilleran' (cf. Blust 2013).

11 Originally referred to as 'South Mangyan' in Blust (1991).

Sundanese and Old Javanese in Java and Lampung in Sumatra are also included in Smith's Western Indonesian languages in this study.

The Batak microgroup is located in North Sumatra around Lake Toba (Woollams 1996:1) and belongs to the Sumatran subgroup.

The languages from each group examined in this study are listed in the following table:

Table 3: Languages considered in this study

Subgroup	Languages
Philippine	
- Northern Luzon	Ilocano, Pangasinan, Bontok, Dupaningan Agta, Central Cagayan Agta
- Greater Central Philippine	
- Central Philippine	Tagalog, Cebuano, Aklanon, Hiligaynon
- South Mindoro	Hanunoo
- Manoboic	Cotabato Manobo, Sarangani Manobo, Western Bukidnon Manobo, Binukid, Matigsalug Manobo
- Gorontalo-Mongondowic	Mongondow
- Sangiric	Sangir
Celebic	
- Tomini-Tolitoli	Pendau
- Kaili-Pamona	Uma, Da'a, Pamona
- Muna-Buton	Muna
- Bungku-Tolaki	Tolaki
- Wotu-Wolio	Wolio
South Sulawesi	Makassarese
Smith's Western Indonesian	
- Barito, Sama-Bajaw	West Coast Bajau
- Berawan-Lower Baram	Kiput
- Malayic	Malay, Salako
- (Unclassified)	Sundanese, Old Javanese, Lampung
Sumatran	Toba Batak, Karo Batak

1.4 Lexical classes, inflectional paradigms, and case-marking systems in WISP languages

This section provides a brief introduction to lexical classes, inflectional paradigms, and case-marking systems in WISP languages for readers who are not familiar with the languages, along with a clarification of the terminology used in reference to these systems. These three systems should be discussed for the following reasons: First, the distinction between lexical classes of linguistic items, such as roots and words, is controversial in some WISP languages. Second, most WISP languages exhibit multiple voice alternations that are unique from a cross-linguistic perspective. Third, case-marking systems must be mentioned in connection with the multiple voice alternations.

1.4.1 Lexical classes

This study employs the term 'lexical class' to refer to the classes of roots. Such root classes are classified solely based on the semantic characteristics of the root. This classification differs from the general method employed in the literature on word class classification, which is usually based on both semantic and formal (morphological and syntactic) criteria (Nichols 2016).

The primary reason for employing this classification is that the nature of word classes in some of the WISP languages is a matter of debate (cf. Himmelmann 2008; Gil 2013). Although we do not discuss such issues due to the current purpose of this study, the semantic varieties and syntactic behaviors of *si- words (i.e., words derived with reflexes of *si-) still require clarification regarding the semantic distinctions of the roots or bases to which the prefix is attached. In classifying types of words or roots, the use of terms such as 'verb' and 'noun' is commonly accepted; however, these terms are not suitable for the current purpose because 1) they may invoke potentially inaccurate syntactic associations, and 2) the use of these 'traditional' terms varies substantially from author to author. For example, with regard to the latter reason, the traditional terms may refer to the semantic, morphological, syntactic, or pragmatic features of a given contentful item (cf. Himmelmann 2017), and the item can be a root, a stem, or a word (cf. Lehmann 2008). This potential variety of the uses of terms may lead to unnecessary confusion for readers. Hence, we adopted the following practical solution in order to avoid such issues.

In making distinctions among lexical classes of the roots, it is sufficient to use somewhat loosely defined semantic terms to describe the semantic features of roots. Thus, we employ three terms, namely, 'action-denoting', 'object-denoting', and 'quality-denoting'. Action-denoting roots refer to roots that encode actions, such as 'walk' and 'hit'. Object-denoting roots refer to roots that indicate objects, such as 'man' and 'flower'. Quality-denoting roots refer to roots that encode qualities, such as 'red' and 'beautiful'. These three terms correspond approximately to the terms 'verb', 'noun', and 'adjective' in the traditional parts-of-speech terminology, as shown in the following table:

Table 4: Lexical classes for the WISP data

Lexical class of a root	Item encoded by the root	Corresponding 'part of speech'
Action-denoting	Action	Verb
Object-denoting	Object	Noun
Quality-denoting	Quality	Adjective

It should be noted that the terms 'predicate' and 'argument' are employed to describe the

syntactic behavior of items in a sentence context in WISP languages. However, when discussing other languages, the original terms found in the literature, such as 'verb', 'noun', and 'adjective', are used due to the lack of sufficient information for converting the terminologies.

1.4.2 Inflectional paradigms in WISP languages

The term 'inflectional paradigm' is employed here to refer to the systems in which voices, moods, and aspects are marked morphologically by affixes in WISP languages. The systems in question in WISP languages approximately correspond to what has been termed as 'inflectional paradigms' in the grammatical traditions in Western European languages, such as Latin, Greek, and English, in the sense that the majority of action-denoting roots have their formal varieties that differ in some grammatical categories, such as voice and mood.

The inflectional paradigms consist of three grammatical categories: voice, mood, and aspect. Among these categories, voice is the most significant because almost all WISP languages have voice alternations, which can be identified as having been descended from the Proto-Malayo-Polynesian language. The other two categories are less significant because some of the items in the categories 1) are not present or 2) have been superseded by other items in some WISP languages.

The inflectional systems vary in our sample WISP languages; however, according to Ross (2002), the following morphemes can be identified as the constituents in the systems:

Table 5: Proto-Malayo-Polynesian inflectional paradigm after Ross (2002)[12]

	Actor voice	Patient voice	Locative voice	Thematic voice (Circumstance voice)
Indicative				
Neutral	*<um>	*-en	*-an	*i-
Realis (Perfective)	*<um><in>	*<in>	*<in> -an	*i-<in>
Progressive (Imperfective)	*<um>CV-[13]	*CV- -en	*CV- -an	*i-CV-
Non-Indicative				
Subjunctive 1 (Atemporal)	*Ø	*-a	*-i	*-án
Subjunctive 2 (Projective)	*-a	*-aw	*-ay	[unknown]

In general, Philippine languages reflect most of these affixes. By contrast, Sulawesi and Western Indonesian languages do not reflect many of these affixes as being productive in nature.

The following subsections explain voices, moods, and aspects in WISP languages.

12 Some of the original terms employed in Ross (2002) have been renamed in the current study. The original terms are listed in brackets.

13 Originally written as 'R-'; however, his examples suggest that it is CV-.

1.4.2.1 Voices

One of the typological peculiarities of voice alternations in WISP languages is that many of these languages have more than two basic voices, namely, the categories referred to as the 'active voice' and the 'passive voice' in the majority of (traditional) grammars of other languages. The multiple voice system consists of one actor voice construction and more than two undergoer voice constructions; both actor and undergoer voices are marked. To explain the mechanisms of the voice alternations in question in more detail, we employ terms that clarify the semantic role of the subject that the voice form assigns, such as 'actor voice', 'undergoer voice', and 'locative voice', instead of traditional and conventional terms for voices, such as 'active voice'.

Despite the multiplicity of voice alternations, for simplicity, we begin by assuming that WISP languages have two major voice alternations that comprise actor and undergoer voices, which correspond approximately to active and passive voices in a traditional account, respectively. An actor voice form assigns an actor argument to a subject[14], and an undergoer voice form assigns an undergoer argument to a subject. The Proto-Malayo-Polynesian language distinguishes three undergoer voice forms: patient, locative, and thematic, which encode patient, location, and theme (including instrument, comitative, and beneficiary) arguments as a subject, respectively (Ross 2002). The actor voice is typically marked with the infix *<um> or the initial segment *m-*, as in *maR- and *maN-. On the other hand, the patient, locative, and thematic voices are marked with *-en, *-an, and *i-, respectively.

Consider the following Tagalog examples. In all the examples, the arguments in question are placed in the Predicate-Actor-Patient-Location-Beneficiary order. The voice-marking of the predicate specifies the semantic role of the subject, and the nominative marker *ang* introduces the following word as the role and argument specified by the predicate. For example, in the actor voice construction below, the prefix *mag-* determines the actor as the subject, and the nominative marker *ang* indicates that the following word *babae* 'woman' is the actor and the subject in the sentence. The same logic can be applied to the following examples in patient, locative, and thematic voice constructions. The patient voice marker *-in* in *a-alis-in* 'will be taken' specifies the patient as the subject, the locative voice marker *-an* in *a-alis-an* 'is the place where (a patient) will be taken' assigns the location *ang sako* 'bag (NOM)' to the subject, and the thematic voice marker *i-* in *i-pag-a-alis* 'is the target where (a patient) will be taken' specifies the beneficiary (*ang bata* 'child (NOM)') as the subject:

Tagalog
Actor voice

(4)	**Mag-**aalis[15]	**ang**	**babae**	ng	bigas	sa	sako	para	sa	bata.
	mag-CV-alis	ang	babae	nang	bigas	sa	sako	para	sa	bata
	AV-PROG-take.away	NOM	woman	GEN	rice	LOC	sack	for	OBL	child

'The woman will take some rice out of a/the sack for a/the child.'

(Schachter 1976:494, glossing modified)

14 The term 'subject' here refers to a syntactic pivot, following the terminology used by Himmelmann (2002).

15 Originally written as 'mag-salis'; however, this is apparently a typographical error because according to English (1986:36–38), it is *alis* that means 'take away; leave', and its reduplicated form in question must be *mag-a-alis*.

Patient voice

(5)	Aalis**in**	ng	babae	**ang**	**bigas**	sa	sako	para	sa	bata.
	CV-alis-in	nang	babae	ang	bigas	sa	sako	para	sa	bata.
	PROG-take.way-PV	GEN	woman	NOM	rice	LOC	sack	for	OBL	child

'A/the woman will take the rice out of a/the sack for a/the child.'
['The rice will be taken out of a/the sack for a/the child by a/the woman.']
(Schachter 1976:495, glossing modified)

Locative voice

(6)	Aalis**an**	ng	babae	ng	bigas	**ang**	**sako**	para	sa	bata.
	CV-alis-an	nang	babae	nang	bigas	ang	sako	para	sa	bata
	PROG-take.way-LV	GEN	woman	GEN	rice	NOM	sack	for	OBL	child

'A/the woman will take some rice out of a/the sack for a/the child.'
['The sack will be the place out of which some rice is taken for a/the child by a/the woman.']
(Schachter 1976:495, glossing modified)

Thematic voice

(7)	**I**pag-aalis[16]	ng	babae	ng	bigas	sa	sako	**ang**	**bata**.
	i-pag-CV-alis	nang	babae	nang	bigas	sa	sako	ang	bata
	TV-AD-PROG-take.away	GEN	woman	GEN	rice	LOC	sack	NOM	child

'A/the woman will take some rice out of a/the sack for the child.'
['The child will be for which some rice is taken out of a/the sack by a/the woman.']
(Schachter 1976:495, glossing modified)

On the other hand, Sulawesi and Western Indonesian languages in general have reduced voice systems consisting of two or three voice forms. For example, Malay has two voices, namely, actor and undergoer voices. The actor voice is marked with actor voice prefixes, such as *meN-* and *ber-*, and the undergoer voice is marked with the undergoer voice prefix *di-*, as in the following examples:

Malay
Actor voice with *meN-*

(8)	Dia	**men**jemput	saya.	
	dia	meN-jemput	saya	
	3.SG	AV-meet	1.SG	

'He met me.'
(Sneddon 1996:247, my glossing)

Undergoer voice with *di-*

(9)	Saya	**di**jemput	oleh	dia.
	saya	di-jemput	oleh	dia
	1.SG	UV-meet	by	3.SG

'I was met by him.'
(Sneddon 1996:248, my glossing)

In addition, some Sulawesi and Western Indonesian languages have so-called 'conjugated verbs'

16 Originally written as 'ipag-salis'; however, this is apparently a typographical error for the same reason as *mag-a-alis*. The reduplicated form in question must be *i-pag-a-alis.*

(Dutch *vervoegde vormen* 'conjugated forms')[17], which refer to action-denoting forms with procliticized or encliticized pronouns. The voices of these forms are sometimes difficult to identify via the forms themselves. Thus, they are often identified in relation to other voice forms with overt voice markers.

The conjugated pronominal forms have two formal types 1) those with procliticized or fronted pronouns and 2) those with encliticized pronouns. We term the former 'fronted pronominal forms' and the latter 'encliticized pronominal forms'. Of these forms, however, the former type is of primary concern from a Philippinist's perspective because arguments do not precede predicates in the majority of Philippine languages unless a negation marker, an auxiliary word, or an inversion marker is present.

Fronted pronominal forms consist of fronted, often procliticized, pronouns and action-denoting words. The action-denoting words often lack overt voice markers. The semantic role of the fronted pronouns is typically an actor, as in the following examples:

Malay

(10) Buku ini tidak akan **kami** baca.
book this not will 1.PE read
'This book will not be read by us.'

(Sneddon 1996:249, my glossing)

(11) Buku sejarah ini belum **dia** baca.
book history this not.yet 3.SG read
'He hasn't read this history book yet.'

(Sneddon 1996:250, my glossing)

(12) Buku itu sudah **ku=**baca.
book that already 1.SG.AC=read
'That book, (is what) I have already read.'

(Sneddon 1996:249, my glossing)

Uma (Celebic, Kaili-Pamona)

(13) **Ku=**oli' onse etu.
1.SG.AC=buy rice that
'I bought that rice.'

(Martens 1988:170, glossing modified)

However, it is also observed that the role is an undergoer, particularly a patient or a theme, in a few known constructions in Sulawesi languages. For example, in the following Wolio sentence, the procliticized pronoun is an undergoer:

Wolio (Celebic, Wotu-Wolio)

(14) **A=**to-bawa i banua.
3.UG=UV-bring LOC house
'They are brought to the house.' or
'S/he was brought to the house.'

(van den Berg 1996:104, glossing modified)

On the other hand, encliticized pronominal forms refer to predicates that consist of encliticized

17 The English term 'conjugated verb' is taken from van den Berg (1996:86).

pronouns and action-denoting words. In contrast to the fronted pronominal constructions, some encliticized pronominal forms contain overt voice markers; however, forms without overt voice markers are also observed. The semantic role of the pronouns in these constructions is primarily an actor, as in the following examples:

Malay

(15) Buku sejarah ini belum di-baca=**nya**.
book history this not.yet UV-read=3.SG.GEN
'He hasn't read this history book yet.'

(Sneddon 1996:250, my glossing)

Da'a (Celebic, Kaili-Pamona)

(16) Ni-oli'=**ku** ose etu.
UV.RLS-buy=1.SG.AC rice that
'The rice is bought by me'

(Martens 1988:170, glossing modified)

Uma (Celebic, Kaili-Pamona)

(17) Mo-keno-**a**.
AV-run-1.SG.AC
'I run.'

(Martens 1988:169, glossing modified)

However, the role of the pronouns can also be an undergoer, particularly a patient or a theme, in some constructions:

Uma (Celebic, Kaili-Pamona)

(18) Nu=weba'=**a**.
2.SG.AC=hit=1.SG.UG
'You hit me.'

(Martens 1988:169, glossing modified)

Tolaki (Celebic, Bungku-Tolaki)

(19) Ku=wutiwuti=**ko**.
1.SG.AC=deceive=2.SG.UG
'I deceived you.'

(Mead 2002:157, glossing modified)

From a Philippinist's perspective, these encliticized pronominal forms would not be as remarkable as the fronted pronominal forms because pronouns are placed after predicates in Philippine languages (cf. Zorc 1977). However, we employ the encliticized pronominal forms as analyzed because we do not have any phonological data to argue against the validity of the enclitic analyses of the pronouns.

The origins of the fronted and encliticized pronominal forms have not been explored thoroughly, although some authors, such as Wolff (1996), van den Berg (1996), Himmelmann (1996), and Mead (2002), have attempted to provide some suggestions for their origins.

1.4.2.2 Applicatives

In Austronesian linguistics, the term applicative (abbreviated as APPL) is specifically employed to refer to the two Post-Proto-Malayo-Polynesian suffixes *-i and *-kan, or the cognates or formal and functional equivalents of the Malay suffixes *-i* and *-kan* (cf. Ross 2002). These suffixes are found in Western Indonesian and Sulawesi languages and may appear with any voice forms. The reflexes of these suffixes have varied functions in WISP languages: However, in general, we can state that the suffix *-i introduces a locative argument and the suffix *-kan adds an instrumental or benefactive argument. In actor voice constructions, the argument introduced by the applicative suffix is typically placed immediately after the derived word.

For example, the actor voice form *menulis* 'write' of the action-denoting root *tulis* 'write' in Malay takes a patient, such as *surat* 'letter', as an object:

Malay
Underived

(20)	Dia	menulis	surat.	
	dia	meN-tulis	surat	
	3.SG	AV-write	letter	

'He writes a letter.'
cf. *tulis surat* 'write a letter' (Echols and Shadily 1994:591)

(My sentence)

When the applicative suffix *-i* is attached to the word *menulis* 'write', the object is a location upon which the action of writing is performed:

Malay
Locative *-i*

(21)	Dia	menulis**i**	kertas	itu.	
	dia	meN-tulis-i	kertas	itu	
	3.SG	AV-tulis-I	paper	that	

'He writes on that paper.'
cf. *menulisi (kertas)* 'write on paper'(Sneddon 1996:89)

(My sentence)

On the other hand, when the applicative suffix *-kan* is attached to the word *menulis* 'write', the object is either an instrumental or a beneficiary:

Malay
Instrumental *-kan*

(22)	Dia	menulis**kan**	pensil	tumpul.	
	Dia	meN-tulis-kan	pensil	tumpul	
	3.SG	AV-write-KAN	pen	blunt	

'He writes with a blunt pen.'

(Sneddon 1996:83, my glossing)

Benefactive *-kan*

(23)	Dia	menulis**kan**	ayahnya	surat.	
	dia	meN-tulis-kan	ayah=nya	surat	
	3.SG	AV-write-KAN	father=3.SG.GEN	letter	

'He writes a letter for his father'

(Sneddon 1996:81, my glossing)

The applicative suffix *-kan* also adds an object that indicates a displaced theme, as in the following example:

Displaced theme indicated by *-kan*

(24)	Dia	menulis**kan**	pikirannya	di	kertas	itu.
	dia	meN-tulis-kan	pikiran=nya	di	kertas	itu
	3.SG	AV-write-KAN	thought=3.SG.GEN	LOC	paper	that

'He wrote down his thoughts on that paper.'
cf. *menuliskan pikirannya di kertas* 'write down o[ne]'s thoughts on paper' (Echols and Shadily 1994:591)

(My sentence)

However, the applicative suffix *-kan* of this use is described as being 'optional' in Sneddon (1996) and can be omitted:

Malay

(25)	Utomo	mengantar(**kan**)	ibunya	ke	pasar.
	Utomo	meN-antar-kan	ibu=nya	ke	pasar
	Utomo	AV-accompany-KAN	mother=3.SG.GEN	to	market

'Utomo accompanied his mother to the market.'

(Sneddon 1996:84, my glossing)

In undergoer voice constructions, the subject has the semantic role that the applicative suffix specifies. For example, the applicativized word *menyerahi* takes the recipient *kami* 1.PE' as the primary object in the following sentence in an actor voice construction:

Malay

(26)	Kepala	kantor	menyerah**i**	kami	tugas	itu.
	kepala	kantor	meN-serah-i	kami	tugas	itu
	head	office	AV-hand-I	1.PE	task	that

'The office head handed us that task.'

(Sneddon 1996:251, my glossing)

By contrast, in an undergoer voice construction with the word *diserahi* with the applicative suffix *-i*, the subject is *kami* '1.PE':

Malay

(27)	Kami	diserahi	kepala	kantor	tugas	itu.
	kami	di-serah-**i**	kepala	kantor	tugas	itu
	1.PE	UV-hand-I	head	office	task	that

'We were handed that task by the office head.'

(Sneddon 1996:251, my glossing)

An undergoer voice construction with *-kan* takes a subject of a different role from those with -i. For example, the applicativized word *menyerahkan* takes the theme argument *tugas itu* 'that task' as the primary object in an actor voice construction:

Malay

(28)	Kepala	kantor	menyerah**kan**	tugas	itu	kepada	kami.
	kepala	kantor	meN-serah-kan	tugas	itu	kepada	kami
	head	office	AV-hand-KAN	task	that	to	1.PE

'The office head handed that task to us.'

(Sneddon 1996:251, my glossing)

In a similar manner to the applicative construction with the word *diserahi*, the subject in an undergoer voice construction with the applicativized word *diserahkan* 'was handed' is the argument *tugas itu* 'that task':

Malay

(29)	Tugas	itu	diserah**kan**	oleh	kepala	kantor	kepada	kami.
	tugas	itu	di-serah-kan	oleh	kepala	kantor	kepada	kami
	task	that	UV-HAND-KAN	by	head	office	to	1.PE

'That task was handed by the office head to us.'

(Sneddon 1996:252, my glossing)

1.4.2.3 Moods and aspects

In this section, we discuss moods and aspects that are marked with reflexes of Proto-Malayo-Polynesian mood and aspect forms. According to Ross (2002), three indicative moods (neutral, perfective, and imperfective) and two non-indicative moods (projective and atemporal) can be identified. In the present study, for convenience, we regard his 'imperfective' mood marker as a progressive aspect marker because it denotes ongoing actions and may also appear with the 'perfective' marker *<in> to indicate ongoing actions that are occurring in reality. In addition, his terms 'perfective' and 'projective' are termed 'realis' and 'subjunctive', respectively.

The indicative mood forms are our central concern because they are found most frequently in Philippine languages. The forms focused upon here are summarized in the following table:

Table 6: Indicative inflectional forms in Proto-Malayo-Polynesian (cf. Ross 2002 and Zorc 1977)

	Actor voice	Patient voice	Locative voice	Thematic voice
Neutral	*<um>	*-en	*-an	*i-
Realis	*<um><in>	*<in>	*<in> -an	*i-<in>
Neutral progressive	*<um>CV-	*CV- -en	*CV- -an	*i-CV-
Realis progressive	*<um><in>CV-	*<in>CV-	*<in>CV- -an	*i-<in>CV-

It should be noted that the functions of the mood- and aspect-marked forms vary from language to language due to reorganizations of the forms in question. In this section, however, we introduce the functions in Tagalog because it is one of the more well-known and -studied WISP languages.

A neutral form is a form without any mood and aspect specifications. In actor voice, it is marked with the form that reflects *<um>. When the root takes the lexical class prefix *paR-, the actor voice form is *maR-, which originates from *p<um>aR-. This form generally indicates that the action has not yet begun (cf. Reid 1992). In Tagalog, the neutral form is used in an imperative sentence:

Tagalog

(30)	**Mag**luto	ka	nga	ng	kape.
	mag-luto	ka	nga	nang	kape
	AV-cook	2.SG.NOM	please	GEN	coffee

'Please make some coffee!'

(Schachter and Otanes 1972:47, my glossing)

The realis mood is the mood that is marked with *<in>. This infix appears as *naR- when it appears with actor voice infix *<um> and the action-denoting class prefix *paR-, being the result of the formal reduction of *p<um><in>aR-. It generally indicates that the action has already begun (cf. Reid 1992). In Tagalog, this mood form denotes 1) past tense, 2) present perfect mood, and 3) past perfect mood:

Tagalog

Past

(31)	**Nag**luto	ng	pagkain	ang	nanay.
	nag-luto	nang	pagkain	ang	nanay
	AV.RLS-cook	GEN	food	NOM	mother

'Mother cooked some food.'

(Schachter and Otanes 1972:67, my glossing)

Present perfect

(32)	**Nag**luto	na	ng	pagkain	ang	nanay.
	nag-luto	na	nang	pagkain	ang	nanay
	AV.RLS-cook	already	GEN	food	NOM	mother

'Mother has cooked some food.'

(Schachter and Otanes 1972:67, my glossing)

Past perfect

(33)	**Nag**luto	na	ng	pagkain	ang	nanay [when I arrived].[18]
	nag-luto	na	nang	pagkain	ang	nanay
	AV.RLS-cook	already	GEN	food	NOM	mother

'Mother had cooked some food (when I arrived).'

(Schachter and Otanes 1972:67, my glossing)

The progressive aspect is typically marked with partial root reduplication in Philippine languages. In Tagalog, the progressive reduplicative form is CVV /CV:/ with vowel lengthening; however, for simplicity, it is conventionally transcribed as CV-, unless vowel length distinction is otherwise specified. When it appears with neutral mood forms, it may indicate future progressive events:

Tagalog

Future (progressive)

(34)	**Maglu**luto	ng	pagkain	ang	nanay	bukas.
	mag-CV-luto	nang	pagkain	ang	nanay	bukas
	AV-PROG-cook	GEN	food	NOM	mother	tomorrow

'Mother will cook (etc.) some food tomorrow.'

(Schachter and Otanes 1972:67, my glossing)

18 Tagalog *noong dumating ako* 'when I arrived' (when arrived I). It is, however, omitted here for simplicity.

When the progressive-indicating reduplication appears with a realis mood form, the inflected form typically encodes the present progressive and past progressive actions:

Tagalog
Present Progressive

(35)	**Naglu**luto	(na)	ng	pagkain	ang	nanay.
	nag-CV-luto	na	nang	pagkain	ang	nanay
	AV.RLS-PROG-cook	already	GEN	food	NOM	mother

'Mother is cooking some food (now).'

(Schachter and Otanes 1972:67, my glossing)

Past Progressive

(36)	**Naglu**luto	ng	pagkain	ang	nanay [when I arrived].
	nag-CV-luto	nang	pagkain	ang	nanay
	AV.RLS-PROG-cook	GEN	food	NOM	mother

'Mother was cooking some food (when I arrived).'

(Schachter and Otanes 1972:67, my glossing)

It should be noted, however, that even within Philippine languages, the uses of these mood and aspect forms differ due to restructuring in their voice and mood paradigms. For example, Ilocano encodes present progressive actions with the neutral mood progressive aspect form, such as *<um>CVC-* in *t<um>ak-takdér* 'is standing' (Reid 1992:69–70).

In contrast to the Philippine languages, Western Indonesian and Sulawesi languages either 1) lack such mood and aspect distinctions or 2) have a reduced mood and aspect paradigm. For example, the four indicative mood and aspect distinctions are not present in Malay. As a result, Malay employs the historically neutral form *maR-, reflected as *ber-,* as a mood- and aspect-neutral form.

In addition to these indicative forms, Ross (2002) suggests two series of Proto-Malayo-Polynesian non-indicative moods, namely, atemporal and projective. The atemporal mood forms refer to the forms that may indicate one or more of the following three basic functions: 1) imperative, 2) verb subordinate to auxiliaries, and 3) expressing non-initial sequential events in narrative (Ross 2002:37). On the other hand, the projective forms refer to the forms that are 'evidently the finite form[s] used for irrealis events, i.e., intention, possibility and exhortation' (Ross 2002:37). It should be noted that both forms are renamed and slightly redefined in the Proto-Austronesian reconstruction in Ross (2009).

However, the WISP languages only have one series of non-indicative mood affixes. A brief examination of Marinduque Tagalog, Bisayan, Manoboic, and Gorontalo-Mongondowic languages suggests the following forms:

Table 7: Subjunctive markers in some Greater Central Philippine languages

	Actor voice	Patient voice	Locative voice	Thematic voice
Form	Ø	-a	-i, -ay	-an

These forms, except for *-ay*, correspond to the atemporal forms in the Proto-Austronesian and Proto-Malayo-Polynesian paradigms in Ross (2002). The form *-ay* can be suggested by the presence of the Mongondow locative voice imperative form *-ay* (data from Lobel 2013). The form appears to reflect Ross's projective form for the locative voice constructions; however, its function suggests

that it corresponds to the one denoted by the atemporal form *-i.

Two uses are commonly observed for these forms: imperative and subjunctive (dependent). The following examples in Aklanon (Bisayan) show the imperative use:

Aklanon
Actor voice, imperative

(37) Bakáɬ it gatas.
buy.AV.SUBJUNCTIVE GEN.INDEF milk
'Buy some milk.'

(Zorc 1968:69, my glossing)

Patient voice, imperative

(38) Bakɬ**á** ro kałamay.
bakaɬ-á ro kałamay
buy-PV.SUBJUNCTIVE NOM.DEF brown.sugar
'Buy that brown sugar.' ['May that brown sugar be bought.']

(Zorc 1968:68, my glossing)

Locative voice, imperative

(39) Łabh**i** ro sałwaɬ.
Łabah-i ro sałwaɬ.
wash-LV.SUBJUNCTIVE NOM.DEF trousers
'Wash these trousers!' ['May the(se) trousers be washed.']

(Zorc 1968:68, my glossing)

Thematic voice, imperative

(40) Dawat**án** ro asín.
dawat-án ro asín
pass-TV.SUBJUNCTIVE NOM.DEF salt
'Pass the salt.' ['May the salt be passed.']

(Zorc 1968:68, my glossing)

Another common use is referred to here as 'subjunctive'. It shows the dependency of a word to another preceding predicative word. Subjunctive forms can often be observed in negative constructions where the negative word occupies the word-initial position:

Aklanon
Past subjunctive

(41) Owá' nakon pagbakɬ**á** ro isda.
Owá' nakon pag-bakaɬ-a ro isda.
not 1.SG.GEN AD-buy-LV.SUBJUNCTIVE NOM fish
'I did not buy the fish.'
['The fish is not bought by me.']

(Zorc 1968:195, my glossing)

They are also attested after interrogatives:

Aklanon

(42)	Siqín	nímu	qig-bákɬ-**a**	ro	reló.
	where	2.SG.GEN	UV.RLS-buy-PV.SUBJUNCTIVE	NOM	watch

'Where did you buy the watch?'
[Where was the watch bought by you?]

(Zorc 1977:126, my glossing)

Some Western Indonesian languages also have non-indicative mood suffixes. For example, Salako has the subjunctive suffix *-àʔ*, and Old Javanese has the subjunctive suffix *-a*:

Salako (Malayic)

(43)	Aku	dah	baiʔ	ba-lawakng-**àʔ**	kau.
	1.SG	already	not.want	AV-marry-SUBJUNCTIVE	2.SG

'I don't feel like marrying you [anymore]!'

(Adelaar 2005c:52, glossing modified)

Old Javanese

(44)	Mahyun	lumakw**a**.
	mahyun	<um>laku-a
	wish	AV-go-SUBJUNCTIVE

'He wishes to go.'

(Kern 1902:175, my glossing)

1.4.3 Case-marking systems

This section concerns itself with case-marking functions of phrase markers in WISP languages. It will provide basic information concerning the case-marking (grammatical relation- and semantic role-marking) performed by the phrase markers. It will also introduce the terminology used for describing cases in the following chapters.

Phrase markers here refer to prepositions, such as Tagalog *ang* 'nominative', as in *ang babae* 'NOMINATIVE:woman (*babae*)', and *si* 'nominative for personal names', as in *si Juan* (NOMINATIVE:Juan). They encode cases, that is, grammatical relations and semantic roles of the following items, which are often object-denoting (i.e., nominal). In some languages, they encode specific features with regard to other distinctions, such as singular and plural, and presence and absence (or 'dead or alive') distinctions.

Phrase markers indicating grammatical relations are characteristic of Philippine languages. However, the forms and functions of the phrase markers vary from language to language, as in the following examples in Tagalog, Ilocano, Dupaningan Agta, and Central Cagayan Agta:

Tagalog

(45)	Ibinigay	**ng**	laláke	**ang**	libro
	TV.RLS:give	GEN	man	NOM	book
	sa	bátà	**sa**	paaralan.	
	OBL	child	LOC	school	

'The man gave the book to the child in school.'
['The book is given to the child by the man in school.']

(Reid and Liao 2004:466, glossing modified)

Ilocano (Northern Luzon)

(46) Nangan **ti** kabsat=ko **ti** innapoy
AV.RLS:eat CORE friend=1.SG.GEN CORE cooked rice

(i)ti balay **ti** kaarrúba=k.
NON-CORE house CORE neighbor=1.SG.GEN
'My friend ate rice at my neighbor's house.'

(Reid and Liao 2004:466, glossing modified)

Dupaningan Agta (Northern Luzon)

(47) I-demat=**na** [babbey]GEN [**i** kape]NOM [**ha** lalaki]OBL.
TV-bring=3.SG.GEN woman DEF coffee OBL man
'A woman will bring the coffee to a man.'

(Robinson 2008:81, glossing modified)

(48) Nag-hida **i** anak=aya **ha** padut.
AV.RLS-main.course DEF child=MEDIAL.SPECIFIC OBL fish
'That child ate fish.'

(Robinson 2008:98, glossing modified)

Central Cagayan Agta (Northern Luzon)

(49) P<in>aligat **na** abbing-en **ya** kábayuq-en.
PV:RLS-hit GEN child-that.GEN NOM horse-that.GEN
'That child hit that horse.'

(Healey 1960:34, glossing modified)

(50) Mag-gáni **ta** pagay **ta** bagetay.
AV-harvest OBL rice OBL hill
'He's harvesting rice up the hill.'

(Healey 1960:47, glossing modified)

Table 8: Case-marking by phrase markers in some WISP languages (Robinson 2008:81)

	Topic	Syntactic subject [actor in actor voice and undergoer in undergoer voice]	Actor of a non-AV verb [actor in undergoer voice]	Possessor	Semantic object of an actor voice verb [undergoer in actor voice]	Oblique [including locative]
Tagalog	**ang** Nominative		**ng /nang/** Genitive			**sa** Oblique ([Locative])
Ilocano	**ti** Core argument					**iti** Non-core
Dupaningan Agta	**Ø** [**i** (?), YK][19] Nominative		**na** (singular), **di** (plural) Genitive		**(ha)** Oblique	
Central Cagayan Agta	**ya** Nominative		**na** Genitive		**ta** Oblique	

In contrast to the Philippine languages, Western Indonesian and Sulawesi languages encode grammatical relations mostly by means of word order. For example, Toba Batak, a Western Indonesian language in North Sumatra, does not have case-markers for subjects and objects in most of the sentence types. It has the personal phrase marker *si*, which can be identified as a cognate of the Tagalog personal name nominative marker *si*. However, it does not have a morphosyntactic function; rather, it functions as a personal name marker in the language. The case-marking of the two types of arguments are marked through word order and voice markers. For example, the following Toba Batak sentences describe the same situation from different points of view:

Toba Batak (Sumatran, Batak)

(51) Mang-ida **si** Ria **si** Torus.
AV-see PN Ria PN Torus
'Torus sees/saw Ria.'
(Schachter 1984:123, quoted in Ross 2002:55, glossing modified)

(52) Di-ida **si** Torus **si** Ria.
PV-see PN Torus PN Ria
'Torus sees/saw Ria.'
(Schachter 1984:123, quoted in Ross 2002:55, glossing modified)

When the root *ida* 'see' appears with the actor voice marker *mang-*, the inflected word *mang-ida* 'see' assigns the sentence-final argument *si Torus* 'Torus (as a personal name)' to the subject role. However, when it takes the undergoer voice marker *di-*, the word *di-ida* 'be seen' specifies the sentence-final argument *si Ria* 'Ria (as a personal name)' as a subject.

As can be observed in the examples above, the majority of the phrase markers in WISP

19 Although Robinson treats the phrase-initial particle as a definite marker, the form and the function suggest that it may reflect the nominative phrase marker with *si (or a phrase marker that replaced the nominative phrase marker, if the restructuring of case-marking segments has occurred).

languages appear to be monosyllabic, and the cases are often distinguished by the initial consonants of the phrase markers. Such initial case-marking segments can also be observed in pronouns, demonstratives, and interrogatives.

The initial consonants are suggested at the Proto-Malayo-Polynesian or Proto-Austronesian level by a few authors. Based only on the initial consonants in phrase markers, for example, Blust (2015) suggests *s for nominative, *n for genitive, *d for locative, and *k for oblique:

Table 9: Case-marking segments assumed in the current study (cf. Blust 2015)

	Nominative	Genitive	Locative	Oblique
Form	*s	*n	*d	*k

To elucidate the functions of the case-marking segments, we employ this functional paradigm and its terminology. It should be noted that, in contrast to most of the Proto-Malayo-Polynesian forms for voice, mood, and aspect, these case-marking forms are not always reflected as having the reconstructed functions. For example, while we can find the nominative segment *s in the Tagalog third person singular nominative pronoun *siya*, what we can find in the Malay third person singular non-genitive pronoun *dia* is the locative segment *d, despite the fact that the pronoun serves as a subject in actor voice constructions in the same manner as the Tagalog pronoun.

The term ‘nominative case’ refers to the case that denotes the subject in a given sentence. The semantic role can be an actor or an undergoer, which is determined by the voice of the predicate in the sentence. In Tagalog, the nominative case is indicated by *ang*:

Tagalog
Nominative marker indicating an actor in an actor voice construction

(53) Nagtanong **ang** bata sa kapit-bahay.
nag-tanong ang bata sa kapit-bahay
AV.RLS-ask NOM child OBL neighbor
‘The child asked the neighbor.’

(De Guzman 2000:227, glossing modified)

Nominative marker indicating an undergoer in an undergoer voice construction

(54) Tinanong ng bata **ang** kapit-bahay.
<in>tanong nang bata ang kapit-bahay
PV:RLS-ask GEN child NOM neighbor
‘The child asked the neighbor.’
[‘The neighbor is asked by the child.’]

(De Guzman 2000:227, glossing modified)

The term ‘genitive case’ refers to the case that functions as 1) a postposed possessive marker, 2) an actor marker in an undergoer voice construction, and 3) a patient or theme marker in an actor voice construction with a predicate marked as an actor voice. Consider the following Tagalog examples with *=nya* / *niyá* ‘3.SG.GEN’, *ni* ‘PN.GEN’, and *nang* ‘GEN (for common nouns)’:

Tagalog

Genitive marker as a postposed possessive marker

(55) ang báhay=**nya**
NOM house=3.SG.GEN
'his house [(NOM)]'

(Bloomfield 1917:172, my glossing)

(56) sa báhay **ni** Pédro
LOC house PN.GEN Pedro
'Pedro's house [(LOC)]'

(Bloomfield 1917:230, my glossing)

(57) ang púno **nang** ságing
NOM tree GEN banana
'the banana tree [(NOM)]'

(Bloomfield 1917:172, my glossing)

Actor marker in an undergoer voice construction

(58) Ang alak ay ininóm **niyá**.
ang alak ay <in>inom niya
NOM wine INV PV:RLS-drink 3.SG.GEN
'The wine was drunk by him.'

(Lopez 1977a [1937]:62, my glossing)

(59) Ang alak ay ininóm ni Pedro.
ang alak ay <in>inom ni Pedro
NOM wine INV PV:RLS-drink PN.GEN Pedro
'The wine is drunk by Pedro.'

(Lopez 1977a [1937]:62, my glossing)

Undergoer marker in an actor voice construction

(60) Si Pedro'y uminóm **nang** alak.
si Pedro ay <um>inom nang alak
PN.NOM Pedro INV AV.RLS-drink GEN wine
'Pedro drank wine.'

(Lopez 1977a [1937]:62, my glossing)

The term 'locative case' refers to the case that indicates locations, as in Malay *di rumah* 'at home'[20] (LOC house) and Cebuano *Prísku sa báybay* 'It is refreshing at the beach' (refreshing LOC beach)[21]. The locative case marker may also be employed in temporal expressions, as in Cebuano *sa máyu* 'in May, next may'[22] (LOC May).

The term 'oblique case' refers to the cases that encode roles that are not marked by the nominative, genitive, and locative cases. It should be noted that the term is meant to refer to the residues of the roles that are indicated by the nominative, genitive, and locative markers mentioned above.

20 Data from Echols and Shadily (1994:143).
21 Data from Zorc (1977:93).
22 Data from Zorc (1977:96).

An oblique marker may also serve as an accusative marker to indicate an undergoer (patient or theme) role in an actor voice construction in some sets of phrase markers in some languages, in the same manner as a genitive marker in some actor voice constructions. For example, Tagalog employs the personal name oblique marker *kay* to encode the undergoer in an actor voice construction, as in the following example:

Tagalog

(61)	Nagtanong	si	Pedro	**kay**/*ni	Juan.
	nag-tanong	si	Pedro	**kay**/*ni	Juan
	AV.RLS-ask	PN.NOM	Pedro	PN.OBL/*PN.GEN	Juan
	'Pedro asked Juan.'				

(De Guzman 2000:227, glossing modified.
The name Pablo is replaced with Juan for clarity.)

It also indicates a variety of semantic roles, such as recipient (e.g., Tagalog *kay* 'dative for personal nouns'), instrument (e.g., Cebuano *sa súndag* 'with a bolo'[23]; *sa* is an oblique and locative marker for common nouns), comitative (Limos Kalinga (Philippine, Northern Luzon) *kan Helena* 'with Helena'[24]), beneficiary (e.g., Cebuano *pára kang lúsi* 'for Lucy'[25]), time (e.g., Cebuano *kani-qádtu* 'long time ago, once upon a time'[26], source (ablative) (e.g., *kan-ámuq* 'from us'[27]), and direction (allative) (e.g., *kang tatay* 'to father (*tatay*)'[28]).

In addition to these case-marking segments, it should be noted that Philippine languages have a particle that has been termed 'linker' (or 'ligature'). The linker often appears in the shape *-ng, *nga, or *na (cf. Zorc 1977) and is placed between a head and a dependent. For example, Tagalog linkers are *-ng* and *na*, both of which are formally different from the genitive common noun marker *nang*. The preferred order of the head and the dependent is dependent-head in Bisayan (Zorc 1977:91), as well as in Tagalog (cf. Bloomfield 1917:163); however, the head-dependent order is also acceptable in some constructions with a linker, as in Tagalog *ang mabúti-ng aklàt* 'the (a) good book' (good LK book; also *ang aklàt na mabúti*)[29], and Cebuano *babáyi-ng maningísdaq* 'lady fisherman (woman LK fisherman; also *mangingísdaq nga babáyi)*[30].

In addition to the terms for cases, we employ two terms that directly refer to macro-roles, namely, actor (AC) and undergoer (UG) when the current data are not sufficient to identify their cases in connection with Philippine case-marking systems. This treatment is primarily performed on some data for a few Sulawesi languages.

Finally, it should be noted that exceptionally we use the terms 'ergative' and 'absolutive' for Makassarese data quoted in Chapter 3, following the terminology by Jukes (2006).

23 Data from Zorc (1977:89).
24 Data from Ferreirinho (1993:13).
25 Data from Zorc (1977:90).
26 Data from Zorc (1977:97).
27 Data from Zorc (1977:89).
28 Data from Zorc (1977:90).
29 Bloomfield (1917:163).
30 Zorc (1977:91).

1.5 Conventions

The conventions used in the current study are as follows:

Phonological

N	homorganic nasal with nasal assimilation and nasal substitution
N#	homorganic nasal with nasal assimilation but no nasal substitution
R	the Proto-Malayo-Polynesian and Proto-Austronesian phoneme *R, which was probably an alveolar or uvular trill (Blust 2013:554)
C	consonant
V	vowel
$	syllable boundary

Morphological

X-	X is a prefix.
-X	X is a suffix.
X=	X is a proclitic.
=X	X is an enclitic.
<X>	X is an infix.

It should be noted that pronouns that attach to inflected forms are processed as clitics to distinguish them from other affixes for orthographic convenience.

It should also be noted that unless otherwise mentioned, examples of lexical items and PMP forms employed here are from Blust and Trussel (ongoing). However, we omit the final proto-phoneme *x[31], as observed in *amax ‘father’ because we do not know what it represents.

Orthographical

Various spellings are used in grammars and previous studies concerning the WISP languages. Unless otherwise indicated, the orthography employed in the sources is maintained in the current study. However, when the source has an orthography or a character that is difficult to distinguish or process, some modifications are implemented based on the following rules:

- A schwa (/ə/) is written as <e>.
- The phoneme /e/ is written as <é> when the distinction from a schwa is relevant.
- The phoneme /u/ is written as <u> (instead of <oe> in some Dutch sources).
- A velar nasal /ŋ/ is written as <ng>.
- The phoneme /j/ is written as <y> (instead of <j>).

Orthographic changes are made in the following languages: Tagalog <ng>, Sangir <e, é, u>, Mongondow <u, y>, Kiput <ng>, Malay <ng>, and Salako <ng>. The velar stops in Blust and Trussel (ongoing) are represented here as <ng>. The Aklanon <e> representing a voiced velar fricative in Zorc (1968) is converted to <ɬ>, following the orthography employed in Zorc (1977).

31 The final *x was first added in the revision on 18 June 2017, as far as we can tell. It appeared as a contrastive segment between two phonologically (and semantically) similar roots, such as the pairs *batux ‘stone’ and *batu ‘to throw’, and *sikux ‘elbow’ and *siku ‘tail, end of something long’.

The letter <e> in Proto-Austronesian and Proto-Malayo-Polynesian forms represents schwas.

It should be noted that no change has been made on the Cotabato Manobo orthography by Kerr (1988), in which the letter <e> represents the phoneme /e/, despite the presence of the schwa (transcribed as <é>). This treatment is because schwas are rarer than mid-front unrounded vowels in our data on this language.

No change has also been made to the representation of glottal stops when they are transcribed as an apostrophe <'> (e.g., Pamona *kina'a* /kinaʔa/ 'rice'), a diacritical mark (e.g., Tagalog *batâ* /batáʔ/ 'child'), the letter <q> (e.g., Lampung *sakiq* /sakiʔ/ 'sick'), the letter <k> (e.g., Malay *anak* /anaʔ/), or the glottal stop symbol <ʔ> (e.g., Salako *ba-sing-koʔopm* 'to put entirely into one's mouth'. However, we convert the representation of the glottal stops in Sangir and Mongondow to the letter <q> because the two languages use none of these transcriptions.[32]

On a related note, a word-initial glottal stop, such as in Tagalog *alis* [ʔális] 'to leave' is not transcribed unless it is phonemic and transcribed in the original literature. Likewise, syllable-initial glottal stops in lexical items, as in Tagalog *daan* [da.ʔan] 'path, road'[33] and those in a morpheme boundary, as in Tagalog *pang-aral* [paŋʔáral] 'preaching, counsel'[34] are not transcribed unless they are also transcribed in the original.

Stress and vowel length, which are often indicated with diacritics (as in Tagalog), are not indicated unless they are marked in the original or are contrastive between underived and derived words.

It should be noted that due to the orthographic differences among the previous studies, some data (e.g., Tagalog) exhibit different spelling conventions regarding glottal stops, stress, and vowel length.

The initial phoneme in a sentence-initial word is capitalized unless it is noted that it is a clause preceded by another clause.

In Tagalog, the standard orthography employs the spellings *ng* for *nang* 'GEN' and *mga* for manga 'PLURAL'. These spellings are not changed.

About cited materials and glossing rules

For simplicity, we omit any accents in language names. For example, Nêlêmwa is written as Nelemwa.

The phrase 'data from', as in 'Data from Adelaar (1984:418)', refers to the fact that we only cite the data, not the author's arguments concerning the data (if they are present).

The modifications indicated by the phrase 'glossing modified' primarily refer to the changes of glossing for 1) voice markers and 2) phrase markers.

The phrase 'spelling modified' refers to the modifications of 1) orthographies that are not addressed in this section, or those of 2) omission of functional glossing, such as [ ]$_{o,}$ in Tinrin (Oceanic) [toni]$_o$ 'Tony [as an object]', as described in Evans (2003:244). However, we do not mention the orthographic changes between a hyphen and an equality sign (=), which are employed here for indicating a morpheme boundary and a boundary between a clitic and a word. For example, the spelling *ku-oli'* in the Uma sentence *Ku-oli' onse etu* 'I bought that rice'[35] is changed to *ku=oli'*,

32 A glottal stop is represented by a dot below the preceding vowel in the Sangir orthography by Adriani (1893) and by the Arabic letter hamza in the Mongondow orthography by Dunnebier (1929).

33 Data from Blust (2013:180).

34 Data from Blust (2013:180).

35 Data from Martens (1988:170).

as in *Ku=oli' onse etu*. Likewise, for instance, the spelling *eg=se=ikagi=yay* 'are speaking to each other' in Cotabato Manobo is altered to *eg-se-ikagi-yay* with hyphens in the current study. Readers who are interested in the original glossing are advised to consult the sources.

The phrase 'my glossing' refers to the fact that the glossing at issue is conducted by the current author by reference to other parts of the same literature, as well as her Austronesian grammatical and etymological knowledge.

Unless otherwise semantically distinctive with other forms with other lexical class prefixes, we omit the glossing of lexical class prefixes in actor voice prefixes. For example, the Tagalog actor voice *paR- class prefix *mag-* is glossed as 'AV (actor voice)'.

1.6 Organization of the remainder of the current study

The organization of the following chapters is as follows: Chapter 2 introduces and discusses the situations and constructions relevant to the discussion of the prefix *si-. Chapter 3 deals with formations with the prefix *si- with action-denoting (i.e., verbal) roots and their functions. Chapter 4 introduces possible uses of *si- forms with other types of roots and words and their functions. Chapter 5 examines the possible cognates of *si- forms in Formosan and Oceanic languages. In particular, it discusses the relations between 1) Kavalan *sim-* and WISP *si- forms, and 2) reflexes of the Proto-Oceanic form *paRi- and WISP *si- forms. Finally, Chapter 6 presents concluding remarks.

2 Situation types and their constructions

The primary objective of the current chapter is to present the definitions of the categories (i.e., situations and constructions) used in the investigations of *si- forms in the WISP languages elaborated upon in the following chapters. The categories that are addressed here include reciprocal, actor plural, actor plural, and comitative. For convenience, we begin our discussion with the reciprocal category due to the rich accumulation of relevant typological studies.

It has been observed in different languages that markers encoding reciprocal situations can also encode various other situations, such as reflexive, actor plural, and iterative ones (Kemmer 1993; Lichtenberk 2000; Nedjalkov 2007a, 2007b). Similarly, it is evidenced in typological studies that some reciprocal markers are employed to form various non-verbal expressions, including nouns meaning 'comrade, friend' (König and Gast 2008:9), nominal dual constructions (e.g., 'two children' with the root 'child'), converse kinship pair constructions (e.g., 'a child and a parent' with the root 'child'), and spatial adverbs (e.g., 'facing each other' with the root 'front') (Nedjalkov 2007b).

Several attempts have been made to explain such polysemies of reciprocal markers. For example, some authors have suggested a variety of shared synchronic semantic features, with which they explain the reciprocal polysemy. Kemmer (1993) claims that the 'low elaboration of events' is a common feature among some reciprocal, actor plural, and reflexive constructions. Lichtenberk (2000), on the other hand, argues that the synchronic semantic common features among most of the various categories marked by the Oceanic reciprocal marker *paRi- are 'plurality of relations' and 'low degree of elaboration'. However, neither Kemmer nor Lichtenberk explores the diachronic relations that exist among these categories by using these common features.

Other authors have suggested various cross-linguistic generalizations regarding the directions in which reciprocal and associated verbal categories develop. For instance, Heine and Miyashita (2008) claim that some reciprocal markers are developed from reflexive markers, but that reflexive markers do not develop into reciprocal markers. Moyse-Faurie (2008), on the other hand, challenges this argument, stating that the development pathway from a reflexive marker to a reciprocal marker is plausible. Similarly, Nedjalkov (2007b:248) claims that reciprocals can evolve from reflexives, and develop into actor plurals (his 'sociatives').

As such, Heine and Miyashita (2008), Moyse-Faurie (2008), and Nedjalkov (2007b) have different views on these diachronic changes. If we suppose that the variation in their arguments lies in the differences of their sample languages, the first requirement for our exploration of the historical changes of the categories marked by a given morpheme is a hypothesis regarding how the situations encoded by the categories are related using the data in our own sample languages.

For example, the authors appear to presuppose that some of the categories in question have a single source in a language or a linguistic family or subgroup. However, our WISP data suggest that this assumption is not always valid because it is observed that the prefix *si- occurs in a variety of formations. While all the forms have the common segment *si-, they differ with regard to affixes that appear with the segment *si- according to their meanings. For instance, some functions tend to be encoded with the formation including the prefix *maR-, and other functions are expressed with the formation including the prefix *ka-, as in the following table:

Table 10: Some *si- formations and their functions[36]

	*si- with *maR- (e.g., *maR-si-)	*si- with *ka- (e.g., *si-ka-)	*si-
Action-denoting roots	Reciprocal, actor plural, reflexive	N/A	Reciprocal, actor plural, reflexive
Quality-denoting roots	Equal quality of two entities (e.g., '(A and B) are beautiful.' with the root 'beautiful')	A singular entity with the same quality (as someone else) (e.g., '(A) is as beautiful as (B).' with the root 'beautiful')	A singular entity of the same quality as someone else
Kinship terms	Converse kinship pair (e.g., 'father and child' with the root 'father')	N/A	Converse kinship pair

Among such *si- forms, we designate the form *maR-si- as the starting point for our discussion of the conceptual apparatus that will be employed in this work because it is one of the forms that are observed frequently across WISP languages. The form *maR-si- is employed to indicate 1) reciprocal, 2) actor plural, and 3) reflexive situations. These situations can be defined briefly as follows: A reciprocal situation refers to a situation where plural actors act on each other, an actor plural situation refers to a situation where plural actors do not act on each other, and a reflexive situation refers to a situation where an actor performs an action on himself/herself. Of these categories, we consider the reciprocal category to be the basis of our discussion because it is one of the most commonly attested categories expressed by *si- formations, and the cross-linguistic studies in reciprocal constructions are many and can easily be obtained.

The semantic relationships of these three categories can be established when we assume a paradigmatic relationship among their semantic relationships with the following two semantic features: 1) the plurality of actors and 2) the actor-cum-undergoer[37] role. The plurality of actors refers to the state-of-affairs where the participants having an actor role in a situation are plural. It is a feature shared by reciprocal and actor plural situations. On the other hand, the actor-cum-undergoer role refers to the role that encodes both the roles of the actor and the undergoer in a situation. It is a feature shared by reciprocal and reflexive situations.[38] These two parameters enable us to analyze reciprocal, actor plural, and reflexive situations with the following features:

(62) Reciprocal situation: [+ actor plural], [+ actor-cum-undergoer]
Actor plural situation: [+ actor plural], [– actor-cum-undergoer]
Reflexive situation: [– actor plural], [+ actor-cum-undergoer]

Furthermore, we employ a syntactic parameter concerning the argument realization of the plural participants of the same macro-role, to identify whether the participants are encoded in a single argument expression or two (discontinuous) argument expressions. For instance, we consider actor plural and comitative actor constructions to be different with regard to their argument structures. While in the former, the plural actors are encoded as one argument (i.e., actor subject), as in *John*

36 It should be noted that the asterisked forms in it only indicate the immediate proto-forms of the prefix combinations attested in modern languages.

37 The term 'actor-cum-undergoer' is taken from Rapold (2011:69).

38 It should be noted that, for simplicity, the situations in which plural actors perform each action to themselves, such as *John and Bill wash their hands individually* are ignored in the current study because they would play minor roles regarding the set of the reciprocal, actor plural, and reflexive situations.

and Bill run together, in the latter, they are encoded in an actor subject and a comitative argument expression, as in *John runs together with Bill,* with the comitative marker 'with'.

The same logic can be applied to constructions that encode reciprocal situations. In order to distinguish the two argument realizations for the situations, the current study employs the terms 'reciprocal construction' and 'reciprocal comitative actor construction' for describing the constructions that encode reciprocal situations. The term 'reciprocal construction' refers only to the reciprocal construction with an actor argument, as in the French[39] sentence *Pierre et Jean se battent* '[Pierre and Jean] are fighting'[40] or 'Pierre and Jean fight each other' [(Pierre and Jean REC fight.3.PL.PRESENT)]. On the other hand, the term 'reciprocal comitative actor construction' refers to the one with a subject encoding a singular actor and an argument encoding a comitative actor, as in the French sentence *Pierre se bat avec Jean.* 'Pierre fights (with) Jean.' [(Pierre REC fight.3.SG.PRESENT with Jean)].

If we take these two types of the argument realizations into account, reciprocal, actor plural, reflexive, comitative actor, and reciprocal comitative actor constructions can be analyzed to have the following semantic and syntactic features:

(63) Reciprocal construction: [+ actor plural], [+ actor-cum-undergoer], [one argument]
Actor plural construction: [+ actor plural], [– actor-cum-undergoer], [one argument]
Reflexive construction: [– actor plural], [+ actor-cum-undergoer], [one argument]
Comitative actor construction: [+ actor plural], [– actor-cum-undergoer], [two arguments]
Reciprocal comitative actor construction: [+ actor plural], [+ actor-cum-undergoer], [two arguments]

The remainder of this chapter details the definitions and encoding patterns of the situation types that have been discussed above, and the other categories that are observed as being related to the reciprocal constructions and relevant to our WISP discussion.

2.1 Reciprocal situations and constructions

A reciprocal situation is defined here as a situation where there is more than one actor, who is also an undergoer for other actors, as in the situation described in the sentence *John and Bill hit each other.* The role of the plural actors in a reciprocal situation is termed an 'actor-cum-undergoer'. In this study, a reciprocal construction refers specifically to a construction that encodes a reciprocal situation, with the actors encoded in subject function.[41] It should be noted that, as mentioned earlier, when actors in a reciprocal situation are encoded in two discontinuous arguments that are a subject and a non-subject, the construction encoding the reciprocal situation is treated as a subtype of a comitative construction and is termed 'reciprocal comitative actor construction'.

According to the lexical properties of roots, the encoding of reciprocal situations can be classified into two types. If a given root does not denote that the action indicated by the root is performed by plural actors in a reciprocal manner, both the actor plurality as well as the actor-cum-undergoer role should be indicated by functional items. For example, the English word '*hit*' does

39 The French examples in this paragraph are cited from Guentchéva and Rivière (2007:565).

40 This sentence is originally translated as 'Jean and Pierre are fighting.' in Guentchéva and Rivière (2007:565).

41 It should be noted that except for the canonical type of reciprocal situations, with one reciprocal actor argument as a subject, there are also some semantic variations that may be termed reciprocal situations, which differ in, for example, the number of participants, saturation types, symmetry, and simultaneity (cf. Majid et al. 2011). Such differences, however, are beyond the scope of the current study.

not necessarily involve two actors who hit each other; it can encode the action of the situation *John hit a ball in the baseball game*, or *John hit Bill (but Bill did not hit John back.)*. Hence, the sentence requires the reciprocal indicator 'each other' to indicate the reciprocal situation of hitting.

However, some roots contain the semantic features of reciprocal situations. Such roots are here termed 'lexical reciprocal' roots. These roots may or may not take reciprocal markers. In other words, the presence of the reciprocal semantic features and that of reciprocal markers are not interlinked. For example, the English word *meet* encodes the situation where plural actors are involved in the action of meeting with each other when it appears with plural actors. This use of the word *meet* can be regarded as lexical reciprocal. The word *meet* in this use takes the reciprocal marker *each other*, as in the sentence *they met each other;* however, the sentence without the reciprocal marker, such as *they met* is also acceptable (Miller 1993:194). In contrast, the English word *love* is not considered to be a lexical reciprocal word because it requires the reciprocal marker, as in the sentence *they love each other*[42], when the word denotes the mutual state of affairs of what the word indicates. The sentence **they love*[43] does not signify the mutual action between the participants encoded by the pronoun *they*.

Furthermore, provided that the semantic components of reciprocal situations are the plurality of actors and the actor-cum-undergoer role, lexical reciprocal roots that take reciprocal markers may be further classified into the following three types. The first type is roots that have the two semantic components, but take reciprocal markers to encode reciprocal situations. The second type is roots that have the semantic component of plurality of actors and require an actor-cum-undergoer role's marker for the encoding of a reciprocal situation. The third type is roots that have the semantic component of actor-cum-undergoer role and require an actor plural marker for the encoding of a reciprocal situation. These types are summarized in the following table:

Table 11: Role of reciprocal markers by properties of roots

Features entailed in the semantic properties of roots	Plurality of actors	Actor-cum-undergoer role	Examples
Plurality of actors and Actor-cum-undergoer role	Not necessary	Not necessary	English They met.
Plurality of actors	Not necessary	Required	Kuuk Thaayorre Pul runc-**e**-r. 'They two ([i.e., the two of them]) collided with one [an]other.'
Actor-cum-undergoer role	Required	Not necessary	Buryat Cecegmaa bid xoyor tani-**lc**-aad xeden džil bol-loo. 'Several years have passed since C[ecegmaa] and I got acquainted.'

42 The sentence is from Miller (1993:194).
43 The sentence is from Miller (1993:194).

A similar concept of the roots of the first type is deponential reciprocal or 'reciproca tantum' in the terminology of Nedjalkov (2007a:14). The term refers to a reciprocal word that consists of a root and a reciprocal marker, but where the root is not employed without the reciprocal marker (Nedjalkov ibid.). Examples of this sort include the French phrases *se disputer* 'to quarrel' (< **disputer) and *se bagarrer* 'to fight' (< **bagarrer) (Guenchéva and Rivière 2007:597).

The roots of the second and third types might be found in a language with a reciprocal-reflexive polysemous marker and one with a reciprocal-actor plural polysemous marker, respectively. These types are detailed below.

When the root conveys the semantic feature of the plurality of actors in a given context, the role of a reciprocal marker is to add the semantic feature of the actor-cum-undergoer role to the plural actors. For example, the Kuuk Thaayorre suffix *-e* encodes a reflexive situation when it takes a singular subject:

Kuuk Thaayorre (Pama-Nyungan)
(64) Ngay nhaanhath-**e**-ø.
1.SG.NOM/ERG watch.RDP-REFL-NONPAST
'I'm looking at myself.'
(Gaby 2008:260, glossing modified)

However, when it appears with a dual subject and some roots that are 'semantically reciprocal' (i.e., lexical reciprocal, YK), such as *runc* 'collide', it encodes a reciprocal situation:

Kuuk Thaayorre (Pama-Nyungan)
(65) Pul runc-**e**-r.
3.DU.NOM/ERG collide-REFL-PAST.PERFECTIVE
'They two ([i.e., the two of them]) collided with one [an]other.'
(Gaby 2008:260, glossing modified)

By contrast, when the root has the semantic feature 'actor-cum-undergoer role', the role of a reciprocal marker is to add the semantic feature 'plurality of actors' to the item that already has the semantic feature of the actor-cum-undergoer role. For example, Buryat has the actor plural marker *lsa-* (and its phonetic variants, including *lca-*), which encodes actor plural situations:

Buryat (Mongolic)
(66) Dordžo Bata xoyor surguuly-d sura-**lca**-dag.
Dorji Bat and/two school-DAT study-ACTOR.PLURAL-ITERATIVE
'Dorji and Bat go to school (lit. 'study at school') together.'
(Nedjalkov et al. 2007:1288, glossing modified)

This language has the reciprocal marker *lda-* (and *ido-* as one of its phonetic variants) for encoding reciprocal situations. Consider the following example:

Buryat (Mongolic)
(67) Dordžo Bata xoyor dzodo-**ldo**-džo.
Dorji Bata and/two beat-REC-CONVERB
'Dorji and Bat are fighting (beating with each other).'
(Nedjalkov et al. 2007:1285, glossing modified)

However, despite the existence of this reciprocal marker, the actor plural marker *lca-* can also denote reciprocal situations:

Buryat (Mongolic)

(68) Cecegmaa bid xoyor tani-**lc**-aad
Cecegmaa I and/two recognize-ACTOR.PLURAL-CONVERB

xeden džil bol-loo.
several year become-PAST
'Several years have passed since C[ecegmaa] and I got acquainted.'
(Nedjalkov et al.2007:1320, glossing modified)

In Buryat, the reciprocal-indicating actor plural marker may appear with roots that can generally be classified as lexical reciprocal ones according to their meanings in other languages, as in *yari-lca-* 'to speak with each other; converse' (< *yari-* 'to speak to, tell someone'), and with roots that cannot, as in *nyudara-lsa-* 'to push each other' (Nedjalkov et al. 2007:1321, cf. p. 1319). If we supposed that the actor plural marker only encoded the actor plural situations and not reciprocal ones at an earlier stage of this language, its reciprocal function could be attributed to the semantic feature of the 'actor-cum-undergoer role' that is inherent to a lexical reciprocal root. In such a case, the reciprocal usage for non-lexical reciprocal roots can be considered to be due to the usage extension of the one for lexical reciprocal roots.

As such, if we assume that the reciprocal situation comprises the plurality of actors and the actor-cum-undergoer role as its semantic features, at least three functions of what has been termed 'reciprocal' markers can be analyzed. However, for simplicity, this study determines the functions of markers in terms of the meanings of words derived with these markers, instead of considering the semantic types of their roots. In other words, regardless of the types of their roots, all the derived words that encode the reciprocal situations and take a single actor argument will be listed under the sections that concern 'reciprocal constructions'. Similarly, some actor plural and reflexive constructions can be formed with 'lexical actor plural' roots, such as *massacre*, and 'lexical reflexive' roots, such as *suicide*. However, this study does not concern itself with whether or not a root entails the categorical semantic features. Thus, for example, any actor plural forms that denote '(plural actors) kill (many people)', with the roots 'kill' and 'massacre', and any reflexive forms that mean 'kill oneself', with the roots 'kill' and 'suicide', are treated as the actor plural and reflexive constructions, respectively.

In addition, it is worth pointing out that a reciprocal situation may also be encoded with a combination of a reflexive (or actor-cum-undergoer role) marker and an event plural marker. For instance, in West Greenlandic, while the prefix *immin-* 'reflexive marker' encodes both the reciprocal and reflexive situations, the combination of the prefix and the iterative prefix *rar-* unambiguously encodes the reciprocal situations. In this construction, the iterative marker serves as an event plural marker. Compare:

West Greenlandic (Eskimo-Aleut)

(69) **Immin-**nut tuqu-pu-t.
REFL-all kill-INDICATIVE-3.PL
'They killed themselves' or 'They killed each other.'
(Fortescue 1984:166, quoted in Maslova 2008:231)

(70) **Immin-**nut tuqu-**rar-**pu-t.
REFL-all kill-ITERATIVE-INDICATIVE-3.PL
'They killed each other.'

(Fortescue 1984:166, quoted in Maslova 2008:232)

The two markers work in different domains. The reflexive-reciprocal marker encodes the actor-cum-undergoer role of the participants in actor function in both situations. On the other hand, the iterative marker encodes the plurality of the events in a reciprocal situation. In a reflexive situation, the action by an actor is performed only once; however, in a reciprocal situation, the number of the actions can be regarded as plural if we consider the plural participants performing the actions individually. For this reason, the combination of the two markers may exclusively indicate reciprocal situations in some languages.

2.2 Actor plural and undergoer plural situations and constructions

Cross-linguistically, it has commonly been observed that reciprocal markers may serve to indicate non-reciprocal situations in which plural actors perform the same actions, such as '(plural actors) sing'. These situations are termed 'actor plural situations', and constructions that encode such situations are referred to as 'actor plural constructions' in this study. The term 'actor plural', rather than other more accepted typological terms, such as 'sociative', is employed to be contrasted with the term 'undergoer plural' that indicates the presence of plural undergoers in a situation. It is further observed that some undergoer voice constructions may also encode actor plural situations in some WISP languages. However, unless otherwise noted, the term 'actor plural construction' refers to the construction that encodes plural actors in subject function in actor voice constructions in a WISP context.

Various terms with slightly different definitions are employed to refer to this type of situation. For example, the term 'sociative' is preferred in the recent typological literature concerning polysemies of reciprocal markers, such as Nedjalkov ed. (2007) and Evans et al. eds. (2011). The term 'sociative' was suggested by Nedjalkov (2007a:33) as a replacement for the terms 'associative', 'collective', and 'cooperative'. Of these terms, 'collective' is widely employed in other literature, such as in Lichtenberk (2000) and Kemmer (1993, 1997). Their definitions and uses of these terms vary, despite their approximate agreement that the situation in question refers to one in which an action is performed by plural actors that are not actor-cum-undergoers.

One of the primary issues concerning the semantic definitions is whether a collective meaning among the plural actors is present. In other words, whether the actor plural construction must have the 'collective' meaning that the plural actors perform an action or actions *together*. For example, Kemmer (1993:99) considers that the collective meaning, rather than a distributive meaning, is present among the plural actors denoted by an actor plural marker (or 'collective' marker in her terminology) that also indicates reciprocal situations. Lichtenberk (2000) also presupposes the presence of 'homogeneity' or 'simultaneity' of the actions performed by the plural actors in his 'collective situations', in contrast to his 'distributed situations' in which the actions are performed in multiple locations or multiple directions. Nedjalkov (2007a:33) also suggests the presence of a collective meaning among plural actors in his 'sociative' meaning. However, Nedjalkov (2007a:38) points out that while 'the meaning of intransitive sociatives can very often be rendered in English more or less adequately by the word "together"', in textual examples, 'they often acquire very subtle semantic nuances […, which] are sometimes difficult to express adequately in English'. He continues that such semantic components 'may be omitted in translation, […] or translated by a

separate word and thus acquire inappropriate emphasis absent in the original' (Nedjalkov ibid.). While allowing that there are meanings (his 'nuances') other than the collective meaning, he considers his 'sociative' situation or construction to have a 'distributive sense', as expressed by the phrase 'here and there' in the English sentence *thousands of the haystacks stand here and there,* which is provided as a translation of a Yakut sentence in an actor plural construction (Nedjalkov ibid., emphasis added).

In this study, the collective meaning is treated as a secondary semantic component because, in relation to its reciprocal polysemy, the critical semantic criterion that differentiates between reciprocal and actor plural situations is whether or not the plural actors also have undergoer roles. For this reason, the situations in which 1) plural actors perform an action/actions (in an unspecified manner) and 2) they perform an action/actions collectively, or 'they do something together' are both treated as '(unmarked) actor plural' situations, unless the formal distinction between the constructions encoding these situational types is present.

Similarly, actor plural situations may have some additional semantic components, which may or may not be encoded in a language. Such components include 1) the distributive meaning among the plural actors, as in the situation *they sing separately* and 2) the meaning of successive actions on the part of plural actors, as in the situation *they sing in turn*. However, despite the presence of these semantic properties, we treat these situations as subcategories of the actor plural category because such semantic differences in actor plural situations are considered to be secondary. In contrast to the treatment of the 'collective actor plural' category, these additional semantic components are terminologically specified, as in the term 'distributive actor plural', because our data demonstrate that some languages differentiate one of these subcategories morphologically from an unmarked actor plural category or between these subcategories.

On the other hand, we employ the term 'undergoer plural situation' to refer to a situation where plural undergoers are present. The undergoer plural situation is encoded in an 'undergoer plural construction' in which plural undergoers are encoded in subject function in an undergoer voice construction (or 'in the passive voice', if we employ a cross-linguistically more common term) in the WISP languages. For example, the sentence *the leaves are bundled together by me* is categorized as an undergoer plural construction because it is in an undergoer voice construction in which plural undergoers are encoded in the subject.

In a similar manner to reciprocal situations that are denoted by lexical reciprocal roots, such as *fight*, actor plural situations are not necessarily marked with a specific marker. In Japanese, for example, some sentences denote both situations with a singular actor and those with plural actors. For instance, the sentence *Kodomo ga hashiru* (child SUBJECT run) can be interpreted as 1) the singular actor situation 'A child runs' or 2) the plural actor situation 'Children run'.[44]

In some languages, the plurality of actors is marked by a marker that is obligatory and is traditionally considered to be inflectional. In English, the presence of the plural actors is encoded by means of inflection. For example, the plural actors in the situation described in the sentence *they sing a song* are marked by 1) the pronoun *they* and 2) the third plural present form *sing* of the verb *sing*. This verb form is not considered to be a derivational form that indicates actor plural situations because the plural marking is obligatory in this language.

Turkic languages effectively illustrate the distinction between an inflectional marker, which encodes the presence of plural actors and agrees with the actor argument, and a derivational marker that denotes the actor plural situations. The Turkic languages cited here have 1) the formal distinction between third person singular and plural agreement markers, in which the latter is marked with the form *-la(r), and 2) the actor plural and reciprocal marker *-š. For example,

44 These examples are produced by the current author whose native language is Japanese.

Karachay-Balkar has the third plural marker *-la* and the actor plural marker *-š*. These two markers co-occur, and the form with the two markers signifies a slightly different meaning from the one with the third person plural marker, as in the following examples with the root *dəbərtla* ‘set off at a gallop’:

Karachay-Balkar (Turkic)

(71) Ala dəbərtla-də-la. ‘They set off at a gallop.’ (ala ‘3.PL’, -də ‘PAST’)
Ala dəbərtla-**š**-də-la. ‘They all set off at a gallop [in a hurry, without order].’
(cf. Ol dəbərtla-də ‘He set off at a gallop.’ (ol ‘3.SG’))
(V. Nedjalkov and I. Nedjalkov 2007:1001)

In Kirghiz, however, the formal distinction between the third plural singular and plural agreement markers is not present, as in the following comparison between the Kirghiz and Tuvan past forms with the root *al* ‘take’:

(72) Kirghiz (Turkic)
al-də-**Ø** ‘s/he took / they took’

Tuvan (Turkic)
al-də-**Ø** ‘he took’
al-də(**-lar**) ‘they took’
(Nedjalkov 2007d:1171 and Tuular 2007:1236)

In the following Kirghiz example, the suffix *-əs* encodes an actor plural situation:

Kirghiz (Turkic)

(73)	Ajša	menen	Kaləjša	kül-üp	žat-**əš**-tə.
	Ajša	and	Kalə**j**ša	laugh-CONVERB	AUXILIARY-ACTOR.PLURAL-3.PAST

‘Ajsha and Kalyjsha laughed.’
(Nedjalkov 2007d:1255)

It should be noted that when a language has different forms for situations that have plural and dual actors, the terms ‘actor plural’ and ‘actor dual’ are employed to differentiate the two situations and constructions. The actor plural situation refers to a situation in which dual actors are present, as in the sentence *the two of them sing*; the actor plural situation refers to a situation in which non-dual plural actors are present, as in the sentence *the three of them sing*. It should also be noted that the term ‘undergoer dual’ is not employed in this study due to a lack of sufficient data in support of its presence in WISP languages.

2.3 Reflexive situations and constructions

A reflexive situation is defined as a situation in which the entity referred to as an actor also serves as an undergoer in the same event, as expressed by the sentence *John killed himself*. A construction that encodes this reflexive situation is termed ‘reflexive construction’. For example, English reflexive constructions are formed with reflexive pronouns, such as *myself*, *himself,* and *themselves*. In the English sentence *John killed himself*, the actor of the action of killing (i.e., ‘the killer’) is the entity referred to by the word *John*, and its undergoer (i.e., ‘the one who was killed’) is the one referred to by the reflexive pronoun ‘himself’, which, in turn, refers to the same entity indicated by the word *John*.

As with reciprocal and actor plural situations, not all reflexive situations are marked by the reflexive markers. For example, according to Kemmer (1993:53), the English word ‘wash’ takes a

reflexive pronoun as an object only when the subject is non-human, as in the sentence *the cat washes himself*. The data provided by Kemmer illustrate that the reflective action of washing performed by a human actor is indicated by 1) an intransitive construction, as in the sentence *the person washed* or 2) the construction with the word *get* and a past participle, as in the sentence *the person got washed*.

Reflexive and intransitive constructions sometimes exhibit formal unity. Reflexive situations are observed to be expressed in the same construction as intransitive constructions for some types of situations. For example, the French reflexive pronoun *se* (allomorph *s'*) encodes reflexive situations, as in the sentence *il **se** lave* 'He washes [himself]'[45]. This reflexive construction is also employed to indicate some situations in which an inanimate actor experiences an action, as in the sentence *la porte **s**'ouvre* 'The door opens'[46]. Similarly, some reflexive constructions also indicate various passive- or result-like situations in which the (inanimate) participants are affected by actions by other participants, as in the French sentence *le livre **se** vend bien* 'The book sells well' [(The book REFL sells well)]. By contrast, some intransitive constructions without a specific reflexive marker encode reflexive situations, as do the English translations for these sentences.

It should be noted that this study does not employ the notion of 'middle', as in 'middle situation', 'middle meaning', 'middle voice', and 'middle marker', which is suggested by Kemmer (1993) based on semantic and formal properties that are claimed to be different from reflexive meaning and markers proper. According to Kemmer (1993), the 'middle situations' refer to (reflexive) situation types, such as grooming actions, changes in body postures, and 'spontaneous actions' that include 'grow' and 'change'. Furthermore, Kemmer (1993:25–26) observes that in some languages, 'reflexive' and 'middle' markers differ and that when this distinction is present, the 'middle' marker is a 'light form' and the 'reflexive' marker is a 'heavy form'. In other words, the 'middle' marker is morphologically shorter than the reflexive marker proper, as in the Russian pair *sebja* 'REFLEXIVE' and *-sja* 'MIDDLE' and the Dutch pair *zichzelf* 'REFLEXIVE' and *zich* 'MIDDLE' (Kemmer ibid.).

However, we do not employ the distinction between reflexive and 'middle' situations and construction because, as Palmer (1995) states, the distinction between the 'middle' meanings and intransitive actions (i.e., one-participant actions) is less convincing because the putative 'middle' meanings are typical for situations denoted by intransitives. In addition, the distinction between the 'middle' and reflexive situations is also questionable because, as mentioned earlier, some 'middle' situations indicated by intransitive constructions in one language are indicated by reflexive constructions in another, as in the English and French expressions *he washes* and *il se lave* and *the door opens* and *la porte s'ouvre*. Even if such a distinction is formally manifested, as in Dutch and Russian, it is considered to be due to a result of subcategorization of the reflexive category, if we define the reflexive category semantically as the category that encodes self-directed actions. For this reason, the distinction is unnecessary in the discussion of polysemy with reciprocal or actor plural situations. Thus, rather than employing the putative 'middle' situations as an independent category, this study treats them as reflexive situations.

2.4 Comitative constructions

Comitative constructions are defined as constructions that encode plural participants of the same macro-role in two discontinuous arguments. This study distinguishes two types of comitative constructions: 1) comitative actor constructions and 2) comitative undergoer constructions.

The term 'comitative actor construction' is employed to refer to the construction with an actor

45 Data from Kemmer (1993:59).
46 Data from Bril (2005:31).

argument and a comitative actor argument, as in the sentence *John runs together with Bill*. The actor argument in this construction can be a subject in an actor voice construction, such as the subject *John* in the sentence *John brought the table with Bill cooperatively* or a non-subject in an undergoer voice construction, such as the argument *by John* in the sentence *the table is brought by John cooperatively with Bill.*

Similarly, the term 'comitative undergoer construction' is employed to refer to the construction with an undergoer argument and a comitative undergoer argument, as in the sentence *a notebook is brought together with a pencil by John*. (cf. in an undergoer plural construction: *a notebook and a pencil are brought together by John*.) However, in contrast to the comitative actor construction, we use this term to refer only to the construction in which the primary undergoer is encoded in subject function in an undergoer voice. In other words, we do not treat the sentence *John brought a notebook together with a pencil* as a comitative undergoer construction because the main undergoer 'a notebook' and its comitative undergoer 'a pencil' are encoded in a non-subject argument in an actor voice construction.

It should be noted that there may be differences in involvement between a primary actor and a comitative actor or between a primary undergoer and a comitative undergoer. However, for simplicity, we consider comitative constructions to be different argument realizations of participant plural constructions, not considering some variations with regard to the involvement of an action between an actor and its comitative actor, or that between an undergoer and its comitative undergoer.[47] In other words, we do not differentiate the situations in which 1) the primary actor or undergoer plays the primary role, and the comitative actor only assists his or her action and in which 2) the primary and comitative actors or undergoers have the same role and degree of involvement. Thus, the term 'comitative situation' is not employed in this study.

In addition, as mentioned earlier, the plural actors in a reciprocal situation may also be expressed in an actor argument and a comitative actor argument, as in the sentence *Pierre se bat avec Jean.* 'Pierre fights (with) Jean'.[48] [(Pierre REC fight.3.SG.PRESENT with Jean)]. In this study, this construction is referred to as a 'reciprocal comitative actor construction' and is regarded as a subtype of the comitative actor constructions when dealing with our WISP data. Thus, for example, the construction of the French sentence *Pierre et Jean se battent* '[Pierre and Jean] are fighting'[49] or 'Pierre and Jean fight each other' [(Pierre and Jean REC fight.3.PL.PRESENT)] is classified as a reciprocal construction. However, the construction of the sentence *Pierre se bat avec Jean* 'Pierre fights (with) Jean' [(Pierre REC fight.3.SG.PRESENT with Jean)] is classified as a comitative actor construction.

2.5 Polysemies of reciprocal constructions

Cross-linguistically, polysemies of reciprocal markers are often reported (cf. Nedjalkov 2007a, 2007b; König and Gast 2008; Lichtenberk 2000). Among the polysemous patterns, reciprocal-reflexive, reciprocal-actor plural, and actor plural-comitative actor polysemies are the primary concerns of this study.

The relationships among the categories in these polysemies can be explained by the three components that are proposed in the initial part of this chapter. They are 1) the actor-cum-undergoer role, 2) the plurality of actors, and 3) the difference in the number of the arguments that encode the plural actors. For example, the common feature 'actor-cum-undergoer role' can be suggested for

47 Nedjalkov (2005a), for example, considers the degrees of involvement in actions by comitative actors and differentiates the two categories that are termed 'comitative' and 'assistive'.

48 The French examples in this paragraph are cited from Guentchéva and Rivière (2007:565).

49 This sentence is originally translated as 'Jean and Pierre are fighting.' in Guentchéva and Rivière (2007:565).

both reciprocal and reflexive situations when they are indicated by the same marker. Along the same lines, we can assume the plurality of actors to be a common feature between the reciprocal and actor plural categories and between the actor plural and comitative actor categories.

Some common features can be omitted. For example, the common syntactic feature 'only one argument for encoding actor(s)' may also be suggested in the reciprocal-reflexive and reciprocal-actor plural polysemies. However, it can be omitted because the difference in the number of arguments that encode the plural actors is relevant only when the polysemous marker also encodes the comitative actor construction, where plural actors appear in two arguments.

The examples below illustrate some of the polysemies and their common components:

Reciprocal-reflexive polysemy: [+ actor-cum-undergoer] ([+ one argument])
German (Germanic, Indo-European)
(74) Sie achten **sich**.
3.PL.NOM respect.3.PL REFL/REC
'They respect themselves/each other.'
(Nedjalkov 2007b:187, glossing modified)

Shisumbwa (Niger-Congo (Bantu))
(75) **-i**-shima 'to love oneself/each other' -shima 'to love'
-i-ihaga 'to kill oneself/each other' -ihaga 'to kill'
-i-gaya 'to hate oneself/each other' -gaya 'to hate'
-i-tema 'to cut oneself/each other' -tema 'to cut'
(Capus 1898:64, quoted in Nedjalkov 2007b:262)

Reciprocal-actor plural polysemy: [+ actor plural] ([+ one argument])
Yakut (Turkic)
(76) tapta-**s**- 'love each other' Reciprocal
köt-**üs**- 'fly together' Actor plural
(I. Nedjalkov and V. Nedjalkov 2007a:1136 and 1132, quoted in Nedjalkov 2007b:337)

Actor plural-comitative actor polysemy [+ actor plural]
Abaza (Northwest Caucasian)
Actor plural
(77) ja-**cə**-gI-ga-tI.
that-ACTOR.PLURAL-we-carry-PAST
'We carried that together.'
(Tabulova 1976, quoted in Nedjalkov 2007a:44, glossing modified)

Comitative actor
(78) j-r**ə**-**cə**-z-ga-tI.
that-they-COMITATIVE.ACTOR-I-carry-PAST
'I carried that with them.'
(Tabulova 1976, quoted in Nedjalkov 2007a:44, glossing modified)

Evenki (Tungusic)
Actor plural

(79) Nuŋartən tatkit-tula ŋene-**ldə**-re-Ø.
they school-ALLATIVE go-ACTOR.PLURAL-NONFUTURE-3.PL
'They went to school together.'
(I. Nedjalkov and V. Nedjalkov 2007b:1608, quoted in Nedjalkov 2007a:43, glossing modified)

Comitative actor

(80) Nuŋan tatkit-tula ŋene-**ldə**-re-n.
he school-ALLATIVE go-COMITATIVE.ACTOR-NONFUTURE-3.SG
'He went to school with somebody.'
(I. Nedjalkov and V. Nedjalkov 2007b:1608, quoted in Nedjalkov 2007a:43, glossing modified)

As illustrated by these polysemies, reciprocal, actor plural, reflexive, and comitative actor constructions are suggested to be related by some shared features between various parts of the set. Such common features assist us in exploring the directions of functional changes. For example, when a language has a marker that forms reflexive, reciprocal, and actor plural constructions, and comparative and historical studies concerning the language suggest that the reflexive function is original, the semantic components of the three categories imply that the reflexive function first developed into the reciprocal function, and then into the actor plural function.

2.6 Other categories relevant to this study

This section discusses some other categories that are observed as being marked by reciprocal markers in typological studies and that are considered to be relevant to the current study concerning our WISP data. These categories are 1) nominal duality, 2) nominal plurality, 3) converse kinship pairs, 4) associative meanings, and 5) symmetrical locational relationship.

2.6.1 Nominal duality

Some reciprocal markers are also employed to indicate the duality of the items indicated by a nominal (i.e., object-denoting) root (Nedjalkov 2007c:360). For example, the Southern Paiute reciprocal marker *na-*, as in *na-ɣwɪ'pa-* 'to hit each other',[50] is also used as a dual marker for nominal roots with the (nominal) plural marker *-ŋwï*, as shown below:
Southern Paiute (Uto-Aztecan)

(81) **na-**rï'χïv^{w}ï-ŋwï 'two friends' tïɣïv^{w}ï- 'friend'
na-va'vi-ŋwï 'two brothers' pavi- 'elder brother'
na-ɣï'mantsi-ŋwï 'two who are strangers to each other' qïmantsi- 'stranger'
(Sapir 1930:109–10, quoted in Nedjalkov 2007c:360)

Ainu also exhibits the polysemy of the reciprocal and dual markers, as in the following examples. In this language, the reciprocal marker is *u-*, as in *u-nukar* 'to see each other'(< *nukar* 'to see'):

50 Nedjalkov (2007c:358).

Ainu (Language Isolate)

(82)	**u**-irwak	‘both brothers’	irwak	‘brother’
	u-anun	‘mutual strangers’	anun	‘stranger’

(Alpatov et al. 2007, quoted in Nedjalkov 2007c:354, 360)

In line with the duality function, a pair relationship, namely, a relationship between two items that are generally considered to be a pair, may also be indicated by the reciprocal marker. Consider the following Ainu examples with the roots that mean ‘foot’ and ‘eye’. Both items are regarded as being a pair relationship when they refer to a part of a human body:

Ainu (Language Isolate)

(83)	**u**-kema	‘both feet’	kema	‘foot, feet’
	u-sik	‘both eyes’	sik	‘eye’

(Alpatov et al. 2007, quoted in Nedjalkov 2007c:362)

With regard to the relationship between reciprocity and duality, Sapir (1930:110, quoted in Nedjakov 2007c:359) claims that ‘[t]he idea of reciprocity leads naturally to that of duality of terms involving mutual relationship’. However, these two categories should be differentiated, as demonstrated by the fact that some languages encode reciprocal situations with dual and non-dual plural actors differently, as in the following Hualapai examples:

Hualapai (Yuman-Cochimí)

(84)	’u:k	‘to see’
	jij’u:**b**k	‘(two persons) to look at each other’
	jij’u:**v**k	‘(three or more persons) to look at one another’

(Watahomigie et al. 2001:331, quoted in Heine and Miyashita 2008:177)

2.6.2 Nominal plurality

Morphological and semantic similarities between the dual and plural nominal expressions in relation to a reciprocal marker are demonstrated in the following Southern Paiute data. In this language, the plurality of the items indicated by an object-denoting root is indicated by *na-na-*, which is the reduplicated form of the reciprocal and dual marker *na-* (Nedjalkov 2007c:360):

Southern Paiute (Uto-Aztecan)

(85)	**na**-va´vi-ŋwï	‘two brothers’
	na-na´-vavɪ-ŋwï	‘three or more brothers.’

(Sapir 1930, quoted in Nedjalkov 2007c:360)

2.6.3 Converse kinship pairs

The term ‘converse kinship pair’ refers to a pair that consists of the kinship member indicated by the root and its counterpart, such as ‘mother and (her) child’ with the root ‘mother’.[51] This category has something in common with the nominal dual function mentioned above in the sense that the markers of both categories appear with object-denoting roots and derive object-denoting words that

51 This grammatical phenomenon has suffered from a lack of terminological consensus among researchers: It is referred to using a variety of terms, such as ‘dyad(ic)’, ‘kinship proprietives’, ‘kinship duals’, ‘reciprocal plurals’, ‘collective nouns’, and ‘kinship pairs’ in Australianist literature, and has been ‘invariably treated as special cases of reciprocals as applied to kinship nominals’ in non-Australianists’ literature (Evans 2003:2). Of these terms, perhaps the most neutral term that does not refer to the connection with other verbal categories (e.g., reciprocals or collectives) or parts of speech by their terms would be dyadic or Evans’ coinage ‘dyad’.

encode the duality of the items denoted by the root. However, this kinship category deserves to be treated separately from the dual category because it is regarded as an independent category in relation to the reciprocal category in typological studies.

Cross-linguistic observations demonstrate that some converse kinship pair constructions share the same markers as the reciprocal constructions in a language (Evans 2003, Nedjalkov 2007c). For example, in Mono, the reciprocal marker *na'na-* also encodes a converse kinship pair:

Mono (Uto-Aztecan)

(86) **na'na**-pet**ɨ** 'parent and daughter' pet**ɨ** 'daughter'
(cf. ***na'na-****'waqa* 'to talk with each other')

(Nedjalkov 2007b:361 and 358)

It should be noted that this construction is different from simple dual forms of object-denoting roots. Simple dual constructions and converse kinship pair constructions differ with regard to the property of the dual items indicated by the constructions. The simple dual markers may appear with object-denoting roots, such as the kinship root that denotes 'mother' and the inanimate root that indicates 'cheek', and derive words that signify the presence of the two items with the same property denoted by the root, such as 'two mothers' and 'two cheeks'. However, the converse kinship pair marker often appears exclusively with kinship terms and derives forms that encode the presence of the two items with the role indicated by the root and its counterpart, such as 'mother' and 'child (mother's counterpart)' with the root 'mother'. In other words, the two items indicated in the converse kinship pair construction are different, in contrast to the 'pure' dual construction.

When the two categories are indicated by the same marker, it could be suggested that it is the relationship of 'mother-and-child' that is shared by the two roles, and that the use of the converse pair is due to the 'pure' duality indicated by a derivational marker and the semantic property of the kinship root. This hypothesis may be true in some languages in which the items 'mother' and 'child' are expressed by the same lexical item. However, the hypothesis is invalid when a language makes a terminological distinction between the two relatives encoded in the converse pair construction. Because the distinction is also present in our WISP data, it is necessary to distinguish between the two categories in this study.

The necessity of the distinction between reciprocal and converse kinship pair categories can also be demonstrated by the fact that other components are present in addition to reciprocal markers in some converse kinship pair constructions. For example, the Ainu converse kinship pair construction *u- (-i)-kor* comprises the reciprocal marker *u-*, the verb *kor* 'have', and the possessive marker *-i*:

Ainu (Language Isolate)

(87) **u**-ona-**kor** 'to be related as father and son' ona 'father'
u-po-**kor** 'to have a son-parent relationship' po 'son, child'
u-ak-**i**-**kor** 'to have a younger brother–older brother or younger brother–older sister relationship' ak 'younger brother'

(Alpatov et al. 2007, quoted in Nedjalkov 2007c:355 and 367)

In Southern Paiute, the converse kinship construction is indicated by the form *na- -ŋwï*, which consists of the dual and reciprocal marker *na-* and the plural marker *-ŋwï*, as shown below:

Southern Paiute (Uto-Aztecan)

(88) **na**-vi'a-**ŋwï** 'mother and daughter' pia (vi'a) 'mother'
na-ŋ'w'tsi-**ŋwï** 'father and son' moa(tsi) (ŋ'w'tsi) 'father'

(Sapir 1930:109–10, quoted in Nedjalkov 2007c:361)

These examples suggest that even if a reciprocal marker is employed to form converse kinship pair constructions, it is worth considering whether other components are present prior to concluding that it is the reciprocal marker that derives the converse kinship constructions. The requirement of the additional markers can be explained if we assume that there is a semantic difference between the items for reciprocal and converse kinship pair constructions: Roots that appear with the reciprocal constructions are deemed to be action-denoting (verbal), and those that are used for the converse kinship pair constructions are kinship terms that are object-denoting (nominal). The coding of these two lexical classes are different in the majority of languages (cf. Nichols 2016). This tendency suggests the possibility that if an object-denoting root appears with an affix for action-denoting roots, there would be another marker to convert the object-denoting root into an action-denoting root. Due to the fact that the converse kinship pair indicates an entity that is more likely to be encoded by an object-denoting word rather than by an action-denoting word, one of these markers must have functions that change the lexical classes from the action-denoting class to the object-denoting class. Such markers may be 1) a genuine class changing marker or 2) specific to the object-denoting roots or words, such as a plural marker. In the latter case, the class changing of roots or words is effected by the presence of the marker.

In addition to the nominal encoding, some verbal encoding, such as a copulative marker 'TO BE' or a possession marker 'TO HAVE', may also be required when converse kinship pairs are expressed predicatively, such as '(plural actors) be a father and his child' or '(plural actors) have a father-and-child relationship'. The presence of a verbal marker for the predicative use of the expressions for the converse kinship pairs would be particularly true when a language requires such a verbal encoding to express predicates derived from nouns, as observed in the Ainu segment *(-i)-kor* 'have' in the converse kinship pair construction *u- (-i)-kor*.

2.6.4 Associative

This study employs the term 'associate' to indicate 'an individual that possesses the same item (object or quality) as someone else', and the term 'associative meaning' to refer to this meaning. The associative meaning appears to be slightly different from the other polysemous categories with reciprocal situations and constructions that are addressed in this study. However, cross-linguistic data demonstrate the presence of some formal commonalities between reciprocal and associative markers. In Kirghiz, for instance, the segment *-la-š* (and its allomorphs, such as *-lo-š*, *-le-š*, *-ta-š*, and *-da-š*) appears with some object- and quality-denoting roots to form words that indicate the 'similarity of two (or more) persons or entities with respect to the feature named by the root' (Nedjalkov 2007d:1266). Consider:

Kirghiz (Turkic)
With object-denoting roots

(89)	tür-**dö-š**	‘of the same shape, homogeneous’	tür	‘appearance, shape’
	öŋ-**dö-š**	‘of the same colour, alike’	öŋ	‘face, complexion, colour’
	zaman-**da-š**	‘contemporary’	zaman	‘time, epoch’
	žal-**da-š**	‘(persons) of the same age/year of birth’	žal	‘year’
	din-**de-š**	‘co-religionist’	din	‘belief, religion’
	klass-**ta-š**	‘class-mate’	klass	‘class’
	žamaat-**ta-š**	‘member in the same community’	žamaat	‘community, society’
	žol-**do-š**	‘fellow-traveller, comrade’	žol	‘road, way’
	ata-**la-š**	‘born of the same father’	ata	‘father’
	olžo-**lo-š**	‘(one) sharing the loot, bag’	olžo	‘loot, bag’
	enči-**le-š**	‘co-heir, co-parcener’	enči	‘share in inheritance’
	kəzmat-**ta-š**	‘colleague; assistant’	kəzmat	‘service, work’
	bötölkö-**le-š**	‘boon companion’	bötöl	‘bottle’
	čada-**la-š**	‘neighbouring’	čada	‘border, boundary’

With quality-denoting roots

(90)	čama-**la-š**	‘equal in strength, power’	čama	‘strength, power’
	boj-**lo-š**	‘(persons) equal in height, age’	boj	‘height’
	teŋ-**de-š**	‘an equal’	teŋ	‘equa

(Nedjalkov 2007d:1269)

The segment *-la-š* comprises the (originally denominal) suffix *-la* (e.g., *sojul-da* ‘to beat with a cudgel < *sojul* ‘cudgel’ (Nedjalkov 2007d:1264)) and the reciprocal, actor plural, and comitative actor marker *-(V)š* (e.g., *ur-uš* ‘to fight’ < *ur-* ‘to beat/hit’ (Nedjalkov 2007d:1259)). It also derives reciprocal verbs from nouns denoting actions, as in *arəz-da-š* ‘to quarrel’ (< *arəz* ‘quarrel’) and *aŋgme-le-š* ‘to converse’ (*aŋgme* ‘conversation’) (Nedjalkov 2007d:1264–1265).

2.6.5 Symmetrical locational relationship

Some reciprocal markers derive words that indicate a symmetrical relationship with regard to two locations from locational roots of various lexical classes, such as roots indicating ‘next to’, ‘near’, and ‘side’. In Yakut, the symmetrical relationship is marked with *-las* (and its allomorphs, such as *-tes*), that comprises the reciprocal suffix *-s* (e.g., *ələ-s* ‘to take/grasp each other’ < *əl* ‘to take’[52]) and the denominal marker *-laa* (and its allomorphs, such as *-taa*):

Yakut

(91)	əksa-**las-**	‘to be next to each other’	əksa	‘nearby, close (to)’
	serges-**tes-**	‘to line up, walk next to [each other]’	serge	‘next to, near’
	eŋer-**des-**	‘to live next to [each other]’	eŋer	‘side, edge’
	ojoγos-**tos-**	‘to be side by side with sb’	ojoγos	‘side, rib’

(I. Nedjalkov and V. Nedjalkov 2007a:1150)

52 I. Nedjalkov and V. Nedjalkov (2007:1142).

3 The prefix *si- in WISP languages with action-denoting roots

This chapter discusses forms with the prefix *si- that take action-denoting roots along with their respective functions in WISP languages.

*3.1 Reconstructing *si- formations with action-denoting roots*

The purpose of this section is to introduce the basic morphological structure of formations with the prefix *si- to prepare a working hypothesis for a more detailed discussion of the derivation with *si- from action-denoting roots.

When reconstructing formations with a derivational affix in WISP languages, it is important to hypothesize the overall construction in which the affix in question occurs, in order to avoid the inclusion of homophonous affixes. More specifically, we need to consider the forms and functions of other affixes that may precede or follow it, in addition to scrutinizing those of the target affix itself. For this reason, prior to proceeding any further, it is advisable to investigate the affixes that co-occur with the prefix *si- and the functions they have.

As mentioned in Chapter 1, no explicit reconstruction has been proposed for the constructions with *si- to date. However, Adelaar (1992a, 2005a) suggests the proto-form *si- in discussing other issues and states that the following Malay and Salako formations reflect it. For this reason, we begin our reconstruction based on his arguments and data. Some data that he adduces in support of his reconstruction of *si- in Adelaar (1992a) are as follows:

Malay

(92)

bər-si-təkan lutut	‘with one’s arms (leaning) on one’s knees’ [lutut ‘knee’]	mənəkan (< təkan) ‘to press’
bər-sə-tumpu	‘to take off against each other (e.g., in [a] tug-of-war game)’	tumpu ‘take-off, abutment’
bər-si-təgang urat leher	‘(to cause each other’s neck muscles to become tense =) to fight with each other’ [urat ‘muscle’, leher ‘neck’]	təgang ‘tended, tense’
bər-sə-tubuh	‘to have sex[ual intercourse]’	tubuh ‘body’

(Adelaar 1992a:395–6)

Salako

(93)

ba-si-muhà	‘to wash one’s face’	muhà	‘face’
ba-sing-komor	‘to rinse one’s mouth’	komor	[‘gloss unknown’]
ba-sing-koʔopm	‘to put entirely into one’s mouth’	koʔopm	[‘gloss unknown’]
sin-soor	‘to throw oneself [on]to the floor out of recalcitrance, roll oneself [across] the floor’	soor	[‘gloss unknown’]

(Adelaar 1992a:396)

In his paper in 1992, Adelaar attributes the proto-form *si- to the Proto-(West-)Malayo-Polynesian

level and suggests that it marks ‘reciprocal (and medial?) verbs [i.e., reciprocal (and reflexive) verbs; the question mark is original, YK]’. Furthermore, he mentions that these Malay and Salako forms correspond to Toba Batak *mar-si-*, which ‘[denotes] reciprocity [(i.e., derives reciprocal predicates from action-denoting roots, YK)], and to Minangkabau *ba-si-*, which has ‘a wider meaning’ (Adelaar 1992a:395–396).

Adelaar’s proposal concerning the proto-form *si- is insightful in its own right. However, his suggestion contains three issues that should be addressed in order to enhance our understanding of *si- forms.

First, he does not mention the fact that the prefix *si- occurs with different roots that belong to varying lexical classes in the examples he provides. While the majority of the *si- forms in his data comprise action-denoting roots, there are a few roots that are object- or quality-denoting. For example, the Malay root *tubuh* ‘body’ in *bər-sə-tubuh* ‘to have sex[ual intercourse]’ and the Salako root *muhà* ‘face’ in *ba-si-muhà* ‘to wash one’s face’ are object-denoting, while the Malay root *təgang* ‘tended, tense’ is quality-denoting. Notwithstanding the differences in lexical classes, however, all the forms in question derive action-denoting words.

Second, he does not explore the function of the prefix that precedes *si-, such as *bər-* in the Malay form *ber-si-,* although he recognizes the presence of the prefix, stating that the Malay prefix *si-* ‘is almost always preceded by *bər-*’ (Adelaar 2005a:132). As he observes, almost all Malay and Salako examples that he adduces contain *bər-* and *ba-* that precede the prefix *si-* or *sə-*. The presence of the prefixes in his data leads us to consider whether a simple reconstruction of *si- is warranted or an affix combination with *si- should be reconstructed. This issue depends crucially on the evaluation of examples regarding the occurrence of the *si- forms without *maR-. Furthermore, the strong tendency of *si- to co-occur with the actor voice markers requires explanation, regardless of how these examples are analyzed.

The functions of the prefixes *bər-* and *ba-* have been sufficiently analyzed in their own right. For example, it is known that the modern Malay prefix *bər-* derives intransitive predicates from action-, quality-, and object-denoting roots (Adelaar 1984:410–411); however, in Classical Malay, it derives both intransitive and transitive predicates in actor voice constructions (Roolvink 1965:333–334, also quoted in Adelaar ibid.). Similarly, the Salako prefix *ba-* derives intransitive predicates from action- and quality-denoting, numeral, and unanalyzable roots (Adelaar 2005c:42). If we take these facts into account, it can be stated that the prefixes in question in both languages are employed to form action-denoting predicates in actor voice constructions that encode actors in subjects. Consequently, we can assume that the reciprocal and reflexive situations encoded in the forms with the prefix *si- are encoded in actor voice constructions, despite the lack of examples in sentence context in Adelaar’s data for these languages.

Furthermore, the existing data and arguments can sufficiently account for the diachronic formal and functional properties of the two prefixes. According to Adelaar (1992a:395), Malay *bər-* and Salako *ba-* reflect the Proto-Malayic prefix *mar- or *mər-. This Proto-Malay prefix reflects Proto-Austronesian *maR- (Adelaar 1984:418)[53]. Because Proto-Austronesian *maR- is reflected as Proto-Malayo-Polynesian *maR- (Ross 2002), we can state that the Malay and Salako forms reflect the Proto-Malayo-Polynesian prefix *maR-. Thus, the diachronic path can be formulated as follows: Proto-Austronesian *maR- > Proto-Malayo-Polynesian *maR- > Proto-Malayic *mar- or *mər- > Malay *bər-* and Salako *ba-*. This Proto-Malayo-Polynesian prefix *maR- consists of the actor voice infix *<um>, which forms predicates in actor voice constructions, and the action-denoting lexical class prefix *paR- (cf. Ross 2002).

53 Adelaar (1984) suggests a slightly different Proto-Malayic form. However, the formal difference in his suggestion is irrelevant for the claim that the Proto-Malayic form that is reflected as Malay *bər-* reflects Proto-Austronesian *maR-.

The prefix *maR- has a paradigmatic relationship with other undergoer voice forms with the prefix *paR-, in a similar manner to the actor voice form *<um> regarding other undergoer voice forms (cf. Ross 2002). The voice paradigms with the actor voice forms *<um> and *maR- are illustrated in the following tables:

Table 12: Proto-Malayo-Polynesian voice and mood forms (cf. Ross 2002)

	Actor voice	Undergoer voice		
		Patient voice	Locative voice	Thematic voice
Neutral mood	*<um>	*-en	*-an	*i-
Realis mood	*<um><in>	*<in>	*<in> -an	*i- <in>

Table 13: Proto-Malayo-Polynesian voice and mood forms with *paR- (cf. Ross 2002)[54]

	Actor voice	Undergoer voice		
		Patient voice	Locative voice	Thematic voice
Neutral mood	*maR- (< *p<um>aR-)[55]	*paR- -en	*paR- -an	*i-paR-
Realis mood	*naR- (< *p<um><in>aR-)[56]	*p<in>aR-	*p<in>aR- -an	*i-p<in>aR-

Due to the fact that the prefix *maR- forms actor voice constructions, it is assumed that the prefix *si- is derivational and has a function other than indicating the actor voice constructions at the Proto-Malayo-Polynesian level. This argument, however, does not indicate that it is the prefix *si- on its own that forms the reciprocal and reflexive constructions. Rather, it is more probable that it is the combination of *maR- and *si- that forms the two constructions.

The final issue that should be addressed regarding Adelaar's proposal is that he does not employ any data from Philippine and Sulawesi languages, although he reconstructs a proto-form that can be traced back to somewhere near to the Proto-Malayo-Polynesian stage. For reconstructing Proto-Malayo-Polyesian morphemes based solely on WISP languages, it is ideal to use data in languages in all the three areal groups. In particular, comparison between Philippine and Western Indonesian languages appears to play a significant role in the reconstruction.

If we search for the form that contains *si- and *maR- in Philippine languages to obtain more data for examining the question raised on the basis of the Malay and Salako data, we can find the Tagalog form *mag-si-*. Phonologically, it is justified to assume that this form reflects the Proto-Malayo-Polynesian *maR-si- because, according to Blust (1991:90), the Tagalog phoneme /g/ reflects Proto-Malayo-Polynesian *R). However, in contrast to the Malay and Salako forms, the Tagalog form appears with an action-denoting root and encodes a non-reciprocal situation in which plural actors in subject function perform the action indicated by the root. For example, in the following Tagalog sentence, the word *nag-si-alis* '(plural actors) left' consists of the root *alis* 'leave' and the form *nag-si-*, which is the realis form of *mag-si-*. The word appears with the plural actor subject *sila* '3.PL.NOM', and encodes the situation in which the action of 'leaving' is performed by

54 Ross (2002:50) reconstructs the patient voice and locative voice forms without *paR-. The current reconstruction with *paR- is based on Sarangani Manobo forms, such as PV *peg- -en*, and LV *peg- -an* (cf. DuBois 1976:n.p.), in which the prefix *peg-*, which is considered a reflex of *paR-, is present.

55 Blust (2013:374).

56 Blust (2013:375).

plural actors:

Tagalog

(94) Silá'y nag**si**alis.
sila ay nag-si-alis
3.PL.NOM INV AV.RLS-SI-leave
'They went away.'

(Lopez 1977a [1937]:91, my glossing)

In spite of the functional difference, we posit that the Tagalog form *mag-si-* is a cognate of Malay *bər-si-*, Salako *(ba-)si(N#)-*, and other reciprocal and reflexive forms with the prefix *si- in other Western Indonesian languages that Adelaar adduces for the following two reasons: 1) the sound correspondences between Proto-Malayo-Polynesian and Tagalog and 2) the cross-linguistic tendency that the same morpheme may encode both reciprocal and actor plural situations, which is mentioned in Chapter 2. If we consider that the Malay, Salako, and Tagalog forms originate from the same source, we can arrive at the assumption that the proto-form *si- is a derivational prefix used with *maR- to indicate reciprocal, reflexive, and actor plural situations in actor voice constructions.

Furthermore, the presence of the actor voice marker in *maR-si- forms raises the question of whether the prefix *si- also occurs with other voice-marking affixes. As to this question, Cotabato Manobo provides evidence that a reflex of the prefix *si- may appear with all the four basic voice forms. These forms indicate either plural actors or plural undergoers:

Cotabato Manobo

Actor voice *eg-se- -ay*: Reciprocal (the 'they' speaking to each other are actors and plural)

(95) **Eg-se**-ikagi-**yay** da.
AV-SI-speak-EP 3.PL.NOM
'They are speaking to each other.'

(Kerr 1988:97, my glossing)

Patient voice *eg-se- -en*: Undergoer plural (the *libi* palm leaves are undergoers and plural)

(96) **Eg-se**-lapin-**en** di sa libi.
UV-SI-layer-PV 3.SG.GEN NOM libi
'He is interlaying the *libi* palm leaves.'
['The *libi* palm leaves are being interlaid by him.']

(Kerr 1988:99, my glossing)

Locative voice *eg-se- -an*: Actor plural (the 'they' crowding in are actors and plural)

(97) **Eg-se**-linud-**an** da aken.
UV-SI-crowd-LV 3.PL.GEN 1.SG.NOM
'They crowd in upon me.'

(Kerr 1988:101, my glossing)

Thematic voice *eg-se-*: Undergoer plural (the tins are undergoers and plural)

(98) **Eg-se**-tipon ku sa me-doo lata.
TV:UV-SI-gather 1.SG.GEN NOM QD-many tins
'I gather the tins together (in one spot)'
['Many tins are gathered by me.']

(Kerr 1988:101, my glossing)

These examples have two features that require explanation before a comparison can be drawn with the other data examined thus far. First, the initial phonemes of the voice markers with *paR- are absent in Cotabato Manobo. These include 1) the initial consonant /m/ in the actor voice prefix *maR-, 2) the initial consonant /p/ in the undergoer voice form *paR- in the patient voice form *paR- -en and the locative voice form *paR- -an, and 3) the initial phonemes /i-p-/ in the thematic voice form *i-paR-. Second, the vowel *a in the forms *maR- and *paR- appears to be other vowels, such as /e/, /ə/, and /o/, in some Manoboic languages including Cotabato Manobo. Compare the voice markers in Tagalog, Cotabato Manobo, and three other Manoboic languages (Sarangani Manobo, Dibabawon Manobo, and Binukid) in the following tables:

Table 14: Voice markers with *paR- in Tagalog (cf. Schachter and Otanes 1972)

	Actor voice	Patient voice	Locative voice	Thematic voice
Neutral	mag-	pag- -in	pag- -an	i-pag-
Realis	nag-	p<in>ag-	p<in>ag- -an	i-p<in>ag-

Table 15: Voice markers with *paR- in Cotabato Manobo (cf. Kerr 1988)

	Actor voice	Patient voice	Locative voice	Thematic voice
Neutral	eg- /eg-/	eg- -en /eg- -en/	eg- -an /eg- -an/	eg- /eg-/
Realis	mig-	<in>	<in> -an	ig-

Table 16: Voice markers with *paR- in Sarangani Manobo (cf. DuBois 1976)

	Actor voice	Patient voice	Locative voice	Thematic voice
Neutral	meg- /məg-/	peg- -en /pəg- -ən/	peg- -an /pəg- -an/	i-peg- /i-pəg-/
Realis	mig-	pig-	pig- -an	im-peg- /im-pəg-/

Table 17: Voice markers with *paR- in Dibabawon Manobo (cf. Forster 1983:47 fn.)

	Actor voice	Patient voice	Locative voice	Thematic voice
Neutral	og-	og- -on	og- -an	ig-
Realis	nig-/mig-	pig-	pig- -an	in-

Table 18: Voice markers with *paR- in Binukid (cf. Post and Gardner 1992)

	Actor voice	Patient voice	Locative voice	Thematic voice
Neutral	mag-	pag- -en	pag- -an	ig-
Realis	mig-	pig-	pig- -an	iN-

Recalling that 1) the prefix *maR- is employed to form actor voice constructions, and that 2) the prefix *paR- is used to form undergoer voice constructions, the following hypotheses can be suggested. First, the form *eg-se- -ay* originates from the Pre-Cotabato-Manobo form *meg-se- (-ay)[57] with the form *meg-* that reflects the actor voice prefix *maR-. Second, the forms in undergoer voice constructions can be traced back to the Pre-Cotabato-Manobo forms that consist of *peg-se-

57 It is uncertain whether *-ay can be reconstructed.

and an undergoer voice marker, such as *peg-se- -en, *peg-se- -an, and *i-peg-se-.

Based on these hypotheses, we can reconstruct the following Pre-Cotabato-Manobo forms with the initial consonants *m and *p that distinguish the actor and undergoer voice forms.

Table 19: Pre-Cotabato-Manobo formations with *si-

	Actor voice	Undergoer voice		
		Patient voice	Locative voice	Thematic voice
Neutral	*meg-se- (-ay)	*peg-se-	*peg-se- -an	*i-peg-se-
Realis	*m<in>eg-se- (-ay)	*p<in>eg-se-	*p<in>eg-se- -an	*i-p<in>eg-se-

Furthermore, the Manoboic data in the tables mentioned above demonstrate that the reflexes of the realis infix *<in> appear as the vowel *i in realis forms in Manoboic languages, as in Cotabato Manobo *mig-*. This fact suggests that the proto-phoneme *i in the prefix *si- is expected to be reflected as the Proto-Manoboic phoneme *i.

If we consider the vowel changes and the reflex of the proto-phoneme *i in affixes in these Manoboic languages, we can suggest the Proto-Manoboic actor voice neutral mood form *mag-si- (-ay) and other inflected forms:

Table 20: Proto-Manoboic formations with *si-

	Actor voice	Undergoer voice		
		Patient voice	Locative voice	Thematic voice
Neutral	*mag-si- (-ay)	*pag-si- -en	*pag-si- -an	*i-pag-si-
Realis	*m<in>ag-si- (-ay)	*p<in>ag-si-	*p<in>ag-si- -an	*i-p<in>ag-si-

These formations, in turn, suggest the following Proto-Malayo-Polynesian paradigm with *si-.[58]

Table 21: The Proto-Malayo-Polynesian formations with *si-

	Actor voice	Undergoer voice		
		Patient voice	Locative voice	Thematic voice
Neutral	*maR-si-	*paR-si- -en	*paR-si- -an	*i-paR-si-
Realis	*m<in>aR-si-	*p<in>aR-si-	*p<in>aR-si- -an	*i-p<in>aR-si-

It should be noted, however, that most of these forms are not apparent in Western Indonesian and Sulawesi languages. In particular, the majority of such languages lack the overt reflexes of the undergoer voice markers, such as *-en, *-an, and *i-, as inflectional markers. As a result, the actor and undergoer voice distinction is manifested only by means of the initial consonants, such as *m and *p:

58 This paradigm may require further evidence of all assumed formations; however, it is nevertheless adequately supported by the present discussion.

Table 22: Tentative reconstruction of a Post-Proto-Malayo-Polynesian voice paradigm with *paR- in Western Indonesian and Sulawesi languages (cf. Ross 2002)

	Actor voice	Undergoer voice
Neutral	*maR-	*paR-
Realis	*m<in>aR-	*p<in>aR-

In such languages, the *si- forms are assumed to reflect the following forms:

Table 23: Tentative reconstruction of Post-Proto-Malayo-Polynesian formations with *si- in Western Indonesian and Sulawesi languages

	Actor voice	Undergoer voice
Neutral	*maR-si-	*paR-si-
Realis	*m<in>aR-si-	*p<in>aR-si-

Based on these hypothesized basic morphological structures of formations with the prefix *si-, the following sections will identify *si- forms with action-denoting roots and elucidate their formal and functional features. The discussion will be structured along the following formations:

- *maR-si- actor voice forms that indicate either actor plural, reciprocal, or reflexive situations (Section 3.2)
- *paR-si- forms in undergoer voice constructions (Section 3.3)

3.2 Formations with *maR-si- with action-denoting roots in actor voice constructions

This section concerns forms that appear with action-denoting roots and reflect the form *maR-si-. The forms in question are found in a large number of WISP languages and exhibit formal diversity, as shown in the following table:

Table 24: *maR-si- forms with action-denoting roots[59]

Language	Subgroup	Form	Meaning
Ilocano	Philippine, Northern Luzon	ag-si-	Actor plural
Pangasinan	Philippine, Northern Luzon	mag-si-	Distributive actor plural
		mag-si- -an	Successive distributive actor plural
Tagalog	Philippine, GCP, Central Philippine	mag-si-([pag-/paN-])	Actor plural
Cebuano	Philippine, GCP, Central Philippine, Bisayan	(mag-)(i)sig-([pa-])	Distributive actor plural
Aklanon	Philippine, GCP, Central	mag-si(g)-	Distributive actor plural

59 Kitada (2015) argues that the Suwawa sociative-progressive form *gi- -a* originates from the *si- form *maR-si- and the "pluractional" suffix *-an, which corresponds to the 'event plural' suffix *-an in the current study. However, since the form lacks the segment /si-/ or /sV-/, it is not considered in this study.

	Philippine, Bisayan		
Cotabato Manobo	Philippine, GCP, Manoboic	eg-se- -ay	Actor plural Reciprocal
Mongondow	Philippine, GCP, Gorontalo-Mongondowic	mo-si-	Actor plural
Sangir	Philippine, Sangiric	ma-siq-([paN-])	Distributive actor plural
Pendau	Celebic, Tomini-Tolitoli	mo-si-	Actor plural Reciprocal Comitative
Da'a	Celebic, Kaili-Pamona	mo-si-([pe-])	Actor plural Reciprocal
Muna	Celebic, Muna-Buton	si- -ha	Actor dual
		si-CVCV- -ha	Actor plural
Makassarese	South Sulawesi	aC-si- (= assi-), si-	Reciprocal Actor plural
West Coast Bajau	Smith's Western Indonesian, Barito, Sama-Bajaw	si-	Reciprocal
Kiput	Smith's Western Indonesian, Berawan-Lower Baram	se-, se-p-, se-l-	Reciprocal Reflexive
Malay	Smith's Western Indonesian, Malayic	ber-si-, ber-se-	Reciprocal
Salako	Smith's Western Indonesian, Malayic	(ba-)si(N#)-	Reciprocal Reflexive
Sundanese	Smith's Western Indonesian	si-	Reciprocal Reflexive
Old Javanese	Smith's Western Indonesian	(m)a-si-	Reciprocal
Lampung	Smith's Western Indonesian	s(a)- -an	Reciprocal
Toba Batak	Sumatran, Batak	mar-si-	Reciprocal Reflexive
Karo Batak	Sumatran, Batak	si-(RDP-) -en	Reciprocal
		si-(RDP-)([N(#)-/er-]) =na	Distributive actor plural

3.2.1 Common features in *maR-si- across languages

This section deals with some phonological and morphological features that are commonly observed in many reflexes of *maR-si- across languages. These features include 1) reflexes of Proto-Malayo-

Polynesian *R, 2) loss of initial segments, 3) vowel changes, 4) the change of /m/ to /b/ in *maR-, 5) additional prefix-final consonants after *si-, and 6) the omission of *maR-.

3.2.1.1 Reflexes of the Proto-Malayo-Polynesian phoneme *R in *maR-

The Proto-Malayo-Polynesian phoneme *R is reflected as a variety of phonemes in its daughter languages (Blust 2013:588). It is also present in *maR- in *si- formations in actor voice constructions. Although the presence of the prefix *maR- has been recognized by some authors, such as Adelaar (1984), Ross (1995, 2002), and Blust (2003a), no comparative literature has exclusively addressed the issue of the phoneme *R in the prefix so far. The current section aims to demonstrate that despite their superficial formal differences, the final phoneme in *maR- in *maR-si- can be considered to have the same proto-phoneme.

To identify whether reflexes of *R in *maR-si- forms are 'regular' (i.e., as expected from a proto-form in a top-down manner), we compare them to those in 1) reflexes of the prefix *maR- and 2) preconsonantal position in lexical words because of the following two assumptions. First, the prefix *maR- appears often with consonant-initial roots and stems with consonant-initial derivational prefixes, such as *pa- 'causative', as in Tagalog *mag-pa-* 'causative in actor voice constructions'. Second, even if the reflexes of *R in *maR- differ before the consonant- and vowel-initial roots and stems, it is assumed that the one that appears with the consonant-initial roots and stems is more likely to be generalized as a standard reflex of *R and may replace the original reflex of *R in *maR- before vowel-initial roots and stems. This assumption is based on the fact the consonant-initial roots and stem are more frequently attested in the majority of WISP languages.

Previous comparative studies regarding the phoneme *R, such as Conant (1911), focus only on its development in word-initial position (i.e., *R/#_), intervocalic (or 'word-medial') position (i.e., *R/V_V), and word-final position (i.e., *R/_#). However, little attention has been paid to the phoneme *R in preconsonantal position (i.e., *R/_C), which is less often attested and may exhibit a slightly different phonological behavior from *R in other positions. For example, Mongondow reflects *R in word-initial, intervocalic, and word-final positions as /g/, as in *gabii* 'night' (< *Rabiqi 'ibid.'), *dugi* 'thorn' (< *duRi 'ibid.'), and *dongog* 'hear' (< *dengeR 'ibid'), respectively; however, it reflects a preconsonantal *R as a zero (i.e., /Ø/), as in *botoi*[60] 'paddle'[61] (< *beRsay 'ibid.'; not **bogtoi with /g/ as the reflex of *R).

In line with this phonological difference because of positions in lexical words, it is observed that some WISP languages reflect the proto-phoneme *R differently in lexical words and prefixes. For example, Malay reflects the proto-phoneme *R as a zero (/Ø/), as in *bubur* 'porridge' (< *buRbuR 'rice porridge') in preconsonantal position and as /r/ elsewhere in lexical words, as in *ratus* 'hundred' (< *Ratus 'ibid.'), *duri* 'thorn' (< *duRi 'ibid.') and *dengar* 'hear' (< *dengeR 'ibid'). On the other hand, it reflects the Proto-Malay-Polynesian *maR- as the following three forms: 1) *ber-* before the majority of vowel- and consonant-initial roots, such as *ber-angkat* 'leave, depart' (*angkat* 'lift, rise')[62] and *ber-lari* 'run' (< *lari* 'ibid.')[63], 2) *be-* before roots with initial /r/, as in *berenang* 'swim' (< *renang* 'ibid.') and before roots that begin with the sound combination *CerC*, as in *be-serta* 'participate' (< *serta* 'ibid.')[64], and 3) *bel-* in *bel-ajar* 'study' with the root *ajar* 'study'[65]. The point that should be considered in these data is the difference in the realizations of *R before a

60 cf. *s > t/_{a, e, u, o} (Dunnebier 1929).
61 Data from Dunnebier (1929:330).
62 Data from Echols and Shadily (1994:21).
63 Data from Echols and Shadily (1994:330–331)
64 Sneddon (1996:8). The translations of the word and root are taken from Echols and Shadily (1994:457).
65 Sneddon (1996:9). The translations of the word and root are taken from Echols and Shadily (1994:7).

consonant in lexical words and prefixes. In other words, for instance, if the same phonological principles are applied to both lexical and functional items, 1) the proto-phoneme *R in *maR- should have been realized as a zero (/Ø/) before consonant-initial roots, such as **be- as in **be-lari (instead of *ber-lari* 'run'), in the same manner as the one in *buRbuR 'rice porridge' > Malay *bubur* 'porridge', or 2) the proto-phoneme *R should have been reflected as /r/ in **burbur 'porridge' (instead of *bubur*), in the same manner as the one in *ber-lari* 'run'. However, these principles are not observed in Malay.

The following lists illustrate the correspondence between the proto-phoneme *R in *maR-si- and the preconsonantal proto-phoneme *R in other environments. For comparison, reflexes of roots with the preconsonantal *R and reflexes of *maR- are also listed. In most of the instances discussed in this section, the reflexes of *R in *maR-si- are identical to those in the prefix *maR- that is directly affixed to lexical roots. Consider:

(99) *R > g/_C:

Language	*maR-si-	*maR-	*R/_C in lexical words	
Tagalog	mag-si-	mag-	bigkís 'bundle'	(< *beRkas 'bundle')
Ilocano	ag-si-	ag-	ma-bogbóg 'to pulp'	(< *buRbuR 'rice porridge')
Aklanon	mag-si(g)-	mag-	búgsay 'paddle'	(< *beRsay 'paddle')
Cebuano	(mag-)(i)sig-	mag-	bugsáy 'oar, paddle'	(< *beRsay 'paddle')
Cotabato Manobo	eg-se- -ay	eg-	bagkes 'bundle'[66]	(< *beRkas 'bundle')

(100) *R > Ø/_C:

Language	*maR-si-	*maR-	*R/_C in lexical words	
Mongondow	mo-si-	mo(g)-	botoi[67] 'paddle'	(< *beRsay 'paddle')
Sangir	ma-siq-	meq-	bekiseq 'bundle'[68]	(< *beRkas 'bundle')
Pendau	mo-si-	mo-	m-bengi 'night'[69]	(< *beRngi 'night')
Da'a	mo-si-	mo-, me-, ma-[70]	m-bengi 'night'[71]	(< *beRngi 'night')
Old Javanese	(m)a-si-	ma-	wengi 'night'	(< *beRngi 'night')
Salako	(ba-)si(N#)-	ba(r)-	bubur 'porridge'[72]	(< *buRbuR 'rice porridge')
Malay	ber-si-, ber-se-	ber-, be-, bel-	bubur 'porridge'	(< *buRbuR 'rice porridge')

66 Kerr (1988:96).

67 cf. *s > t/_{a, e, u, o} (Dunnebier 1929).

68 Sneddon (1984:71).

69 Quick (2003:123).

70 This language has these three *m-initial actor voice monosyllabic forms with different vowels. As the data cited in van den Berg (1996) and Mead (2002) show, this pattern is widely observed in Celebic languages. Van den Berg considers the forms *ma-*, *me-*, and *mo-* in these languages to reflect his 'Proto-Celebic' forms *ma-, *me-, and *mo-, respectively. However, we have not identified whether these (proto-)forms can be traced ultimately back to 1) Proto-Malayo-Polynesian *maR- or 2) different historical sources, such as *maR- for *ma-*, *me- (/me/) for *me-*, and *mo- for *mo-*, which had different functions in Proto-Malayo-Polynesian. For future reference, however, we list all the three actor voice forms, temporarily considering that they are reflexes of *maR-.

71 Barr (1988:43).

72 Adelaar (2005c:242).

(101) *R > r/_C:

Language	*maR-si-	*maR-	*R/_C in lexical words
Toba Batak	mar-si-	mar-	borngi-n ‘night’ (< *beRngi ‘night’)

(102) *R > C (consonant assimilated to the next consonant)/_C:

Language	*maR-si-	*maR-	*R/_C in lexical words
Makassarese	(aC)-si-	aC-	bangngi ‘night’ (< *beRngi ‘night’)

As observed in these examples, the Proto-Malayo-Polynesian phoneme *R in question is reflected as /g/ in Greater Central Philippine languages except for Mongondow. In lexical words, it is reflected as a zero in Salako, Malay, and languages that belong to various microgroups in Sulawesi, and as /r/ in Toba Batak. In most languages, the reflex in lexical words is identical to those in *maR- and *maR-si-. However, some languages exhibit phonological differences, as detailed below.

In Mongondow, the prefix *mo(g)-* is realized as *mo-* before consonants, as in *mo-kali* ‘dig’[73], and *mog-* before vowels, as in *mog-irup* ‘blow’[74]. Thus, the absence of the phoneme /g/ that reflects the proto-phoneme *R in Mongondow *mo-si-* can be expected because the prefix *mo(g)-* precedes the consonant-initial prefix *si-*.

Similarly, the Salako prefix *ba(r)-* is realized as *ba-* before consonant-initial roots, as in *ba-tanang* ‘to look at each other’ (< *tanang* ‘to see O[bject]’), and before some vowel-initial roots, as in *ba-inuʔ* ‘to have a mother’ (< *inuʔ* ‘mother) (Adelaar 2005c). The prefix is realized as *bar-* with other vowel-initial roots, as in *ber-ajet* ‘to fight, to scratch each other’ (< *ajet* ‘to fight, to scratch O[bject]’) (Adelaar 2005c:42). Therefore, the zero reflex of *R in the segment *ba-* in the Salako form *(ba-)si(N#)-* is expected, if being viewed from the reflex of *maR- in the language.

Sangir reflects the prefix *maR- as *meq-* /məʔ/ with a final glottal stop; however, it reflects the one in *maR-si- as *ma-* with no final glottal stop. The reflex *ma-* in this position is as expected, if being viewed from the fact that this language reflects the phoneme *R as a zero in preconsonantal position, as in *bekiseq* ‘bundle’ (< *beRkas ‘ibid.’).

The letter C in Makassarese represents a consonant assimilated to the one that follows. When the root- or stem-initial morpheme is a vowel, however, the phoneme C is realized as a glottal stop (Jukes 2006:70) (i.e., ‘C’ > ʔ/#V_). In this language, the phoneme *R is reflected as /r/ in all environments except a preconsonantal position, as in *rakiʔ* ‘raft of wood or bamboo’ (< *Rakit ‘raft’), *rara* ‘blood’ (< (PMP *daRaq) < PAn *daRaq ‘ibid.’) and *iparaʔ* ‘sibling-in-law’ (< *hipaR ‘ibid.’). It is reflected as the same consonant as the following consonant in preconsonantal position, as in *bangngi* /baŋŋi/ ‘night’ (< *beRngi ‘ibid.’) and *butta* ‘earth, ground, land’ (< *buRtaq ‘earth, soil’). Thus, while the consonant assimilation is a regular reflex of the preconsonantal proto-phoneme *R, its reflex /ʔ/ before vowel-initial roots is unpredictable from the sound correspondences between Proto-Malayo-Polynesian and Makassarese. The reflex of *R in question is otherwise expected as **r, as in **ar-agang (cf. aC-agang /aʔ-agaŋ/ ‘have friends’ < *agang* /agaŋ/ ‘friend’ (data from Jukes 2006)).

Finally, it should be mentioned that the Pangasinan form *mag-si-* exhibits an unexpected reflex of *R, if being viewed from the reflexes of *R in lexical words and the prefix *maR-. This language reflects *R as /l/ in lexical words, as in *bolból* ‘a thin porridge of boiled rice’ (< *buRbuR ‘rice porridge’) (Pangasinan data from Benton 1971a), and as /N#/ in the prefix *maR- (i.e., maN#-) (cf. Blust 2004:90 fn.). For this reason, it is expected to reflect *maR-si- either as **mal-si- or **man-si- (< **maN#-si-). The phoneme /g/ in the form *mag-si-* suggests that the form *mag-si- may be a

73 Data from Dunnebier (1929:334).

74 Data from Dunnebier (1929:355).

loan from a language where the proto-phoneme *R is reflected as /g/, such as Tagalog (*mag-si-*) and other Central Philippine languages in which the immediate proto-form *mag-si- is used for the actor plural constructions.

3.2.1.2 Loss of the initial *m

Some authors have observed that aphaeresis, or loss of the initial segments, may occur in (a series of) affixes in WISP languages. For example, Blust (2003a) observes the gradual aphaeresis of Proto-Austronesian *Sika- 'ordinal numeral' to its Proto-Malayo-Polynesian reflex *ika-, further to *ka-, and finally to *a-, as follows:

> … although a prefix *ka- has previously been noted as a marker of ordinal numerals, the agreement of Paiwan *sika-*, Ilokano *ma-ika-*, Tagalog, Cebuano, Hanunóo *ika-* ,marker of ordinal numerals' suggests that this affix was instead PAN *Sika-, PMP *ika-, with a sporadic loss of *S seen also in several other morphemes such as PAN *Sepat, PMP *epat 'four', and PAN *Si-, PMP *i- 'Instrumental Focus'. Other languages show further sporadic erosion, as with Bontok, Kapampangan, Yakan, Tondano *ka-*, Araki *ha-*, Fijian, Pohnpeian, Marshallese, Kiribati *ka-* 'marker of ordinal numerals' (loss of *Si-), and Batad Ifugao *a-* 'marker of ordinal numerals' (loss of *Sik-).
>
> (Blust 2003a:444)

Another well-known aphaeresis in Austronesian languages is the one in the hypothetical earlier (i.e., (Pre-)Proto-Austronesian, YK) actor voice form *p<um>aR-, which is reflected as Proto-Malayo-Polynesian *maR- with the loss of the initial segment /pu/ (Ross 1995). Similarly, the Proto-Malayo-Polynesian actor voice realis form *m<in>aR- is phonologically reduced as (Post-)Proto-Malayo-Polynesian *naR- (Reid 1992). Furthermore, the actor voice form *maR- is reflected as *aR- with the loss of the initial *m in some languages, such as Ilocano *ag-* (expected **mag-).

The loss of the initial *m is also observed in some *maR-si- forms, as in the following examples. For comparison, reflexes of *maR- and the initial *m in lexical words are also listed. The absence of the initial *m is not phonologically regular because the initial *m is reflected as /m/ in lexical words in all examples employed in the current study.

(103) Loss of initial *m in maR- in *maR-si-

Language	*maR-si-	*maR-	Word-initial *m in lexical words	
Ilocano	ag-si-	ag-	matá 'eye'	(< *mata 'eye')
Cotabato Manobo	eg-se- -ay	eg-	mata 'eye'[75]	(< *mata 'eye')
Makassarese	(aC-)si- ((as-)si-)	aC-	mata 'eye'[76]	(< *mata 'eye')
Old Javanese	(m)a-si-	ma-	mata 'eye'	(< *mata 'eye')

3.2.1.3 Vowel changes

Certain vowels in affixes that are reconstructed based on major Philippine and Western Indonesian languages, such as Tagalog and Toba Batak, appear as different vowels, particularly reflexes of the Proto-Malayo-Polynesian phoneme *e /ə/. For example, the vowel in the prefix *maR- is reflected as *ber-* with the vowel *e* /ə/, despite the fact that the proto-phoneme *a is reflected as *a, as in *mata. Such reflexes are considered to be due to some phonological changes that occurred after the emergence of the Proto-Malayo-Polynesian language. Such unexpected vowels can be observed in some reflexes of the proto-form *maR-si-. For example, the vowel /i/ in *si- in maR-si- appears as a

75 Kerr (1988:109).
76 Jukes (2006:69).

schwa in Malay and Kiput:

(104) Schwa in *si- in *maR-si-

Language	*maR-si-	Word-medial *i in lexical words	
Malay	ber-se-	sakit ‘sick’	(< *sakit ‘sick’)
Kiput	se- /sə-/ se-p- /sə-p-/ se-l- /səl-/	akEt ‘illness’	(< *sakit ‘sick’)

If we assume that the original vowel was /i/, as observed in lexical words, these schwas can be explained as the result of the vowel change from *i to *e /ə/.

In Cotabato Manobo, the prefix *si- appears as *se-* /se-/. In this language, the word-medial proto-phoneme *i is reflected as /i/, as in the following examples:

(105)

Language	*maR-si-	Word-medial *i in lexical words	
Cotabato Manobo	eg-se- -ay /eg-se- -ay/	sakit ‘pain’[77]	(< *sakit ‘sick’)

In Lampung, the prefix *si- is reflected as *s(a)*-. In this language, the proto-phoneme *i is also reflected as /i/ in word-medial position:

(106)

Language	*maR-si-	Word-medial *i in lexical words	
Lampung	s(a)- -an	sakiq ‘sick’[78]	(< *sakit ‘sick’)

In a similar manner to the vowel in the prefix *si-, the vowel in *maR- also appears as the one that is not a regular reflex of *a. For example, the Malay *maR-si- form *ber-si-* has a schwa in the prefix *ber-* that reflects *maR-:

(107) Vowel change of a/ in *maR- in *maR-si-

Language	*maR-si-	*maR-	Word-medial *a in lexical words	
Malay	ber-si- /bər-si-/ ber-se- /bər-sə-/	ber- /bər-/, be- /bə/, bel- /bəl/	mata ‘eye’	(< *mata ‘eye’)

In some languages, the proto-form *maR-si- appears as *mo-si-*:

(108) *maR-si- forms with *moR-

Language	*maR-si-	*maR-	Word-medial *a in lexical words	
Mongondow	mo-si-	mo-	mata	(< *mata ‘eye’)
Pendau	mo-si-	mo-	mata[79]	(< *mata ‘eye’)
Da’a	mo-si-	mo-, me-, ma-	mata ‘eye’[80]	(< *mata ‘eye’)

In Cotabato Manobo, the prefix *maR- in the reflex of *maR-si- appears as *eg-* /eg/ with the phoneme /e/:

(109) *maR-si- forms with *meR- /meR-/

Language	*maR-si-	*maR-	Word-medial *a in lexical words	
Cotabato Manobo	eg-se- -ay /eg-se- -ay/	eg- /eg-/	mata ‘eye’[81]	(< *mata ‘eye’)

77 Kerr (1988:114).
78 Walker (1976:10).
79 Quick (2003:135).
80 Barr and Barr (1988:134).
81 Kerr (1988:109).

3.2.1.4 Phonological change of /m/ to /b/ in *maR-

Adelaar (1992a:163) reports the phonological change of /m/ to /b/ that is specific to the prefix *maR- in Malayic languages, such as Malay *ber-* /bər/ and Minangkabau *ba-*. In a similar manner to the loss of the initial *m discussed earlier, this change is phonologically unexpected because the initial *m in lexical words is reflected as /m/, as in Malay and Minangkabau *mata* 'eye' (< *mata 'ibid.'). The reflex of the Proto-Malayo-Polynesian *maR- in these languages is expected to be **ma(r)- with the initial consonant /m/.

As a rule, this change is also observed in the form *maR-si-, as in the Malay forms *ber-si-* and *ber-se-* and the Salako form *(ba-)si(N#)-* (cf. *mata 'eye' > *matà* 'ibid.'[82]).

3.2.1.5 Additional prefix-final consonants after *si-

Some *maR-si- forms exhibit the presence of consonants after the segment *si. They are /g/ in Aklanon *mag-si(g)-* and Cebuano (*mag-)(i)sig-([pa-])*, /N#/ in Salako *(ba-)si(N#-),* /p/ in Kiput *se-p-*, /l/ in Kiput *se-l-,* and a glottal stop after /si/ in Sangir *ma-siq- /masi?/*. The functions and sources of these consonants require further study.

3.2.1.6 Omission of *maR- in *maR-si-

As mentioned in Section 3.1 with regard to the Salako examples, some reflexes of *maR-si- appear to consist only of reflexes of *si-. This absence of the prefix *maR- in *si- actor voice forms may suggest that the proto-form *si- by itself should be reconstructed as a *si- form in actor voice construction, in addition to *maR-si-. However, some data demonstrate that several *si- forms are reduced types of *maR-si- forms because the function marked by the actor voice marker *maR- survives despite the formal loss of *maR-. For this reason, this phenomenon is termed 'omission', rather than 'loss'.

For example, the following Cebuano sentences illustrate the language-internal alternation between *maR-si- and *si- actor voice forms in sentence context:

Cebuano

(110) **Nag-isig**-pa-úlì ang mga táwu.
AV.RLS-SI-AD-return NOM PL person
'They each returned to their respective homes.'
(Wolff 2012 [1972]:n.p., my glossing)

(111) **Isig**-pa-úlì ang mga táwu.
AV.RLS.SI-AD-return NOM PL person
'They each returned to their respective homes.'
(Wolff 2012 [1972]:n.p., my glossing)

In both sentences, the words *nag-isig-pa-úlì* '(plural actors) each returned' and *isig-pa-úlì* '(plural actors) each returned' take the plural actor subject *ang mga táwu* '(plural) people (in the nominative case)'. For this reason, we can assume that these forms are semantically and syntactically equivalent alternatives.

Another language-internal alternation between *maR-si- and *si- is attested in Makassarese.

82 Data from Adelaar (2005c:278)

Compare the two examples below. In the first example, the form *aC-si-* (= *assi-)* appears with the first person plural inclusive absolutive pronoun *=ki'*. In contrast, the form *si-* appears with *taua ri Marusu'* 'the people of Maros' in the second example. In both examples, however, the two forms derive reciprocal words:

Makassarese

(112) **Assi**janjiki' a'lampa ammuko.
aC-si-janji=ki' aC−lampa ammuko
AV-SI-promise=1.PI.ABS AV-go tomorrow
'We promised each other ([i.e., agreed]) to go tomorrow.'
(Jukes 2006:303, glossing modified)

(113) Taua ri Marusu' **si**kanre-balei.
tau=a ri Marusu' si-kanre bale=i[83]
person=DEF LOC Maros AV.SI-eat.fish=3.ABS
'(at that time) the people of Maros ate each other like fishes.'
(Jukes 2006:309, glossing modified)

However, other languages with *si- forms without the prefix *maR- in actor voice constructions do not have alternative *maR-si- forms. For example, West Coast Bajau has the reciprocal and comitative actor form *si-*; however, it does not have **be-si- with the prefix *be-* that reflects *maR-. Similarly, the Karo Batak form *si-(RDP-)([N(#)-/er-]) =na* 'distributive actor plural' does not alternate with **er-si-(RDP-)([N(#)-/er-]) =na with the prefix *er-* as a reflex of *maR-. Consider:

<u>Reciprocal construction with a *si- form</u>

West Coast Bajau

(114) "Amun kiti **si**-temu aku jadi jomo,"
when 1.PI.NOM AV.SI-meet 1.SG.NOM become person

ling dela e.
say man that
"When we meet together, I will become a human being," said the man.'
(Miller 2007:250, glossing modified)

<u>Comitative actor construction with a *si- form</u>

West Coast Bajau

(115) Tujuan=ni supaya iyo **si**-temu engko' see'=ni.
purpose=3.SG.GEN so.that 3.SG.NOM AV.SI-meet with companion=3.SG.GEN
'His purpose was so that he could meet with his friend.'
(Miller 2007:250, glossing modified)

<u>Distributive actor plural construction with a *si- form</u>

Karo Batak

(116) Opé denga kundul, **si**ngabit-ngabitken kampuhna kita.
Opé denga kundul, si-RDP-N(#)-abit-ken kampuh=na kita.
before still sit SI-??-AD-put.on-APPL sarong=3.GEN 1.PI.NOM
'Before sitting down, we all have to wear a sarong.'
(Woollams 1996:73, glossing modified)

These reciprocal, comitative actor, and distributive actor plural constructions correspond to the

83 This enclitic cross-refers the phrase *taua ri Marosu'* 'the people of Maros'.

following *maR-si- forms with overt reflexes of *maR- in other languages:

Reciprocal construction with a *maR-si- form
Pendau
(117) Jimo **nosi**baro.
Jimo no-si-baro
3.PL.NOM AV.RLS-SI-argue
'They argued.'

(Quick 2003:298, glossing modified)

Comitative actor construction with a *maR-si- form
Pendau
(118) A'u **nosi**baro sono rapi'u.
a'u no-si-baro sono rapi='u
1.SG.NOM AV.RLS-SI-argue with spouse=1.SG.GEN
'I argued with my spouse.'

(Quick 2003:299, glossing modified)

Distributive actor plural construction with a *maR-si- form
Aklanon
(119) **Magsig**buqóɬ kamó qit serbésa.
mag-sig-buqóɬ kamó qit serbésa
AV-SI-get 2.PL.NOM GEN.INDEF beer
'Each of you get your own beer.'

(Zorc 1977:143, glossing modified)

These examples suggest that some forms that reflect the immediate proto-form *si- without *maR- in actor voice constructions are syntactically and semantically identical to the constructions with the reflexes of *maR-si-.

The omission of *maR- in *maR-si- is morphologically unexpected from a Proto-Malayo-Polynesian viewpoint because almost all languages with the actor voice form *si- have overt reflexes of *maR-. It is also phonologically unexpected from a historical perspective because 1) the initial *m and the word-medial *a are reflected as /m/ and /a/ in the lexical words *mata 'eye' and *manuk 'chicken', respectively, and 2) some languages reflect the proto-phoneme *R as a given phoneme other than a zero in reflexes of *maR-.

(120) Language	*maR-si-	*maR-	Word-nitial *m- and word-medial *a
Cebuano	(mag-)(i)sig-	mag-	matá ‘eyes’ (< *mata ‘eye’)
Makassarese	(aC)si-	aC-	mata ‘eye’[84] (< *mata ‘eye’)
Salako	(ba-)si(N#)-	ba(r)-	matà ‘eye’[85] (< *mata ‘eye’)
West Coast Bajau	si-[86]	be- /bə-/	manuk ‘chicken’[87] (< *manuk ‘chicken’)
Sundanese	si-	ba-[88]	mata ‘eye’ (< *mata ‘eye’)
Karo Batak	si-(RDP-) ([N(#)-/er-]) =na si-(RDP) -en	er- /ər-/	manuk ‘chicken’[89] (< *manuk ‘chicken’)
Lampung	s(a)- -an	ma-[90]	manuq ‘chicken’[91] (< *manuk ‘chicken’)
Muna	si- -ha	mo-, me-	manu ‘chicken’[92] (< *manuk ‘chicken’)
Kiput	se-, se-p-, se-l-	[unknown]	mateh ‘eye’[93] /matəh/ (< *mata ‘eye’)

3.2.2 Other forms and developments in *maR-si- formations

The Cebuano form *(mag-)(i)sig-* has the segment *i-* before *sig-*. However, this segment deserves further research.

In the Karo Batak distributive actor plural construction marked by *si-(RDP-)([N(#)-/er-]) =na*, the form *si-(RDP-)([N(#)-/er-])* appears with the determiner *=na* after the following argument, as in *si-n-daram pangan=na* ‘each to look for their own food’ (< *ndarami pangan* 'to look for food’) (Woollams 1996:73). This determiner also functions as the third person singular and plural genitive pronoun, as in *baju=na* ‘his coat’ and *agi=na* ‘their little brother’ (Woollams 1996:115). However, in the distributive actor plural construction *si-(RDP-)([N(#)-/er-]) =na*, the determiner does not alternate with other genitive pronouns, such as *=(n)ta* ‘1.PI.GEN’ (Woollams 1996:73). This phenomenon is in contrast to other sentences indicating the pronominal plural actors that perform on their items. For example, in the following example, the actors indicated by the first person plural inclusive actor pronoun *si=* conduct an action of bringing (*baba*) on their item *pergogo=nta* ‘our dispute’. In this construction, the possession (or relatedness) of the item *pergogo* ‘dispute’ is marked with the first person plural inclusive genitive marker *=(n)ta*:

Karo Batak

(121) Ota	**si**baba	gia	pergogo**nta**	enda	ku	bale	raja, […].
ota	si=baba	gia	pergogo=nta	enda	ku	bale	raja
let's	1.PI.AC-bring	SOFTENER	dispute=our	this	to	court	chief

‘Let’s take this dispute of ours to the court of chiefs […].’

(Woollams 1996:116, glossing modified)

84 Jukes (2006:69).
85 Adelaar (2005c:278).
86 Miller (2007:250) notes that ‘[t]he pronunciation of the “i” vowel in si- often approximates to the schwa, particularly when attaching to roots with schwa or “a” in the first syllable’.
87 Miller (2007:72).
88 Robins (1959:348).
89 [manuʔ] (Woollams 1996:127).
90 Walker (1976:28).
91 Walker (1976:44).
92 van den Berg (2013 [1989]:327).
93 Blust (2003b:74).

However, in the following sentence with the actor subject *kita* '1.PI.NOM', the undergoer *kampuh=na* 'sarong=*na*' is marked with the "third person genitive" pronoun *=na*, instead of the first person plural inclusive genitive pronoun *=(n)ta*, despite the fact that the item *kampuh* 'sarong' refers to each item possessed or employed by the plural participants encoded as *kita*:

Karo Batak

(122)	Opé	denga	kundul,	**si**ngabit-ngabitken	kampuh**na**	**kita**.
	Opé	denga	kundul,	si-RDP-N(#)-abit-ken	kampuh=na	kita.
	before	still	sit	SI-??-AD-put.on-APPL	sarong=3.GEN	1.PI.NOM

'Before sitting down, we all have to wear a sarong.'

(Woollams 1996:73, glossing modified)

The segment *s(a)-* in the Lampung reciprocal form *s(a)- -an* appears to be *s* before vowels, as in *s-anjaw-an* 'to visit each other, trade visits' (< *anjaw* 'to visit') (Walker 1976:30).

3.2.3 Additional morphemes

This section introduces additional morphemes in the formations with reflexes of *maR-si-. They are 1) the action-denoting lexical class prefixes *paR- and *paN-, 2) the event plural suffixes *-an and *-ay, and 3) reduplication.

*3.2.3.1 Action-denoting lexical class prefixes with *maR-si-*

Action-denoting lexical class prefixes, such as reflexes of *paR- and *paN-, may appear after *maR-si- to indicate lexical classes of the action-denoting roots that follow. This function suggests that these prefixes belong to the roots with which they co-occur, and not to the preceding derivational prefix that reflects *si-. For this reason, it would not be appropriate to treat them as part of the derivational formations with *maR-si-. However, they are dealt with here to enhance our understanding of the affix combinations and the affix orders of the *maR-si- formations.[94]

Tagalog provides clear instances of the lexical class prefixes in *maR-si- formations. In this language, the forms *mag-si-pag-* and *mag-si-paN-* are regularly employed for the *mag-si-* forms with action-denoting roots with the action-denoting prefix *pag-* (< *paR-) and *paN-* (< *paN-), respectively. For action-denoting roots with the actor voice infix <*um*>, only *mag-si-* is employed, without lexical class prefixes that intervene between the prefixes and the roots:

94 It should be noted that what we analyze as 'lexical classes' seems to correspond to what Ross (2015) terms 'verb classes' in Formosan languages because both deal with some groups of roots based on the bare and (historically) *<um>-affixed forms, such as 1)*Ø and *<um> and 2) *p- and *m-initial alternations. However, further investigation is required for the identification because there are some differences between our data and his. For example, while our data demonstrate the difference between *paR- and *paN- class forms, such a distinction is not attested in the data that he provides.

Tagalog

(123)	**mag-si-**kanta	‘sing (plural)’	k**<um>**anta	‘sing’
	mag-si-laki	‘grow (plural)’	l**<um>**aki	‘grow’
	mag-si-pag-aral	‘study (plural)’	**mag-**aral	‘study’
	mag-si-pag-luto	‘cook (plural)’	**mag-**luto	‘cook’
	mag-si-pang-isda	‘go fishing (plural)’	**mang-**isda	‘go fishing’
	mag-si-pangailangan	‘need (plural)	**mang**ailangan	‘need’ [kailangan ‘need’]

(Schachter and Otanes 1972:335)

In a similar manner to the Tagalog examples, Sangir also has the lexical class marking system with the following three action-denoting lexical classes: *paN-*, as in *mang-alaq* ‘will take’ (< *<um>paN- with *<um> ‘actor voice’), *peq-,* as in *meq-bera* ‘will speak’(< *<um>peq- with *<um> ‘actor voice’), and a zero, as in *k<um>âng* ‘will eat’. However, the system is marginal in the *maR-si- formation. Some roots take the lexical class prefixes in *ma-siq-* forms, as in *ma-siq-pang-inung* ‘drink simultaneously, but separately’ (cf. *mang-inung* ‘will drink’, which comprises the root *inung* ‘drink’, the lexical class prefix *paN-*, and the actor voice infix *<um>). However, other roots in the formation do not take any lexical class prefixes, as in *ma-siq-bera* ‘each speak for themselves’ (expected **ma-siq-peq-bera). With regard to the root *alaq* ‘take’, the forms *ma-siq-pang-alaq* with the lexical class prefix *paN-* and *ma-siq-alaq* without it can both be observed. The root *kâng* ‘eat’ in the *maR-si- form *ma-siq-kâng* ‘each eat’ does not take any action-denoting class prefix, nor does that in the underived form *k<um>âng* ‘will eat’. Compare:

Sangir

(124)	**ma-siq-**alaq ~ **ma-siq-pang-**alaq	‘each take for themselves’	**mang-**alaq ‘will take’
	ma-siq-pang-inung	‘drink simultaneously, but separately’	**mang-**inung ‘will drink’
	ma-siq-bera	‘each speak for themselves’	**meq-**béra ‘will speak’
	ma-siq-kâng	‘each eat’	k**<um>**âng ‘will eat’

(Adriani 1893)

Similarly, Da’a also has some *si- forms with lexical class prefixes. For example, the root *kutana* ‘ask’ takes the action-denoting lexical class prefix *pe-*, as in the following example:

Da’a

(125) … **ne-**kutana-mo kolombio, ...

<in><um>**pe-**kutana=mo	kolombio
RLS-AV-AD-ask=PERF	giant

‘the giant asked ...’

(Barr 1988b:81, glossing modified)

This lexical class prefix is also present with the form *mo-si-*:

Da’a

(126)	Ira	kana	mo-si-**pe-**kutana.
	3.PL.NOM	must	AV-SI-AD-ask

‘They must ask each other questions.’

(Barr 1988:32a, glossing modified)

However, with the root *koni* ‘eat’, which takes the lexical class prefix *paN-,* as in the first

example below, the lexical class prefix is absent with *mo-si-* (< *maR-si-): The derived form appears as *mo-si-koni*, not the expected form **mo-si-pang-goni, as in the second example below:

Da'a

(127)	Aku	**nang**-goni	loka.
	Aku	<in><um>paN-koni	loka
	1.SG.NOM	RLS-AV-AD-eat	banana
	'I ate a banana.'		

(Barr 1988a:26, glossing modified)

(128)	Naile	kita	**mo-si**-koni	tibo.
	tomorrow	1.PI.NOM	AV-SI-eat	rice
	'Tomorrow we will all eat rice together.'			

(Barr 1988a:31, glossing modified)

Similarly, the Karo Batak form *si-(RDP-)([N(#)-/er-]) =na* 'distributive actor plural' may appear with action-denoting lexical class prefixes, as in *si-er-bahan-bahan permainen=na* 'each to amuse [oneself]' (< *er-bahan permainen* 'to make a game'). In contrast to Tagalog, Sangir, and Da'a, however, lexical class prefixes in question can be observed in all words with this *si- form in the data provided by Woollams (1996:73). For example, the form *si-maba-maba* in *si-maba-maba uis cucin=na* 'each to bring their own clothes to wash' can be analyzed as having the root *baba* and the lexical class prefix *N(#)-*, if we consider the word formation of the undergoer voice form *i-baba*[95] with the undergoer voice prefix *i-*. Likewise, the form *ngabit* in *si-ngabit-ngabitken kampuh=na* 'each put their sarong on' can be assumed to consist of the root *abit* 'put on and the lexical class prefix *N(#)-*, if we consider the Toba Batak root *ábit* in *par-ábit* 'wear; dress oneself with'[96] to be a cognate.

Finally, we may add the Cebuano form (*mag-)(i)sig-([pa-])*, as in *Nag-isig-pa-úlì* '(each of plural actors) go back home' with the base *pa-úlì* 'go back'[97] as an instance of a *si- form with a lexical class prefix.

The following table summarizes the *maR-si- forms with action-denoting lexical class prefixes with action-denoting roots. It should be noted that the action-denoting lexical class prefixes in question are indicated in square brackets ([ ]) to distinguish their idiosyncratic nature from other affixes. It should also be mentioned that unless the lexical class prefixes in question are present or relevant, they are not written outside this section because they are bound more closely to roots that follow than to the prefix *si-.

95 Data from Woollams (1996:83).
96 Data from Nababan (1981:102).
97 Data from Wolff (2012 [1972]:n.p.).

Table 25: Reflexes of *maR-si- with action-denoting lexical class prefixes

Language	Subgroup	Form	Meaning
Tagalog	Philippine, GCP, Central Philippine	mag-si-([pag-/paN-])	Actor plural
Cebuano	Philippine, GCP, Central Philippine, Bisayan	(mag-)(i)sig-([pa-])	Actor plural
Sangir	Philippine, Sangiric	ma-siq-([paN-])	Actor plural
Da'a	Celebic, Kaili-Pamona	mo-si-([pe-])	Actor plural Reciprocal
Karo Batak	Sumatran, Batak	si-(RDP-)([N(#)-/er-]) =na	Distributive actor plural

*3.2.3.2 The event plural suffixes *-an and *-ay*

Our data show that the suffixes *-an and *-ay may appear in some *maR-si- formations. These suffixes, together with their combined form *-anay, are frequently attested in reciprocal, actor plural, and event plural constructions. Compare the following Malay, Cebuano, and Sarangani Manobo examples:

***-an**
Malay
Reciprocal

(129) Mereka **ber**salam**an** waktu bertemu.
mereka ber-salam-an waktu ber-temu
3.PL AV-greet-EP when AV-meet
'They shook hands when they met.'

(Sneddon 1996:107, my glossing)

Event plural
(130) **ber-**lari**-an** 'to run in all directions' lari 'to run'

(Echols and Shadily 1994:331)

(131) Para saksi mata mulai **ber**mati**an**.
para saksi mata mulai ber-mati-an
PL witness eye begin AV-die-EP
'The eye witnesses were beginning to die off.'

(Sneddon 1996:109, my glossing)

***-ay**
Cebuano
Reciprocal
(132) **Nag**lalis**ay** sila.
nag-lalis-ay sila
AV.RLS-argue-EP 3.PL.NOM
'They argued back and forth.'

(Wolff 2012 [1972]:n.p., my glossing)

Sarangani Manobo
Reciprocal

(133)	**Migpe**belaw-belaw**ay**	dan.
	mig-pe-RDP-belaw-ay	dan
	AV.RLS-AD-??-discuss-EP	3.PL.NOM

'They discussed among themselves.'

(DuBois 1976:n.p., my glossing)

***-anay**
Cebuano
Reciprocal

(134)	Nagtistis**ánay**	ang	mga	hubug	sa	tubaan.
	nag-tistis-anay	ang	manga	hubug	sa	tubaan
	AV.RLS-joke-EP	NOM	PL	drunk	LOC	toddy.stand

'The drunks are exchanging jokes at the toddy stand.'

(Wolff 2012 [1972]:n.p., my glossing)

The situations in the examples above can be interpreted as having event plurality. For example, the meaning 'to run in all directions' of the Malay example *ber-lari-an* (< *lari* 'to run') suggests more than one action of running, all of which are performed either by plural actors in diverse directions simultaneously or by a singular actor who runs first in a given direction and then in another direction. On the other hand, for instance, the Cebuano reciprocal word *nag-lalis-ay* 'argued back and forth' refers to the situation in which the action of saying words occurs more than once.

The suffix *-an or *-ay may also be present in *maR-si- forms that are used for actor plural and reciprocal constructions in some languages. If we suppose that 1) every affix has a function that does not overlap with the functions of other affixes in *maR-si- -an or *maR-si- -ay, and that 2) the voice and lexical class in *maR-si- -an and *maR-si- -ay are marked by *maR- and their participant plurality is marked with *si-, we can assume that the function of the suffixes is to indicate the event plurality of derived words. As detailed in Chapter 2, the notion of event plurality refers to a state-of-affairs that involves more than one action, as represented in the sentences *I hit repeatedly* and *We hit each other*; more than one action of hitting occurs in the situations described in the sentences. It is one of the semantic components that can encode some reciprocal and actor plural situations.

The suffixes in question that appear in *maR-si- forms are reflexes of *-an in the majority of the languages in our data, as in Pangasinan *mag-si- -an* 'successive distributive actor plural' and Lampung *s(a)- -an* 'reciprocal'. However, the Cotabato Manobo form *eg-se- -ay* includes the suffix *-ay*, which is considered to reflect the proto-form *-ay.

Furthermore, we can observe further phonological developments in the reflexes of the suffixes. For example, the segment *-en /ən/* in the Karo Batak form *si-(RDP-) -en* can be explained due to the vowel change from *a to a schwa because the form *si- -en* indicates reciprocal situations as the Lampung form *s(a)- -an* does. Because reciprocal forms often also encode actor plural situations, the Karo Batak form can be considered to be a cognate of the Pangasinan form *mag-si- -an*.

The segment *-ha* in the Muna form *si- -ha* 'actor dual' and *si-CVCV- -ha* 'actor plural' can also be considered due to regular developments from the proto-form *-an for the following two reasons. First, the actor dual and plural situations may involve plural actions, which may be marked with a reflex of *-an, *-ay, or *-anay. Second, the final *n in the segment *-an is not present because word-final closed syllables are not allowed in this language. The initial /h/ in the Muna segment *-ha* is considered to be an inserted phoneme for indicating a morpheme boundary. It is also attested in the

formation *ka- -ha* that indicates locations or time of actions denoted by stems, such as *ka-lente-ha* 'place/time of birth' (< *lente* 'be born')[98]. This form is assumed to originate from *ka- -an that derives words indicating the place or time related to the roots (cf. Blust 2003a), as in Malay *ke-raja-an* 'kingdom' (< *raja* 'king')[99] and Tae' (Celebic, Kaili-Pamona) *ka-mate-an* 'time of death' (< *mate* 'die')[100].

Cotabato Manobo *-ay* in *eg-se- -ay* has the following three phonemically conditional allomorphs: *-yay* after the vowel /i/, as in *eg-se-ikagi-yay* '(plural actors) speak each other', *-oy* after the syllable with /o/, as in *eg-se-tipon-oy* '(plural actors) gather', and *-éy* /əy/ after the syllable with a schwa, as in *eg-se-tépéd-éy* '(plural actors) sit side-by-side' (Kerr 1988:96–97).

In some languages, a *maR-si- with an event plural suffix is contrastive with a *maR-si- form without the suffix. Such distinctions can be observed 1) between Pangasinan *mag-si-* 'distributive actor plural' and *mag-si- -an* 'successive distributive actor plural' and 2) between Karo Batak *si-(RDP-)([N(#)-/er-]) =na* 'distributive actor plural' and *si-(RDP-) -en* 'reciprocal'. However, other languages only have *maR-si- forms with an event plural suffix and lack *maR-si- forms without the suffix.

The forms *maR-si- -an and *maR-si- -ay with action-denoting roots are summarized in the following table:

Table 26: *maR-si- forms with *-an or *-ay for action-denoting roots

Language	Subgroup	Form	Meaning
Pangasinan	Philippine, Northern Luzon	mag-si- -an	Successive distributive actor plural
Cotabato Manobo	Philippine, GCP, Manoboic	eg-se- -ay	Actor Plural Reciprocal
Muna	Celebic, Muna-Buton	si- -ha	Actor dual
		si-CVCV- -ha	Actor plural
Lampung	Smith's Western Indonesian	s(a)- -an	Reciprocal
Karo Batak	Sumatran, Batak	si-(RDP-) -en	Reciprocal

3.2.3.3 Reduplication

Various types of reduplicative patterns serve to indicate a variety of meanings in Austronesian languages (cf. Blust 2013:406–432). Blust (2003a) suggest that different patterns of reduplication can be reconstructed, such as Proto-Malayo-Polynesian *Ca-[101] 'deverbal instrumental noun', Proto-Oceanic *CVC- 'progressive aspect', Proto-Malayo-Polynesian full reduplication 'do [verb] in numbers, do in quantity', and full reduplication with *-an* 'simulative'. Of such reduplicative patterns, the current study excludes *si- formations with reduplicative forms that are part of an inflectional paradigm or derivational formations other than the *si- derivations, even if they are observed to appear with a *si- form. For example, we do not consider Tagalog *nag-si-si-* to be a

98 Data from van den Berg (2013 [1989]:297).

99 Data from Sneddon (1996:37).

100 Data from van der Veen (1940), quoted in Blust (2003a:448).

101 The term 'Ca- reduplication' refers to the reduplicative pattern in which the first consonant of the root is reduplicated, and followed by a fixed vowel, which is usually *a, but also *e /ə/ and *o in some languages.

different derivational formation from *nag-si-* because they are inflected forms of *nag-si-* with the progressive marker *CV-*:

Tagalog

(135) Silá'y **nagsisi**alis.
sila ay nag-CV-si-alis
3.PL.NOM INV AV.RLS-PROG-SI-leave
'They are going away.'

(Lopez 1977a [1937]:91, my glossing)

(136) Silá'y **nagsi**alis.
sila ay nag-si-alis
3.PL.NOM INV AV.RLS-SI-leave
'They went away.'

(Lopez 1977a [1937]:91, my glossing)

Some *si- forms include reduplicative patterns. For example, in Muna, the reduplicative form *si- CVCV- -ha* 'actor plural' contrasts with the non-reduplicative *si- form *si- -ha* 'actor dual':

Muna

(137) Do=**si**-filei=**ha**.
3.PL.AC-AV.SI-run.away-EP
'They (dual) ran away together'

(van den Berg 2013 [1989]:319, my glossing)

(138) Ta=**si**-fuma-fumaa-**ha**.
ta=si-CVCV-fumaa-ha
1.PE.AC-AV.SI-PL-eat-EP
'We ([plural] exclusive) eat together'

(van den Berg 2013 [1989]:319, my glossing)

In Karo Batak, both *maR-si- forms *si-(RDP-) -en* 'reciprocal' and *si-(RDP-)([N(#)-/er-]) =na* 'distributive actor plural' may appear with reduplication. Woollams (1996:72) states that the Karo Batak *si- forms with reduplication '[u]sually, but not invariably, […] have a more durative or stative meaning than an eventive one'. However, the function of the reduplication in this formation requires further study due to the lack of any contrastive examples with the same root.

The *maR-si- forms with reduplicative forms mentioned above are summarized in the following table:

Table 27: Formations of *maR-si- and reduplication with action-denoting roots

Language	Subgroup	Form	Meaning
Muna	Celebic, Muna-Buton	si-CVCV- -ha	Actor plural
Karo Batak	Sumatran, Batak	si-(RDP-)[(N(#)-/er-)] =na	Distributive actor plural

3.2.4 Functional variations of *maR-si- with action-denoting roots

This section discusses functional variations of *maR-si- with action-denoting roots. Our data will demonstrate that reflexes of the form *maR-si- derive 1) actor plural, 2) reciprocal, 3) reflexive, and 4) comitative actor constructions from this type of roots. Prior to proceeding any further, however, it is advisable to state the terminology and criteria for the identification and classification of the relevant constructions.

As discussed in Chapter 2, we consider that reciprocal, actor plural, and reflexive situations can be distinguished by the presence and absence of the two semantic properties, namely, 1) the plurality of actors and 2) the actor-cum-undergoer role.

Table 28: Relations among reciprocal, actor plural, and reflexive situations

	Actor plurality	Actor-cum-undergoer role
Reciprocal	Yes	Yes
Actor plural	Yes	No
Reflexive	No[102]	Yes

Of these three situations, reciprocal and actor plural situations can be encoded in comitative constructions in which plural actors in the situations are expressed in subject and comitative functions. However, our data lack examples of non-reciprocal comitative constructions. Reciprocal comitative constructions are treated as 'comitative actor constructions' because the syntactic difference with regard to encoding the plural actors is more relevant than is the semantic one between reciprocal and actor plural situations when they are expressed in the comitative constructions. The categories and their distinctions employed in the current chapter are summarized as in the following table:

Table 29: Relations among reciprocal, actor plural, reflexive, and comitative actor constructions for examining the WISP data in this study

Constructions	Argument structure of the actor role	Actor plurality	Actor-cum-undergoer role
Reciprocal	One argument	Yes	Yes
Actor plural	One argument	Yes	No
Reflexive	One argument	No	Yes
Comitative actor Reciprocal comitative actor	Two arguments	Yes	No (comitative actor) Yes (reciprocal comitative actor)

The following sections provide detailed descriptions of the functions of *maR-si- forms. We begin with actor plural constructions in actor voice constructions because they exhibit the most widespread geographical distribution.

102Reflexive situations in which plural actors perform actions to themselves, such as *They wash themselves,* are not considered here.

3.2.4.1 Actor plural constructions in actor voice

Some *maR-si- forms denote actor plural situations. For example, the Tagalog form *mag-si-* (realis *nag-si-*) derives actor plural words from action-denoting roots. In the following examples, the *maR-si- words *nag-si-alis* '(plural actors) went away' and *nag-si-pag-dasál* '(plural actors) were playing' take plural actor subjects *silá* '3.PL.NOM' and *ang matatandá* 'the old folks (NOM)', respectively:

Tagalog

(139) Silá'y **nagsi**alis.
sila ay nag-si-alis
3.PL.NOM INV AV.RLS-SI-leave
'They went away.'

(Lopez 1977a [1937]:91, my glossing)

(140) Ang matatandá'y **nagsi**pagdasál.
ang ma-CV-tandá ay nag-si-pag-dasál
NOM AV.QD-PL-old INV AV.RLS-SI-AD-play
'The old folks were playing.'

(Lopez 1977a [1937]:91–92, my glossing)

Mongondow has the actor plural form *mo-si-*. In the following example, the word *mo-si-lituq=makow* '(plural actors) be sitting there' appears with the third person plural nominative pronoun *imosia* as the actors:

Mongondow

(141) Kinobayaqan-mai **mosi**lituq=makow kong ganderia (dangkulon)
kinobayaqan-mai mo-si-lituq=makow kong ganderia (dangkulon)
arrived-VENITIVE AV-SI-sit-ANDATIVE at gallery (front veranda)

imosia.
imosia
3.PL.NOM
'When (I) arrived, they were all sitting at the front veranda.'

(Dunnebier 1929:552, my glossing)

Pendau has the actor plural form *mo-si-* (realis *no-si-*). In the following example, the phrase *jimo toroonong moo* 'the six of them' is the subject and refers to plural entities.

Pendau

(142) Jimo toroonong moo ndaumo **nosi**otoi.
jimo to=roonong moo ndau=mo no-si-otoi
3.PL.NOM REL=six here NEG=COMPLETIVE AV.RLS-SI-know
'None of the six knew a thing.'

(Quick 2003:300, glossing modified)

Other examples with the Pendau form *mo-si-* are as follows:

Pendau

(143) **mo-si**-inum 'drink from the same glass' [inum 'drink']
mo-si-'omung 'carry [together]' ['omong 'carry']

(Quick 2003:297)

The Da'a form *mo-si-* denotes actor plural situations. In the following example, the word *mo-si-koni* takes the first person plural inclusive nominative pronoun *kita* as an actor subject:

Da'a
(144) Naile kita **mo-si**-koni tibo.
tomorrow 1.PI.NOM AV-SI-eat rice
'Tomorrow we will all eat rice together.'

(Barr 1988a:31, glossing modified)

The Cotabato Manobo actor plural *maR-si- form is *eg-se- -ay*. In the following example, the third person plural nominative pronoun *da* is the subject:

Cotabato Manobo
(145) **Eg-se**-tipon-**oy** da.
eg-se-tipon-ay da
AV-SI-gather-EP 3.PL.NOM
'They are all gathered together.' ['They all gather.']

(Kerr 1988:96, glossing modified)

Some languages have two *maR-si- actor plural forms, which differ in participant number or event number. For example, the Muna form *si- -ha* exclusively encodes actor dual situations in actor voice constructions with intransitive action-denoting roots. By contrast, actor plural constructions in actor voice with intransitive action-denoting roots are marked with *si-CVCV- -ha* in this language (see Section 3.2.4.5.1). On the other hand, Pangasinan has two *maR-si- forms, namely, *mag-si-* 'distributive actor plural' and *mag-si- -an* 'successive distributive actor plural' (see Section 3.2.4.5.2).

Finally, we may add the Ilocano form *ag-si-*, as in *ag-si-tulong* 'to help out with many people' (< *tulong* 'help') (Rubino 1997:200) to our data concerning the actor plural constructions with *maR-si- forms, although they were not attested in sentence context.

In the available data, *maR-si- forms that denote actor plural situations are found in some Philippine and Sulawesi languages, as illustrated in the following table:

Table 30: Actor plurals in actor voice with *maR-si-

Language	Subgroup	Form	Function
Ilocano	Philippine, Northern Luzon	ag-si-	Actor plural
Tagalog	Philippine, GCP, Central Philippine	mag-si-	Actor plural
Cotabato Manobo	Philippine, GCP, Manoboic	eg-se- -ay	Actor plural
Mongondow	Philippine, GCP, Gorontalo-Mongondowic	mo-si-	Actor plural
Pendau	Celebic, Tomini-Tolitoli	mo-si-	Actor plural
Da'a	Celebic, Kaili-Pamona	mo-si-([pe-])	Actor plural
Muna	Celebic, Muna-Buton	si- -ha	Actor dual
		si-CVCV- -ha	Actor plural

In addition, it should be noted that some *maR-si- forms indicate situations with plural actors with no cooperativity among them. Such situations are termed ‘distributive actor plural situations’, and constructions that encode the situations are termed ‘distributive actor plural constructions’. The absence of the cooperativity is often signaled in the translation with the word ‘each’ for actor arguments, as in ‘each person does something’.

In the discussion regarding the relations between actor plural, reciprocal, comitative actor, and reflexive constructions marked by *maR-si- in Section 3.2.4.6, the distributive actor plural construction marked with *maR-si- is treated as a subtype of the actor plural constructions because both categories encode non-reciprocal situations with plural actors. However, we describe the distributive actor plural construction separately for further study, in case the actor plural and distributive actor plural usages of *maR-si- might be traced back to different sources that should be explored further.

For example, in Aklanon, the *maR-si- forms *mag-si-* and *mag-sig-* are used to indicate ‘individual actions’ (Zorc 1977:143). In the following sentences, the subjects are the plural pronouns *kamó* and *sandá*, respectively:

Aklanon

(146)	**Magsig**buqóɬ	kamó	qit	serbésa.
	mag-sig-buqóɬ	kamó	qit	serbésa
	AV-SI-get	2.PL.NOM	GEN.INDEF	beer
	‘Each of you get your own beer.’			

(Zorc 1977:143, glossing modified)

(147)	**Nagsi**bakáɬ	sanda	it	ɬambong.
	nag-si-bakáɬ	sanda	it	ɬambong
	AV.RLS-SI-buy	3.PL.NOM	GEN.INDEF	shirt
	‘They each bought a different shirt.’			

(Zorc 1968:135, my glossing)

The Cebuano form *mag-isig-* (allomorphs: *mag-sig-* and *isig-*) is also employed to form distributive actor plural words. In the following example, the word *nag-(i)sig-pauli* ‘(each of plural actors) go home’ takes the plural actor subject *ang mga táwu* ‘the people (NOM)’:

Cebuano

(148)	**Nag-isig** paúlì	(/**nag-sig** paúlì)	ang	mga	táwu.
	AV.RLS-SI-AD-return	(AV.RLS-SI-AD-return)	NOM	PL	person
	‘They [many people] each returned to their respective homes.’				

(Wolff 2012 [1972]:n.p., my glossing)

In Pangasinan, the form *mag-si-* encodes distributive actor plural situations. In the following examples, the words *mag-si-kanta* ‘(each of plural actors) sing’ and *mag-si-salíta* ‘(each of plural actors) speak’ appear with the plural pronouns *kayó* and *irá*, respectively:

Pangasinan

(149)	**Magsi**kánta	kayó.
	mag-si-kánta	kayó
	AV-SI-sing	2.PL.NOM
	‘Each of you sing.’	

(Benton 1971b:132, my glossing)

(150) **Magsi**salíta irá.
mag-si-salíta irá
AV-SI-speak 3.PL.NOM
'They will each speak.'

(Benton 1971b:132, my glossing)

Sangir has the distributive actor plural form *ma-siq-([paN-])*. In the following example, the word *ma-siq-kâng* '(each of plural actors) eat' with the root *kâng* 'eat' takes the word *taumata* 'person' as the actor subject in the sentence, and denotes that the item referred to by the word *taumata* should be interpreted as plural (i.e., as 'people'):

Sangir
(151) Taumata e kai harusĕʔ **masiq**kâng u kange,
taumata e kai harusĕʔ ma-siq-kâng u kange
person EMPH EMPH should AV-SI-eat GEN food

tadeaqu mebiaheq.
tadeaqu me-biaheq
in.order.to AV-live
'People should each eat to live.'

(Adriani 1893:112, my glossing)

Karo Batak has the form *si-(RDP-)([N(#)-/er-]) =na* that denotes distributive actor plural situations. In the following example, the word *simaba-maba* '(each of plural actors) bring' takes the actor argument *ia peképar* 'they each' and the undergoer argument *kinibeluhen=na* 'their skill' with the enclitic *=na*:

Karo Batak
(152) **Si**maba-maba kinibeluhen**na** ia peképar.
si-RDP-N(#)-baba kinibeluhen=na ia peképar
AV.SI-??-AD-carry skill=3.GEN 3.NOM both
'They each displayed their skill.'

(Woollams 1996:179, glossing modified)

In the following example, the word *singabit-ngabitken* takes the first person plural inclusive nominative pronoun *kita* as a subject.

Karo Batak
(153) Opé denga kundul, **si**ngabit-ngabitken kampuh**na** kita.
Opé denga kundul, si-RDP-N(#)-abit-ken kampuh=na kita.
before still sit AV.SI-??-AD-put.on-APPL sarong=3.GEN 1.PI.NOM
'Before sitting down, we all have to wear a sarong.'

(Woollams 1996:73, glossing modified)

Other examples with *si-(RDP-)([N(#)-/er-]) =na* are as follows:

(154) **si**-n-daram pangan=**na** ndarami pangan ‘to look for food’
‘each to look for their own food’
si-n-dahi dahin=**na** ndahi dahin ‘to do work,
‘each to do their own work’ to attend to a job’
si-maba-maba uis cucin=**na** maba uis cucin ‘to bring clothes
‘each to bring their own clothes to wash’ for washing’
si-er-bahan-bahan permainen=**na** er-bahan permainen
‘each to amuse [oneself]’ ‘to make a game’
(Woollams 1996:73)

Our data show that *maR-si- forms that indicate distributive actor plural situations are attested in some Philippine and Western Indonesian languages, as summarized in the following table:

Table 31: Distributive actor plural constructions in actor voice with *maR-si-

Language	Subgroup	Form
Pangasinan	Philippine, Northern Luzon	mag-si-
Cebuano	Philippine, GCP, Central Philippine, Bisayan	(mag-)(i-)sig-
Aklanon	Philippine, GCP, Central Philippine, Bisayan	mag-si(g)-
Sangir	Philippine, Sangiric	ma-siq-([paN-])
Karo Batak	Sumatran, Batak	si-(RDP-)([N(#)-/er-]) =na

3.2.4.2 Reciprocal constructions

Some *maR-si- forms indicate reciprocal situations. For instance, the Pendau form *mo-si-* (realis *no-si-*) derives the word *no-si-turu’* ‘(plural actors) agree with each other (in realis mood)’ from the root *turu’* ‘obey’ (cf. *no’u-turu’-i* ‘I have obeyed (it)’ 1.SG.AG.RLS-obey-APPL (Quick 2003:383, glossing modified)). In the following example, the derived word co-occurs with the third person plural nominative pronoun *jimo* as an actor subject, and indicates a reciprocal situation:

Pendau
(155) Jimo **nosi**turu’.
jimo no-si-turu’
3.PL.NOM AV.RLS-SI-obey
‘They agreed’ or ‘They had the same opinion.’
(Quick 2003:299, glossing modified)

Another example in sentence context is the one with the reciprocal word *no-si-baro* ‘(plural actors) argued each other’ with the root *baro* ‘argue’. It should be noted that in contrast to the root *turu’* ‘obey’, this root always appears with the prefix *si-* in the data provided by Quick (2003):

Pendau
(156) Jimo **nosi**baro.
Jimo no-si-baro
3.PL.NOM AV.RLS-SI-argue
‘They argued.’
(Quick 2003:298, glossing modified)

Other reciprocal words with *mo-si-* are as follows:

Pendau

(157)	**mo-si-**sempa’	‘kick [each other]’	[sempa’ ‘kick’]
	no-si-ampuni	‘forgive [each other]’	[ampuni ‘forgive’]

(Quick 2003:297)

The Cotabato Manobo actor plural *maR-si- form *eg-se- -ay* also indicates reciprocal situations with some roots. In the following examples, all the *eg-se- -ay* forms appear with the plural pronoun *da* or *ki* as the subject:

Cotabato Manobo

(158) **Eg-se-**ikagi-**yay** da.
AV-SI-say-EP 3.PL.NOM
‘They are speaking to each other.’

(Kerr 1988:97, glossing modified)

(159) **Eg-se-**tinudu-**ay** da=’t belad.
AV-SI-point.at-EP 3.PL.NOM=LK hands
‘They point at each other.’

(Kerr 1988:97, glossing modified)

(160) **Eg-se-**tabang-**ay** ki <um>tudak katila.
AV-SI-help-EP 1.PI.NOM AV.FUT-plant sweet.potato
‘We help each other plant sweet potato[es].’

(Kerr 1988:97, glossing modified)

The Da’a form *mo-si-([pe-])* also derives reciprocal words. In the following sentence, the subject is the third person plural nominative pronoun *ira*:

Da’a

(161) Ira kana **mo-si-**pe-kutana.
3.PL.NOM must AV-SI-AD-ask
‘They must ask each other questions.’

(Barr 1988a:32, glossing modified)

Karo Batak has the form *si-(RDP)- -en*. This form may derive both intransitive and transitive words. In the next example, s*i-alo-alo-n* ‘(plural actors) followed each other (lit. welcome/greet each other)’ takes the plural actor subject *udan ras kilap* ‘rain and lightning’:

Karo Batak

(162) Udan ras kilap **sialo-**alo**n**.
udan ras kilap si-RDP-alo-en
rain and lightning AV.SI-??-welcome/greet-EP
‘Rain and lightning followed each other.’

(Woollams 1996:178, glossing modified)

Other intransitive words with *si-(RDP)- -en* are as follows:

Karo Batak

(163)	**si**-tatap-**en**	‘look at each other’	tatap	‘look, gaze at’
	si-buat-**en**	‘marry with each other’	buat	‘take’
	si-tanda-**n**	‘know each other’	tanda-i[103]	‘know (a person)’
	si-rawa-**n**	‘scold each other’	rawa-i	‘scold, angry at’
	si-antus-**en**	‘understand each other’	antus-i	‘understand’

(Woollams 1996:71–2)

(164)	**si**-kirim-kirim-**en**	‘correspond with each other’	kirim	‘send’
	si-jerleng-jerleng-**en**	‘stare at each other’	jerleng	‘stare’
	si-sampat-sampat-**en**	‘help each other’	sampat-i	‘help’

(Woollams 1996:72)

The following examples are transitive words with *si-(RDP)- -en*:

Karo Batak

(165) **si**-tanda-**n** rupa
‘to know each other by appearance’
si-beteh-**en** orat tutur
‘to know what kinship terms by which to address each other’

tandai rupa
‘to know, be acquainted with a face’
(b)eteh orat tutur
‘to know what form of address to use to another person’

(Woollams 1996:72)

(166) **si**-benter-benter-**en** nakan
‘to throw rice at each other’

benterken nakan ‘to throw rice’

(Woollams 1996:72)

In West Coast Bajau, the prefix *si-* forms reciprocal words. In the following example, the subject of the word *si-temu* ‘(plural actors) meet (each other)’ is the first person plural inclusive nominative pronoun *kiti*:

West Coast Bajau

(167)	“Amun	kiti	**si**-temu	aku	jadi	jomo,”
	when	1.PI.NOM	AV.SI-meet	1.SG.NOM	become	person
	ling	dela	e.			
	say	man	that			

“When we meet together, I will become a human being,” said the man.’

(Miller 2007:250, glossing modified)

Other reciprocal words with *si-* are as follows:

103It should be noted that the suffix *-i* is a derivational marker that ‘operates primarily on adjectives, intransitive verbs, and nouns, to form transitive verb stems with a locative meaning’ (Woollams 1996:56) and is not present in the forms with *si- -en*.

West Coast Bajau

(168)	**si**-bangga’	‘to meet each other suddenly’	bangga’	‘to meet suddenly’
	si-bara’	‘to tell each other’	bara’	‘to tell’
	si-enda’	‘to look at each other’	enda’	‘to look at’
	si-tumbuk	‘to punch each other’	tumbuk	‘to punch’

(Miller 2007:251)

Makassarese has two reciprocal forms, namely, *aC-si-* (*assi-*) and *si-*, as discussed in Section 3.2.1.6. In the following example, the plural entities indicated by the pronominal enclitic *=ki’* (i.e., ‘we’) serve as the actors of the action *assijanji* ‘promise each other’:

Makassarese

(169)	Assijanjiki’	a’lampa	ammuko.
	aC-si-janji=ki’	aC-lampa	ammuko
	AV-SI-promise=1.PI.ABS	AV-go	tomorrow

‘We promised each other ([i.e., agreed]) to go tomorrow.’

(Jukes 2006:303, glossing modified)

Finally, examples of *maR-si- reciprocal forms in other languages are listed below in isolation. The argument structures for these forms are unknown because no data in sentence context are available in these languages. It may be the case that some forms are also employed to indicate a comitative actor construction with plural actors in two discontinuous arguments (i.e., a subject and a non-subject argument). However, it is unreasonable to assume that some forms do not allow plural actor subjects because the term ‘reciprocal’ usually presupposes a construction with plural actors in subject function as a primary syntactic structure (cf. Nedjalkov 2007a; Lichtenberk 2000; Kemmer 1993). For this reason, they are treated here. Consider:

Malay

(170)	**bər-sə-**tumpu	‘to take off against each other (for instance in tug-of-war game)’	tumpu	‘take-off, abutment’
	bər-sə-kongkol	‘to conspire’	kongkol	‘to discuss, gossip’

(Adelaar 2005a:132)

Toba Batak

(171)	**mar-si**-gulut	‘to quarrel with each other’	[gulut (mar-gulut) ‘argue about something’[104]]

(van der Tuuk 1971:140)

(172)	**mar-si**-adu	‘run in a race; be in a competition’	adu	‘chase; overtake’

(Nababan 1981:99)

Sundanese

(173)	**si**-dakeup	‘to hug each other’	dakeup	[‘hug’ cf. Malay dakap ‘hug’]

(Coolsma 1873:26)

104Warneck (1977 [1906]:97).

Old Javanese

(174) **(m)a-si-**ḍakĕp ‘hug; hold each other in their arms’ [ḍakĕp ‘hug, embrace’]

(Kern 1899:403)

Salako

(175) **sing**-karumak ‘to claw at each other, scratch each other.’ ngarumak [N-karumak][105] ‘to claw, scratch’

(Adelaar 2005c:44)

Kiput

(176) **se-**bukUt	‘punch each other’	mukUt	‘to punch’ (< bukUt)
se-deñek	‘squeeze each other’	duñek	‘to squeeze’
se-p-abit	‘hold each other (coarse)’	m-abit	‘to hold’ (< abit)
se-l-abit	‘hold each other (refined)’	m-abit	‘to hold’ (< abit)
se-l-adek	‘kiss each other’	m-adek	‘to kiss’ (< adek)

(Blust 2003b:18)

Lampung

(177) **sa-**sium-**an**	‘to kiss each other’	sium	‘to kiss’
sa-segoq-**an**	‘to play “hide and seek”’	segoq	‘to hide’
s-anjaw-**an**	‘to visit each other, trade visits’	anjaw	‘to visit’

(Walker 1976:30)

The reciprocal usage of *maR-si- forms is attested in some languages in all Western Indonesian, Sulawesi, and Philippine areas, particularly in the former two areas. The *maR-si- forms are summarized in the following table:

Table 32: Reciprocal constructions with *maR-si-

Language	Subgroup	Form
Cotabato Manobo	Philippine, GCP, Manoboic	eg-se- -ay
Pendau	Celebic, Tomini-Tolitoli	mo-si-
Da'a	Celebic, Kaili-Pamona	mo-si-([pe-])
Makassarese	South Sulawesi	(aC-)si-
West Coast Bajau	Smith’s Western Indonesian, Barito, Sama-Bajaw	si-
Kiput	Smith’s Western Indonesian, Berawan-Lower Baram	se-, se-p-, se-l-
Malay	Smith’s Western Indonesian, Malayic	ber-si-, ber-se-
Salako	Smith’s Western Indonesian, Malayic	(ba-)si(N#)-
Sundanese	Smith’s Western Indonesian	si-
Old Javanese	Smith’s Western Indonesian	(m)a-si-

105The prefix *N-* indicates that the verb is intransitive (Adelaar 2005c:43).

Lampung	Smith's Western Indonesian	s(a)- -an
Toba Batak	Sumatran, Batak	mar-si-
Karo Batak	Sumatran, Batak	si-(RDP-) -en

3.2.4.3 Reflexive constructions

Some *maR-si- forms are employed to indicate reflexive situations. Consider the following Salako example:

Salako
(178) … ià **sin**-soor ka tanàh, nangis.
… ià sin-soor ka tanàh, N-tangis.
… 3 AV.SI-throw LOC floor AV-cry
'(If he asked permission to go to the earth and he did not get it,)
he would throw himself to the ground and cry.'
(Adelaar 2005c:88, glossing modified)

In this sentence, the word *sin-soor* 'to throw oneself' appears with the third person pronoun *ià* as the actor subject in actor voice.

Other *maR-si- forms indicating reflexive situations are as follows:

Toba Batak
(179) **mar-si**-jongjong 'to raise oneself up' jongjong 'to stand'
(van der Tuuk 1971:140)

(180) **mar-si**-ájar 'study [teach oneself]' ájar 'teach ([transitive])'
(Nababan 1981:99)

Kiput
(181) **se**-kelai 'to cool oneself/each other' ngelai [N-kelai][106] 'to cool sth'
se-pidië 'to hang oneself' midië [N-pidië] 'to hang sth up'
se-p-ataay 'commit suicide' m-ataay [N-pataay] 'to die'
se-l-ibet 'turn oneself' m-ibet [N-ibet] 'to turn'
(Blust 2003b:18–19)

Salako
(182) **sin**-soor 'to throw oneself to the floor out of recalcitrance, roll oneself over the floor' soor '[throw, creep]' cf. nyoor (N-soor) 'to grow above ground (root); to creep over the ground (root, snake, child)'
ba-sing-koʔopm 'to put entirely into one's mouth' koʔopm ['gloss unknown']
(Adelaar 2005c:43, 306)

106The prefix *N-* (or *ng-*) indicates actor voice (cf. Blust 2003b:10).

Sundanese

(183)	**si**-duru	‘warm oneself at the fire’	duru	[‘gloss unknown’]

(Kern 1899:402)

It should be noted that some object-denoting roots may also appear with *maR-si- forms to encode reflexive actions that describe the act of sanitizing oneself. Two types of the roots can be identified for this construction: One is those that indicate body parts, such as ‘face’, and the other is those indicating tools for washing, such as ‘water’. Despite the fact that these are not action-denoting roots, they are treated here because their derived words with *maR-si- forms encode reflexive situations in the same manner as action-denoting roots. Consider the following examples in Salako and Sundanese:

Salako

(184)	**ba-si**-muhà	‘to wash one’s face’	muhà	‘face’
	ba-sing-komor	‘to rinse one’s mouth’	komor	[‘gloss unknown’]

(Adelaar 2005c:44)

Sundanese

(185)	**si**-beungeut	‘to wash the face’	beungeut	‘face’
	si-banyu	‘to wash the hands’	banyu	‘water’

(Coolsma 1873:26)

Reflexive *maR-si- forms are attested only in Western Indonesian languages in our data. These forms also indicate reciprocal situations.

Table 33: Reflexive constructions with *maR-si-

Language	Subgroup	Form	Notes
Kiput	Smith’s Western Indonesian, Berawan-Lower Baram	se-, se-p-, se-l-	polysemous with reciprocals
Salako	Smith’s Western Indonesian, Malayic	(ba-)si(N#)-	polysemous with reciprocals
Sundanese	Smith’s Western Indonesian	si-	polysemous with reciprocals
Toba Batak	Sumatran, Batak	mar-si-	polysemous with reciprocals

3.2.4.4 Comitative actor constructions in actor voice

Some *maR-si- words form comitative actor constructions in actor voice constructions. For example, the Pendau actor plural and reciprocal marker *mo-si-* may also appear with a singular subject with its comitative actor argument:

Pendau

(186) Odo **nosi**baro sono ulasang.
odo no-si-baro sono ulasang
monkey AV.RLS-SI-argue COM turtle
'The monkey quarreled with the turtle.'

(Quick 2003:298, glossing modified)

In this example, the actor plural/reciprocal form *mo-si-* (realis *no-si-*) allows the subject encoding the singular entity *odo* 'monkey' due to the presence of its comitative actor argument *sono ulasang* 'with the turtle'.

This comitative actor function is assumed to be related to the reciprocal construction with a plural actor subject indicated by a coordinate phrase with the comitative marker *sono* 'with'. In the following example, the two object-denoting words *odo* 'monkey' and *ulasang* 'turtle' are conjoined with the comitative marker *sono* and form a plural actor argument as the subject of the word *no-si-baro* 'have argued'. All the components in the sentence are identical to the example in the comitative actor construction mentioned above; however, the sentence has a different syntactic structure:

Pendau

(187) Odo sono ulasang **nosi**baro.
odo sono ulasang no-si-baro
monkey with turtle AV.RLS-SI-argue
'The monkey argued with the turtle.'

(Quick 2003:298, glossing modified)

Furthermore, the Pendau form *mo-si-* may take singular subjects as actors even without comitative actor arguments. In this case, the sentence implies that there is another comitative actor that performs the action denoted by the root, with the singular actor encoded as the subject. Compare the following examples:

Pendau

(188) A'u **nosi**baro sono rapi'u.
a'u no-si-baro sono rapi='u
1.SG.NOM AV.RLS-SI-argue with spouse=1.SG.GEN
'I argued with my spouse.'

(Quick 2003:299, glossing modified)

(189) A'u **nosi**baro.
a'u no-si-baro
1.SG.NOM AV.RLS-SI-argue
'I argued (with someone).'

(Quick 2003:298, glossing modified)

In the first example, the word *no-si-baro* '(plural actors) argued' appears with the first person singular nominative pronoun *a'u* as the actor in subject function and its comitative actor argument s*ono rapi'u* 'with my spouse'. However, in the second example, the absence of the comitative actor argument implies that the action of *no-si-baro* is performed with someone else who is not present in the sentence:

Another instance of comitative actor constructions is found in West Coast Bajau. In this language, the reciprocal form *si-* is also employed as a comitative actor marker. In the following

example, the word *si-temu* ‘meet’ appears with the third person singular nominative pronoun *iyo* as an actor subject and the comitative actor argument *engko’ see’=ni* ‘with his friend’:

West Coast Bajau
(190) Tujuan=ni supaya iyo **si**-temu
purpose=3.SG.GEN so.that 3.SG.NOM AV.SI-meet

engko’ see’=ni.
with companion=3.SG.GEN
‘His purpose was so that he could meet with his friend.’
(Miller 2007:250, glossing modified)

All these Pendau and West Coast Bajau examples represent reciprocal comitative actor constructions that encode reciprocal situations, and no comitative actor constructions that encode actor plural situations were found in our data. However, these examples demonstrate that reflexes of *maR-si- may encode plural actors in two arguments in a different manner from reciprocal and actor plural constructions in which the plural actors are expressed in subject function.

Table 34: Comitative actor constructions in actor voice with *maR-si-

Language	Subgroup	Form	Notes
Pendau	Celebic, Tomini-Tolitoli	mo-si-	polysemous with actor plural and reciprocal marker
West Coast Bajau	Smith’s Western Indonesian, Barito, Sama-Bajaw	si-	polysemous with reciprocal marker

3.2.4.5 Number distinction with additional morphemes

This section introduces formations with *maR-si- and *maR-si- with additional morphemes that are different with respect to participant and event numbers.

3.2.4.5.1 Participant number

Muna distinguishes between the categories that denote dual and plural actors with a *maR-si- form and a *maR-si- with CVCV- reduplication. The form *si- -ha* is employed to indicate dual actors in actor voice constructions. In contrast, the form *si-CVCV- -ha* with the CVCV- reduplication is used to indicate non-dual plural actors in actor voice constructions. Compare:

Muna
(191) Do=**si**-filei=**ha**.
3.PL.AC-AV.SI-run.away-EP
‘They (dual) ran away together.’
(van den Berg 2013 [1989]:319, my glossing)

(192) Ta=**si-**fuma-fumaa-**ha**.
ta=si-CVCV-fumaa-ha
1.PE.AC-AV.SI-PL-eat-EP
'We ([plural] exclusive) eat together.'

(van den Berg 2013 [1989]:319, my glossing)

3.2.4.5.2 Event number

Pangasinan has the form *mag-si- -an*, in addition to *mag-si-*. While the form *mag-si-* derives words indicating distributive actor plural situations, *mag-si- -an* derives words denoting successive distributive actor plural situations, or situations in which each of the plural actors performs actions in turn:

Pangasinan
(193) **Magsi**kánta kayó.
mag-si-kánta kayó
AV-SI-sing 2.PL.NOM
'Each of you sing.'

(Benton 1971b:132, my glossing)

(194) **Magsi**kanta**án** kayó.
mag-si-kanta-án kayó
AV-SI-sing-EP 2.PL.NOM
'You should each sing in turn.'

(Benton 1971b:132, my glossing)

Due to the fact that turn-taking is considered to involve plural actions, the suffix *-an* in the formation *mag-si- -an* can be identified as an event plural marker, in a similar manner to Malay *ber- -an* and Cebuano *mag- -ay.*

*3.2.4.6 Functional patterns of *maR-si- in actor voice*

As we have seen in the previous sections, various situations and constructions are marked with *maR-si- forms in actor voice constructions. They are 1) actor plural, 2) distributive actor plural, 3) reciprocal, 4) reflexive, 5) comitative, 6) actor dual, and 7) successive distributive actor plural situations and constructions. Of these functions, the "successive" meaning of the successive distributive actor plural function can be attributed to the event plural suffix *-an* in Pangasinan because the turn-taking meaning of the distributive situations is attributable to the plurality of events. Furthermore, the *maR-si- reflex *mag-si-* is used to indicate actor plural constructions in the language. Similarly, the non-dual plural meaning of the Muna actor plural form *si-CVCV- -ha* is attributable to reduplication because the Muna *maR-si- reflex *si- -ha* denotes actor dual constructions. Further investigation is required to determine why *maR-si- forms without reduplication indicate actor plural constructions in some languages, and actor dual constructions in others. Similarly, further study is required of *maR-si- forms with distributive meaning in some languages. However, we consider the actor dual and distributive actor plural situations and constructions to be variants of actor plural ones because non-singular actors do not act on each other, and the actors are encoded in subject function. Therefore, we only consider four types of constructions as having the functions marked with *maR-si- forms, namely, 1) actor plural, 2) reciprocal, 3) reflexive, and 4) comitative actor.

Based on these functional classifications, we identify the following distributional patterns of *maR-si- functions for investigating the geographical distribution of the *maR-si- forms with action-denoting roots: 1) actor plural only, 2) reciprocal only, 3) actor plural and reciprocal, 4) actor plural, reciprocal, and comitative actor, and 5) reciprocal and reflexive.

Forms with reflexes of *maR-si- that only have the actor plural function are found in all the Western Indonesian, Sulawesi, and Philippine areas. The Philippine region has the most languages of this type. For example, the following forms are found in Greater Central Philippine languages: Tagalog *mag-si-([pag-/paN-])*, Cebuano (*mag-)(i)sig-*, Aklanon *mag-si(g)-*, and Mongondow *mo-si-*. Forms with reflexes of *maR-si- also only have the actor plural function in some other Philippine languages: Ilocano *ag-si-*, Pangasinan *mag-si-* and *mag-si- -an*, Sangir *ma-siq-([paN-])*. This functional type of *maR-si- form is also found in the Sulawesi area, such as Muna *si- -ha* 'actor dual' and *si-CVCV- -ha* 'actor plural', as well as in the Western Indonesian area, such as Karo Batak *si-(RDP-)[(N(#)-/er-)] =na* 'distributive actor plural'.

Forms with reflexes of *maR-si- with only the reciprocal function are found only in Western Indonesian languages. They are Malay *ber-si-* and *ber-se-*, Old Javanese *(m)a-si-*, Lampung *s(a)- -an*, and Karo Batak *si-(RDP)- -en*.

The actor plural-reciprocal polysemy of *maR-si- forms is found in Philippine and Sulawesi languages: Cotabato Manobo *eg-se- -ay*, Da'a *mo-si-([pe-])*, and Makassarese *(aC)si-*.

With comitatives, two types of polysemies have been found in Sulawesi and Western Indonesian languages: Pendau *mo-si-* exhibits actor plural, reciprocal, and comitative actor polysemy, and West Coast Bajau *si-* demonstrates reciprocal and comitative actor polysemy.

The reciprocal-reflexive polysemy is found exclusively in Western Indonesian languages: Kiput *se-, se-p-*, *se-l-*, Salako *(ba-)si(N#)-*, Sundanese *si-*, and Toba Batak *mar-si-*.

Having seen how the four major meanings co-occur geographically, the question arises as to how they are related. In spite of the variety of the meanings, however, the current data cannot identify a form that is specific to a given function. In other words, for example, our data do not suggest that *maR-si- is specific to reciprocal constructions.

If we assume that no other derivational morphemes (except *-an, *-ay and reduplication) were involved in the semantic differences among the *maR-si- forms, then the question arises as to 1) the directionality of the semantic change of the proto-form *maR-si- and 2) which meaning is a retention or an innovation. To investigate the directionality, we must first eliminate comitative actor constructions from our semantic consideration because while actors in reciprocal, actor plural, and reflexive constructions are encoded in subject function in actor voice constructions, those in a comitative actor construction are indicated in two arguments — subject and comitative actor.

The comitative actor constructions are considered to be a result of a Post-Proto-Malayo-Polynesian development for the following reasons. As mentioned in Chapter 1, one of the functions of voice affixes in WISP languages is to mark an argument with a specific semantic role as a subject in a sentence (Himmelmann 2002; Ross 2002). For example, the actor voice marker *<um> indicates an actor argument as a subject: the patient voice marker *-en encodes a patient argument as a subject. In other words, it is only a single argument that is specified as a subject by a voice marker in WISP languages. Because the form *maR-si- can be traced back to the form *<um>paR-si-, it is plausible to assume that the prefix *si- in the form *maR-si- initially indicated something concerning the item encoded in a single argument, which is a subject in an actor voice construction. Comitative actor constructions encoding plural actors in two arguments including a non-subject argument, on the other hand, are assumed to be a result of the division of an argument that consists of a coordinate phrase indicating plural actors in actor plural and reciprocal constructions through the reanalysis of its syntactic structure, as in 'A with B do' as 'A does with B'.

The relationships among the remaining three constructions (i.e., actor plural, reciprocal, and reflexive constructions) can be explained through the two semantic features, namely, 1) plurality of actors and 2) actor-cum-undergoer role, as discussed in Chapter 2. While the feature of the plurality of actors is shared between actor plural and reciprocal situations, and that of the actor-cum-undergoer role is shared between reciprocal and reflexive situations, no common semantic feature is present between actor plural and reflexive situations and constructions. This observation suggests that, if the development of the three functions is unidirectional, the reciprocal function that has common features with both the reflexive and actor plural functions should be interposed between the reflexive and actor plural functions in the paths of the semantic change among the three functions. For this reason, two unidirectional patterns of the semantic change of the proto-form *maR-si- can be suggested. One is actor plural → reciprocal → reflexive, and the other is reflexive → reciprocal → actor plural.

Alternatively, one might suggest that the meaning of *maR-si- at the Proto-Malayo-Polynesian level was reciprocal; in other words, it originally had both semantic features. Considering this assumption, in contrast to the unidirectional patterns mentioned above, we must posit two patterns for explaining the presence of the actor plural and reciprocal functions. They are 1) reciprocal → actor plural and 2) reciprocal → reflexive.

Table 35: Functional patterns of reflexes of *maR-si-

Functional pattern in a language	Language and form	Subgroup
Actor plural only	Ilocano *ag-si-*	Philippine, Northern Luzon
	Pangasinan *mag-si-, mag-si- -an*	Philippine, Northern Luzon
	Tagalog *mag-si-([pag-/paN-])*	Philippine, GCP, Central Philippine
	Cebuano *(mag-)(i)sig-([pa-])*	Philippine, GCP, Central Philippine, Bisayan
	Aklanon *mag-si(g)-*	Philippine, GCP, Central Philippine, Bisayan
	Mongondow *mo-si-*	Philippine, GCP, Gorontalo-Mongondowic
	Sangir *ma-siq-([paN-])*	Philippine, Sangiric
	Muna *si- -ha* (actor dual), *si-CVCV- -ha* (actor plural)	Celebic, Muna-Buton
	Karo Batak *si-(RDP-) [(N(#)-/er-)] =na*	Sumatran, Batak
Reciprocal only	Malay *ber-si-, ber-se-*	Smith's Western Indonesian, Malayic
	Old Javanese *(m)a-si-*	Smith's Western Indonesian
	Lampung *s(a)- -an*	Smith's Western Indonesian
	Karo Batak *si-(RDP)- -en*	Sumatran, Batak
Actor plural and reciprocal	Cotabato Manobo *eg-se- -ay*	Philippine, GCP, Manoboic
	Da'a *mo-si-([pe-])*	Celebic, Kaili-Pamona
	Makassarese *(aC)si-*	South Sulawesi
Actor plural, comitative actor, and reciprocal	Pendau *mo-si-*	Celebic, Tomini-Tolitoli
Reciprocal and comitative actor	West Coast Bajau *si-*	Smith's Western Indonesian, Barito, Sama-Bajaw
Reciprocal and reflexive	Kiput *se-, se-p-, se-l-*	Smith's Western Indonesian, Berawan-Lower Baram
	Salako *(ba-)si(N#)-*	Smith's Western Indonesian, Malayic
	Sundanese *si-*	Smith's Western Indonesian
	Toba Batak *mar-si-*	Sumatran, Batak

3.3 Formations with *paR-si- with action-denoting roots in undergoer voice constructions

As discussed in Section 3.1, some languages have undergoer voice formations with reflexes of *si-. The following table summarizes *paR-si- forms and their functions in undergoer voice as discussed in the current study:

Table 36: *paR-si- constructions in undergoer with action-denoting roots

Language	Subgroup	Form	Function
Cotabato Manobo	Philippine, GCP, Manoboic	eg-se- -en	Undergoer (patient) plural in patient voice Comitative undergoer (patient) in patient voice
		eg-se- -an	Actor plural in locative voice
		eg-se-	Undergoer (theme) plural in thematic voice
Pendau	Celebic, Tomini-Tolitoli	ni-po-si-[poN-] -a’	Actor plural in undergoer voice

3.3.1 Common features in *paR-si- across languages

The current section introduces two features, which are: 1) reflexes of the Proto-Malayo-Polynesian phoneme *R and 2) vowel changes in reflexes of the initial segment *paR-si- in undergoer voice.

3.3.1.1 Reflexes of the Proto-Malayo-Polynesian phoneme *R in *paR-

It is observed that the Proto-Malayo-Polynesian phoneme *R is reflected as /g/ and /Ø/ in *paR-si- undergoer voice forms in our data. In contrast to *maR-si- forms, no irregularity concerning *R is attested. Compare the following examples with *paR-si- forms, reflexes of the prefix *paR- used in undergoer formations, and the preconsonantal *R in lexical words:

(195) *R > g/_C:

Language	*paR-si-	*paR-	preconsonantal *R in lexical words
Cotabato Manobo	eg-se- -en eg-se- -an eg-se-	eg-	bagkes ‘bundle of woods’[107] (< *beRkas ‘bundle’)

(196) *R > Ø/_C

Language	*paR-si-	*paR-	preconsonantal *R in lexical words
Pendau	ni-po-si- -a’	po-	m-bengi ‘night’ (< *beRngi ‘night’)

107Kerr (1988:117).

3.3.1.2 Vowel changes

Some vowel changes are also observed in the segments *paR- and *si- in the undergoer form *paR-si-. For example, both vowels in the segment *paR-si- appear to be mid-front unrounded vowels in Cotabato Manobo *eg-se-* /eg-se-/. Pendau reflects the vowel in the form *paR- as *po-* in *ni-po-si--a'*. In both languages, the reflexes of *paR- are identical with those of *paR- in *paR-si- in form. Compare the following examples with the reflexes of *maR-si-, the prefix *paR- in undergoer voice forms, and word-medial *a in lexical words:

(197)

Language	*paR-si-	*paR-	word-medial *a in lexical words
Cotabato Manobo	eg-se- /eg-se-/	eg- /eg-/	mata 'eye'[108] (< *mata 'eye')
Pendau	ni-po-si- -a'	po-	mata 'eye'[109] (< *mata 'eye')

3.3.2 Other forms and developments in *paR-si- formations

Some additional morphemes that are attested in the actor voice forms discussed earlier, such as action-denoting lexical class prefixes, can also be observed in *paR-si- forms in undergoer voice constructions. For example, the stem *ponuan* in Pendau undergoer voice actor plural form *ni-po-si-ponuan-a'* 'be poured (plural actors)' can be analyzed as comprising the root *tuang* 'pour' and the action-denoting lexical class prefix *poN-* (cf. Quick 2003).

In Cotabato Manobo, the initial consonant /p/ is not reflected in *paR-si- forms in undergoer voice constructions, as in the patient voice form *eg-se- -en* and the locative voice form *eg-se- -an*. In addition, this language also exhibits the absence of an overt reflex of the thematic voice marker *i- in the form *eg-se-*, which reflects *i-paR-si-.

3.3.3 Functional variations of *paR-si- formations with action-denoting roots

This section addresses functional variations denoted by forms with *paR-si- when the forms appear with action-denoting roots.

3.3.3.1 Undergoer plural constructions in undergoer voice

A *paR-si- form may indicate plural undergoers in subject function in an undergoer voice construction. Such a construction is, for example, formed with the Cotabato Manobo patient voice *si- form *eg-se- -en*. In the following examples, the item indicated by the patient subjects *sa elé* 'the *kunai* grass'*, sa libi* 'the *libi* palm', and *siya me-doo kalatas* 'a lot of paper' are plural, respectively:

Cotabato Manobo

(198) **Eg-se-**tepeng-**en** ku sa elé.
UV-SI-even.up-PV 1.SG.GEN NOM kunai.grass
'I will even up the *kunai (cogon)* grass
(the bundles of grass cut for roofing are jostled to bring the stem ends uniformly together)'
[The bundles of *kunai* grass will be evened up by me.]
(Kerr 1988:99, glossing modified)

108Quick (2003:135).
109Kerr (1988:109).

(199) **Eg-se**-lapin-**en** di sa libi.
UV-SI-layer-PV 3.SG.GEN NOM libi.palm.leaves
'He is interlaying the *libi* palm leaves.'
['The *libi* palm leaves are being interlaid by him.']

(Kerr 1988:99, glossing modified)

(200) Sa miyong, **eg-se**-kaen-**en** di siya me-doo kelatas.
DETERMINER cat UV-SI-eat-PV 3.SG.GEN that QD-many paper
'The cat is eating a lot of paper.'
['The cat, a lot of paper is eaten by it.']

(Kerr 1988:100, glossing modified)

In addition to the patient voice form with the prefix *se-*, the Cotabato Manobo thematic voice *si- form *eg-se-* also encodes undergoer plural constructions. In the following sentence, the phrase *sa me-doo lata* 'many tins (NOM)' is the subject and denotes plural items:

Cotabato Manobo
(201) **Eg-se**-tipon ku sa me-doo lata.
TV:UV-SI-gather 1.SG.GEN NOM QD-many tins
'I gather the tins together (in one spot)'
[The many tins are gathered by me.]

(Kerr 1988:101, glossing modified)

The formations with *si- indicating undergoer plural situations in undergoer voice constructions are summarized as follows:

Table 37: Undergoer plural constructions in undergoer voice with *paR-si-

Language	Subgroup	Form	Function
Cotabato Manobo	Philippine, GCP, Manoboic	eg-se- -en	Undergoer (patient) plural in patient voice
		eg-se-	Undergoer (theme) plural in thematic voice

3.3.3.2 Actor plural constructions in undergoer voice

Some *paR-si- forms in undergoer voice constructions are also employed to indicate the plurality of actors in a non-subject argument in undergoer voice constrictions. For example, the Cotabato Manobo form *eg-se- -an* encodes plural actors in a non-subject argument in locative voice constructions. In the following examples, the non-subjects *da* '3.PL.GEN' and *ta* '1.PI.GEN' encode plural actors:

Cotabato Manobo
(202) **Eg-se**-limud-**an** da aken.
UV-SI-crowd-LV 3.PL.GEN 1.SG.NOM
'They crowd in upon me.'

(Kerr 1988:101, glossing modified)

(203) **Eg-se**-tagped-**an** ta sa timun.
UV-SI-cut-LV 1.PI.GEN NOM cucumber
'We cut the cucumber and take half each.'

(Kerr 1988:101, glossing modified)

Pendau provides another instance where a non-subject argument encodes plural actors in an undergoer voice construction with a *paR-si- form. In the following example, the predicate *ni-po-si-poN-tuang-a'* 'is poured' appears with the actor non-subject *nijimo* '3.PL.GEN', and the plurality of the actors is marked with the prefix *si-*:

Pendau
(204) Ogo uo **niposi**ponuan**a'** nijimo api uo.
ogo 'uo ni-po-si-poN-tuang-a' nijimo api 'uo
water yonder UV.RLS-AD-SI-AD-pour-APPL 3.PL.GEN fire yonder
'Together they poured water on the fire. (It is implied that water is taken from one place or container)'

(Quick 2003:283, glossing modified)

The subject in this sentence is the theme *ogo 'uo* 'that water' because the basic word order in Pendau is SVO, and the theme argument is in the subject position in the undergoer voice construction with the applicative suffix *-a'*.

The actor plural constructions in undergoer voice with the prefix *si- are summarized in the following table:

Table 38: Actor plural constructions in undergoer voice with *paR-si-

Language	Subgroup	Form	Function
Cotabato Manobo	Philippine, GCP, Manoboic	eg-se- -an	Actor plural in locative voice
Pendau	Celebic, Tomini-Tolitoli	ni-po-si-([poN-]) -a'	Actor plural in undergoer voice

3.3.3.3 Comitative undergoer constructions in undergoer voice

Undergoer voice constructions with *paR-si- may take a singular undergoer subject with a comitative undergoer non-subject argument. For example, the Cotabato Manobo patient voice *paR-si- form *eg-se- -en* may take a singular actor subject to indicate that the patient has a comitative patient:

Cotabato Manobo
(205) **Eg-se**-tepeng-**en** ku sini unol.
UV-SI-even.up-PV 1.SG.GEN this unol.snake
'I will measure this *unol* snake.'
[This *unol* snake (with something else) will be evened up by me.]

(Kerr 1988:99, glossing modified)

The subject in this example is *sini unol* 'this *unol* snake' and is a singular undergoer. However, according to Kerr (1988:99), this sentence describes the situation in which the snake will be measured by placing a measuring rod alongside it. In other words, the snake and measuring rod will

be placed side by side in the situation. The measuring rod is considered to be the other patient of the action of *eg-se-tepeng-en* 'will be evened up together', and we can interpret that the presence of this concomitant patient is marked by the prefix *se-*. The meaning of 'to measure' appears to be due to this specific situation with the measuring rod, if we consider the fact that the same root is not employed to indicate the action of measuring in the sentence *eg-se-tepeng ku sa elé* 'I will even up the *kunai (cogon)* grass' (will.be.evened.up.together 1.SG.GEN NOM kunai.grass).

*3.3.3.4 Functional patterns of *paR-si- in undergoer voice*

Our WISP data demonstrated that *paR-si- forms in undergoer voice constructions may indicate 1) undergoer plural, 2) actor plural, and 3) comitative undergoer constructions.

Of these three functions, the undergoer and actor plural constructions are common in that plural participants with the same macro-role are encoded in one argument. On the other hand, undergoer plural and comitative undergoer constructions are common in the sense that both encode plural undergoers in undergoer constructions. These relationships are summarized in the following table:

Table 39: Relations among undergoer plural, actor plural, and comitative undergoer constructions

	Argument structure of the macro-role marked as plural	Participant plurality
Undergoer plural	One argument	Undergoer plurality
Actor plural	One argument	Actor plurality
Comitative undergoer	Two arguments	Undergoer plurality

In a similar manner to actor voice constructions with *maR-si-, the comitative undergoer function encoding plural undergoers in two arguments is assumed to have appeared based on the undergoer plural function that encodes plural undergoers in subject function.

The evaluation of the original function of *paR-si- in undergoer voice constructions depends on the original function of *maR-si- in actor voice constructions. If the actor plural function is the original in actor voice constructions with *maR-si-, the function of the prefix *si- is to encode plural participants in subject function. When the prefix *si- encodes plural participants in subject function in undergoer voice constructions, the macro-role of the participants is undergoer. Thus, the original function was undergoer plural.

On the other hand, if the reflexive or reciprocal function is original, the undergoer voice form *paR-si- might initially indicate the plurality of participants in subject function or that of actors in any argument because when the actor plural function of the form *maR-si- in actor voice is developed from another function, it is not certain whether the encoding of the actor plurality is performed in accordance with the syntactic rule that the prefix *si- encodes participant plurality in subject function. Therefore, when the prefix encodes plural participants in subject function, the form *paR-si- originally functioned as an undergoer plural marker. However, when it encodes plural actors in subject or non-subject function, the older function of *paR-si- is considered to have been actor plural.

In contrast to the actor voice *maR-si- forms, *paR-si- forms are rarely polysemous. The clear exception is the Cotabato Manobo *paR-si- patient voice form *eg-se- -en* that denotes 1) undergoer plural and 2) comitative undergoer constructions.

Formal types are classified according to the five voice-marked types: 1) *paR-si- -en with *-en

'patient voice', 2) *paR-si- -an with *-an 'locative voice', 3) *i-paR-si- with *i- 'thematic voice', and 4) *paR-si- without the Proto-Malayo-Polynesian undergoer voice affixes *-en, *-an, and *i-.

The following table illustrates the functional patterns of *paR-si- forms in question.

Table 40: Functional patterns of reflexes of *paR-si- with action-denoting roots

Functional pattern	Form	Subgroup
Undergoer plurals in undergoer voice constructions	***paR-si- -en**	
	Cotabato Manobo *eg-se- -en*	Philippine, GCP, Manoboic
	***i-paR-si-**	
	Cotabato Manobo *eg-se-*	Philippine, GCP, Manoboic
Actor plurals in undergoer voice constructions	***paR-si- -an**	
	Cotabato Manobo *eg-se- -an*	Philippine, GCP, Manoboic
	***paR-si-**	
	Pendau *ni-po-si-([poN-]) -a'*	Celebic, Tomini-Tolitoli
Comitative undergoers in undergoer voice constructions	***paR-si- -en**	
	Cotabato Manobo *eg-se- -en*	Philippine, GCP, Manoboic

3.4 Summary

This chapter analyzed *si- forms with action-denoting roots. It is observed that there are actor voice formations with reflexes of *maR-si- and undergoer voice formations with those of *paR-si-. Among these voice forms, the reflexes of the actor voice form *maR-si- are most widely observed in the WISP languages. Most reflexes exhibit various phonological and morphological changes, such as the loss of initial phonemes. In addition, some *maR-si- forms contain the action-denoting lexical class prefixes *paR- and *paN-, the event plural suffixes *-an and *-ay, and reduplication.

The construction types encoded by reflexes of *maR-si- are divided into four types of constructions: actor plural, reciprocal, comitative actor, and reflexive. Among these constructions, both the actor plural and reciprocal ones are observed in all the Western Indonesian, Sulawesi, and Philippine areas. However, the reflexive function is only attested in Western Indonesian languages. On the other hand, the comitative actor function is only found in one Sulawesi language (i.e., Pendau) and one Western Indonesian language (i.e., West Coast Bajau).

As discussed in Section 3.2.4.6, provided that the *maR-si- forms indicating the four functions are cognate without other morphemes that cause semantic and syntactic changes, the syntactic properties of actor plural, reciprocal, and comitative actor constructions suggest that the comitative constructions have been derived from the actor plural or reciprocal constructions. In addition, the semantic properties of the actor plural, reciprocal, and reflexive constructions marked by *maR-si- suggest two possible semantic change paths (i.e., (1) actor plural → reciprocal → reflexive and (2) reflexive → reciprocal → actor plural). It may also be hypothesized that the original function of *maR-si- was to encode reciprocal situations, and the reflexive and actor plural functions are developed directly from the reciprocal function. However, no further assumption as to the original meaning of the prefix *si- can be deduced only from the data available to us at this point.

In addition to the *maR-si- forms in actor voice constructions, we have also attested *paR-si- forms in undergoer voice constructions in two Philippine and Sulawesi languages — Cotabato

Manobo and Pendau. These forms are employed to indicate undergoer plural, actor plural, and comitative undergoer constructions in undergoer voice constructions.

4 Possible uses of *si- with other types of roots and words in WISP languages

This chapter discusses derivations, which include the form *si-, from other types of roots. The form in question possibly reflects the prefix *si- discussed in the previous chapter, which analyzed *si- formations on the basis of those forms with action-denoting roots.

However, it should be noted that the derivations discussed in the present chapter involve a number of unresolved issues in terms of their forms and functions. Hence, the present discussion only aims to suggest the possibility that some reflexes of *si- are also attested in other derivations.

4.1 With quality-denoting roots

Some *si- forms that co-occur with action-denoting roots also appear with quality-denoting roots in some languages. For example, the Malay form *ber-si-*, which is identical to the reciprocal and actor plural form *ber-si-*, may appear with some quality-denoting roots to derive the words that indicate competitive situations in which plural participants compete with respect to the quality denoted by the roots:

Competitive
Malay
(206) **ber-si-**cepat ‘compete with each other in speed’ cepat ‘to be fast’
ber-si-dahulu ‘to try to outstrip each other’ dahulu ‘before, earlier’
(Ogloblin and Nedjalkov 2007:1456)

In Karo Batak, the reciprocal form *si- -en* also attaches to some quality-denoting roots to form the words that denote a situation in which plural participants in subject function are not equal in terms of the quality indicated by the roots.

Different degrees of quality denoted by the root
Karo Batak

(207)	**Si**kitik**en**	kap	sepatuku	é!
	si-kitik-en	kap	sepatu=ku	é
	SI-small-EP	EMPH	shoes=my	that

‘These shoes aren’t the same size!’
(Woollams 1996:72, glossing modified)

The form *si- -en* may appear with full root reduplication. According to Woollams (1996:72), the two forms are in free variation:

Karo Batak

(208)	**Sigegeh-**gegeh**en**	kita	erdahin	pé.
	si-RDP-gegeh-en	kita	er-dahin	pé.
	SI-??-strong-EP	1.PI.NOM	AV-work	EMPH

‘Not everybody puts the same amount of effort into their work.’
[We are tough in varying degrees for working.]
(Woollams 1996:73, glossing modified)

(209) **si-(gedang-)**gedang-**en**	‘of different lengths’	gedang	‘long’
si-(berat-)berat-**en**	‘of different weight’	berat	‘heavy
si-(kitik-)kitik-**en**	‘one smaller than the other’	kitik	‘small’
si-(bentar-)bentar-**en**	‘not quite the same shades of white as each other’	bentar	‘white’

(Woollams 1996:72)[110]

The possible cognate form of the Karo Batak form *si- -en* is attested in Toba Batak, another Batak language. In this language, the form *mar-si- -i* denotes ‘be unequal in what is expressed by the adjective base [i.e., quality-denoting roots, YK]’ (Nababan 1981:100).

Toba Batak

(210) **mar-si-**tibbó-**i**	‘be of unequal height’	tibbó	‘high’
mar-si-gadjáng-**i**	‘be of unequal length’	gadjáng	‘long’
mar-si-pogós-**i**	‘be of unequal poverty’	pogós	‘poor’

(Nababan 1981:100)

This form shares the first segments *mar-si-* with the reciprocal and reflexive form *mar-si-* in this language. However, in contrast to the form *mar-si- -i*, the segment *-i* does not occur in the *si- formations with action-denoting roots. Nevertheless, the presence of the prefix *mar-* suggests that the voice of this form is actor voice, marked with the prefix *maR- (< *p<um>aR-). If the Toba Batak and Karo Batak forms share the same source, we can assume that the voice of the *si- formation of this function is actor voice.

In other languages, *si- derivations from quality-denoting roots differ clearly from those from actor-denoting roots. For instance, Tagalog has the form *mag(ka-)sing-* that appears with quality-denoting roots and a subject that denotes dual participants to designate the equality of the quality indicated by the root:

Tagalog

(211) **Mag(ka)sing**bago	ang	iyong	kotse	at	ang	akin.
mag-ka-sing-bago	ang	iyo-ng	kotse	at	ang	akin
AV-KA-SI-new	NOM	2.SG.OBL-LK	car	and	NOM	1.SG.OBL

‘Your car and mine are equally new.’

(Schachter and Otanes 1972:238, my glossing)

(212) **Mag(ka)sing**talino	sina	Juan	at	Pedro.
mag-ka-sing-talino	sina	Juan	at	Pedro
AV-KA-SI-intelligent	PN.NOM.PL	Juan	and	Pedro

‘Juan and Pedro are equal in intelligence.’

(Schachter and Otanes 1972:238, my glossing)

The origin of segment /ng/ in the segment *sing-* and the function of the prefix *ka-* in this form deserves further research. However, we consider it likely that the segment *si-* in this formation is a possible cognate of the prefix *si- because it appears with the prefix *mag-* (< *maR-) and the derived words take plural participants in subject function, in the same manner as *maR-si- actor

110These forms are written as *si-gedang(-gedang)-en* by Woollams (1996:72). However, they are rewritten here as *si-(gedang-)gedang-en*. etc., regarding the reduplicant as the preceding segment to the root because it can more easily be compared with other reduplicated forms where the reduplicant precedes the root, as in Tagalog *mag-(ka-)sing-CV-* and Muna *si-CVCV-*.

plural forms, such as Tagalog *mag-si-([pag-/paN-]).*

In addition, the Tagalog dual form *mag-(ka-)sing-* has as its plural form *mag-ka-ka-sing-*:

Dual participants sharing the same quality

(213) **Mag(ka)sing**talino sina Juan at Pedro.
mag-ka-sing-talino sina Juan at Pedro
AV-KA-SI-intelligent PN.NOM.PL Juan and Pedro
'Juan and Pedro are equal in intelligence.'

(Schachter and Otanes 1972:238, my glossing)

Plural participants sharing the same quality

(214) **Magkakasing**talino silang lahat.
mag-CV-ka-sing-talino sila-ng lahat
AV-PL-KA-SI-intelligent 3.PL.NOM-LK all
'They are all equal in intelligence.'

(Schachter and Otanes 1972:238, my glossing)

Tagalog also has the form *mag-ka-sing-CV-* that has the same function (Schachter and Otanes 1972:238)). The *CV-* in this form refers to the CV- reduplication of the root. Hence, the plural form *mag-ka-ka-sing-* can be analyzed as the form with *mag-(ka-)sing-* with the CV- reduplication of the base *ka-sing-ROOT.* In both forms *mag-ka-ka-sing-* and *mag-ka-sing-CV-*, the vowel of the reduplicants *ka-* and *CV-* is short (Schachter and Otanes ibid.). Thus, for the root *talino* 'clever', both the forms *mag-ka-ka-sing-talino* /magkakasingtaliˑnoh/ and *mag-ka-sing-ta-talino* /magkasingtataliˑnoh/ are possible (Schachter and Otanes ibid.).

In addition to these functions, possible reflexes of the form *si- appear with quality-denoting roots to indicate a *singular* entity that shares the same quality with someone else. For example, the Tagalog form *(ka-)sing-* derives a word indicating 'to have the same quality as someone else'.

In the following example, the phrase *si Maria* is a subject and the source of the comparison concerning the quality *ganda* 'beautiful' and the phrase *ni Elena* is the entity that is compared with the source, namely, Maria:

Singular associate *si-

Tagalog

(215) **(Ka)sing**ganda ni Elena si Maria.
ka-sing-ganda ni Elena si Maria
KA-SI-beautiful GEN Elena si Maria
'Maria is as beautiful as Elena.'

(Schachter and Otanes 1972:237, my glossing)

Other Tagalog examples with *(ka-)sing-* are as follows:

Tagalog

(216) **(Ka)sing**hirap ko siya.
ka-sing-hirap ko siya
KA-SI-poor 1.SG.GEN 3.SG.NOM
'He is as poor as I am.'

(Schachter and Otanes 1972:238, my glossing)

(217)	**(Ka)sing**luma	ng	bahay	natin	ang	kanila.[111]
	ka-sing-luma	nang	bahay	natin	ang	kanila
	KA-SI-old	GEN	house	1.PI.GEN	NOM	3.PL.OBL

'Their house is as old as ours.'

(Schachter and Otanes 1972:237, my glossing)

(218)	Hindi	**kasing**talino	ni	Mary	si	John.
	hindi	ka-sing-talino	ni	Mary	si	John
	NEG	KA-SI-intelligent	PN.GEN	Mary	PN.NOM	John

'John isn't as intelligent as Mary.'

(Schachter and Otanes 1972:238, my glossing)

The following formal aspects of this derivation require further research: They are 1) the prefix *ka-* and 2) the phoneme /ng/ after the prefix segment *si-. In addition, when compared to the Tagalog form *mag-(ka-)sing-*, we can observe that this form lacks an actor voice marker. This observation suggests that the two Tagalog forms may be related and only differ in their voice. The voice marking of the form *(ka-)sing-* is not apparent and requires further investigation.

The same function as the one indicated by the Tagalog *(ka-)sing-* is marked with the prefix *se-* in Malay:

Malay

(219)	Saya	**se**tinggi	Ali.
	saya	se-tinggi	Ali
	1.SG	SI-high	Ali

'I am as tall as Ali.'

(Sneddon 1996:182, my glossing)

(220)	Buku	ini	tidak	**se**mahal	yang	itu.
	buku	ini	tidak	se-mahal	yang	itu
	book	this	not	SI-expensive	REL	that

'This book is not as expensive as that one.'

(Sneddon 1996:181, my glossing)

It should be noted that this language has the *si- form *ber-si-* for action-denoting roots in actor voice constructions. The vowel change from *i to a schwa (represented as <e>) is not observed in the form *ber-si-*.

In the preceding discussion, the Malay prefix *se-* in *Saya setinggi Ali.* 'I am as tall as Ali' is considered a possible reflex of *si-. However, alternative analyses are possible. One might also think that *se-* in these constructions is Malay *se-* 'one' or *sama* 'same'. However, the Tagalog form *ka-sing-* of the same function, as in *ka-sing-ganda* 'is as beautiful as' (as in *Ka-sing-ganda ni Elena si Maria.* 'Maria is as beautiful as Elena'; *si* 'nominative', *ni* 'genitive') suggests that the Malay form also originates from a *si- form.

For explaining the semantic connection between *si- forms for action-denoting roots and those for quality-denoting roots, we first need to discuss the category labeled as 'equal quality between plural (or dual) participants'. The category in question can be seen as plural participants sharing the same quality, as in Tagalog *Mag-ka-sing-talino sina Juan at Pedro.* 'Juan and Pedro are equal in

111Originally written as '*kamila*', which is supposed to be a typographical error.

intelligence' (*sina* 'plural.nominative', *at* 'and') and *Mag-ka-kasing-talino sila-ng lahat*. 'They are all equal in intelligence' (*sila-ng lahat* 'they all (nominative)'). The plural participants in this situation are placed in subject position. Thus, this construction is identical to the actor plural construction in which plural participants are encoded in subject position. For this reason, we consider that the *si- or *sV- reflexes of this use for quality-denoting roots are identical to those for action-denoting roots.

The category 'associate sharing the same quality' refers to a singular participant who shares the same quality with someone else. The presence of this meaning for *si- reflexes can be explained by the fact that *si- reflexes can also encode comitative constructions with a singular participant in subject function with a comitative argument (e.g., Pendau *Odo no-si-balo sono ulasang* 'The monkey quarreled with the turtle.' *odo* 'monkey', *sono* 'with', *ulasang* 'turtle'). Therefore, this category can be related to the comitative actor constructions expressed by *si- and action-denoting roots.

To summarize, formations with possible reflexes of *si- with quality-denoting roots may have one of the following four meanings: 1) competitive, 2) different degrees of quality among plural participants, 3) equal quality between dual or plural participants, and 4) an associate who has the same quality with someone else (i.e., 'associative' function). Among the formal types that indicate these four functions, the former three types can be identified as having reflexes of the actor voice prefix *maR-, such as Malay *ber-* in *ber-si-*, Toba Batak *mar-* in *mar- -i*, and Tagalog *mag-* in *mag-(ka-)sing-*. The semantic relations among these three meanings, however, deserve further investigation.

4.2 With object-denoting roots

Possible cognates of *si- are also attested with some object-denoting roots. For example, in Muna, the prefix *si-* appears with an object-denoting root to indicate the situation in which the two participants in subject function share the object indicated by the root. This form is identical to the initial part of the Muna *maR-si- actor dual form *si- -ha*:

Muna
Dual participants sharing an object
(221) Ta-**si**-guru.
1.PE-SI-teacher
'We (dual, exclusive) have the same teacher.'
(van den Berg 2013 [1989]:319, my glossing)

When the participants in subject function are more than two, the form *si-CVCV-* is employed to indicate the situation of sharing an object. This form is also identical to the initial part of the Muna actor plural *maR-si- form *si-CVCV- -ha*:

Muna
Plural participants sharing an object
(222) Ta-**si-guru**-guru.
ta-si-CVCV-guru
1.PI.NOM-SI-PL-teacher
'We (exclusive) have the same teacher.'
(van den Berg 2013 [1989]:319, my glossing)

In Malay, the sharing of an object by plural participants in subject function is marked by *se-*, as

in the following example:

Malay
Plural participants sharing an object
(223) Amir dan Hasan **se-**kelas.
Amir and Hasan SI-class
'Amir and Hasan are in the same class at school.'
(Ogloblin and Nedjalkov 2007:1474, my glossing)

The form *se-* is formally identical to the one that has the associative function, mentioned in the previous section. In the same manner as the one for quality-denoting roots, it illustrates the vowel change from *i to a schwa. It should be noted that this form is also employed to indicate the associative function for object-denoting roots, as will be mentioned below.

In another derivation, *si- appears to denote an associate who either 1) shares an object with someone else, or 2) is of the same quality (e.g., occupation and relationship) as someone else is. These two meanings can be distinguished as follows. If the root and derived word refer to different entities, as do the Aklanon root *klasi* 'class' and its derived word *i-sig-ka-klasi* 'a classmate', it is the object denoted by the root that is shared. Roots for this usage typically signify inanimate items, such as 'class', 'house', and 'name'. By contrast, if they refer to the same entity, it is the quality of the object indicated by the root that is shared. For example, the Aklanon root *maestra* 'teacher' and its derived word *i-sig-ka-maestra* 'a co-teacher; a fellow-teacher' can be considered to refer to the same entity because the quality or occupation of a co-teacher is also a teacher as that of someone else is. Roots for this usage typically denote animate entities, such as 'teacher', 'man', and 'child'. Compare:

Associate with a shared object
Aklanon
(224) **i-sig-ka-**sangáy 'a person with the same name as another' [sangáy 'name']
i-sig-ka-klasi 'a classmate' [klasi 'class']
(Zorc 1968:137)

(225) **ka-sim**ałáy 'someone coming from the same house' [bałáy 'house']
ka-simaryo 'someone coming from the barrio; a barriomate' [baryo 'barrio
[someone coming from the neighborhood; neighbor'] [i.e., neighborhood]']
ka-simanwa 'a town mate' [banwa 'town']
(Zorc 1968:138)

Malay
(226) Amir **se-**kelas dengan Hasan.
Amir SI-class with Hasan
'Amir is in the same class with Hasan.'
(Ogloblin and Nedjalkov 2007:1474, my glossing)

(227) rakan **se**kamarnya
rakan se-kamar=nya
friend SI-room=3.GEN
'his room mate (sic)'
(Sneddon 1996:51, my glossing)

(228) **se**-kantor (dengan) ‘sharing an office with (with)’ [kantor ‘office’ dengan ‘with’]
se-pendapat (dengan) ‘of the same opinion (as)’ [pendapat ‘opinion’]
se-umur (dengan) ‘being the same age (as)’ [umur ‘age’]
(Sneddon 1996:52)

Associate with a shared quality of an object

Cebuano

(229) **i-sig-ka**-tawo ‘fellow being’ tawo ‘man, human being’
i-sig-ka-sakup ‘fellow member’ sakup ‘member’
(Lopez 1977b [1949]:129)

Aklanon

(230) **i-sig-ka**-ungá’ ‘a fellow child; a sibling’ **[**ungá ‘child’]
i-sig-ka-maestra ‘a co-teacher; a fellow teacher’ [maestra ‘teacher’]
(Zorc 1968:137)

Pamona

(231) **ka-si**-tau ˈfellow person’ [tau ˈperson’]
ka-si-to’isilamu ˈfellow Muslim’ [to’isilamu ˈMuslim’]
ka-si-taumaju’a=nya ˈhis fellow patients’ [tau maju’a ˈpatient’, =nya ‘3.SG.GEN’]
(Adriani 1931:267–8)

In these formations, we can observe the presence of 1) the prefix *ka-* in Aklanon and Cebuano *i-sig-ka-*, Aklanon *ka-siN-*, and Pamona *ka-si-*, 2) additional consonants after *si-, such as /g/ in Cebuano *i-sig-ka-* and /N/ in Aklanon *ka-siN-*, and 3) the form *i-* in Aklanon and Cebuano *i-sig-ka-*. Among these forms, the same additional segments are also observed in the *si- forms with action-denoting roots in actor voice constructions in some languages, as in Cebuano *(mag-)(i-)sig-* and Aklanon *mag-si(g)-*. The source of the form *i-* requires further research; however, it formally resembles the Cebuano and Aklanon thematic voice form *i-*.

Finally, it is worth mentioning the presence of the Cebuano form *ma-sig-ka-*, which appears with object-denoting roots. It derives words that denote the plurality or duality of the objects indicated by the root. The derived words with *ma-sig-ka-* are always preceded by a phrase marker and function as object-denoting words (i.e., nouns). In the following sentences, the nominative marker *ang* is placed before the *ma-sig-ka-* words *ma-sig-ka-ligid* ‘(two or more) wheels’ and *mag-si-ka-tumuy* ‘(two) tips’, respectively:

Cebuano

Plural or Dual

(232) Nahiyusan ang **masig ka** ligid.
na-hiyus-an ang masigka-ligid
UV.QD.RLS-flat-LV NOM SI-wheel
‘All of the wheels had a flat tire.’
(Wolff 2012 [1972]:n.p., my glossing)

Dual

Cebuano

(233) Tabúa ang **masig ka** tumuy arun glúhan.
tabu-a ang masigka-tumuy arun gluh-an.
make-PV.SUBJUNCTIVE NOM SI-tip in.order.to glue-LV
'Make both ends meet so they can be glued.'

(Wolff 2012 [1972]:n.p., my glossing)

The derived words with *ma-sig-ka-* also appear with other phrase markers. For example, in the following sentence, the word *ma-sig-ka-kilid* 'sides' appears with the locative phrase marker *sa*. The phrase *sa ma-sig-ka-kilid* signifies the location 'on either side':

Cebuano

(234) Ang irù dúnay itum nga bátuk sa **masig ka** kílid.
ang iru duna=y itum nga batuk sa masigka-kilid
NOM dog have=INV black LK spot LOC SI-side
'The dog has a black spot on either side.'

(Wolff 2012 [1972]:n.p., my glossing)

In this formation, we can observe the presence of 1) the prefix *ka-*, 2) the prefix *ma-*, and 3) the segment /g/ after *si-*, all of which require further investigation.

To summarize, when they attach to object-denoting roots, the formations with the possible reflexes of *si- may denote one of the following four meanings: 1) sharing of an object by dual or plural participants in subject function, 2) a singular associate entity who has the same object that someone else does, 3) a singular associate entity who has the same quality of an object as someone else does, and 4) the plurality or duality of the object denoted by the root.

4.3 With kinship terms

Some kinship terms appear in formations that contain the segment *si- to indicate a pair that consists of what the kinship term denotes and its 'counterpart'. For example, the Cebuano form *ma-sig-ka-* derives the word *ma-sig-ka-asawa* 'husband and wife [< 'wife and her counterpart']' from the word *asawa* 'wife':

Cebuano

(235) **ma-sig-ka-**asawa 'husband and wife' asawa 'wife'

(Lopez 1977b [1949]:132)

A form that has this function may appear with other kinship terms. Consider the following examples in Bontok:

Bontok

(236) **sin**-ama 'father and child' ama 'father'
sin-agi 'brothers' [agi 'relative at the same generation'
< *huaji 'younger sibling', YK]

(Reid 1976:xv)

When the Bontok form *sin-* appears with a root that indicates 'father', the derived word denotes the pair 'father and child', that is, 'father and his counterpart (= child)'. When it appears with a root that

indicates a ‘younger sibling’, the derived word denotes the pair ‘younger and older siblings’, that is, ‘younger sibling and its counterpart (= older sibling)’. As mentioned in Chapter 2, such pairs are termed ‘converse kinship pairs’ in this study. When the derived word signifies a group comprising more than two members, such as ‘father and (his) children’ with the root ‘father’, the group is termed a ‘converse kinship group’.

The converse kinship pair constructions are apparent when the derivational forms take the kinship term with a specification with regard to its generation or age relative to the ego or another deictic center. For example, the lexical item ‘father’ has the specification concerning his generation relative to the ego’s (i.e., one generation higher than the ego): the lexical item ‘younger brother’ has the specification regarding his age relative to the ego’s (i.e., younger than the ego).

Other converse kinship pair formations with an *s- initial segment are as follows:

Sangir

(237) i kami **sengka**tuhang
i kami seng-ka-tuhang
1.PE.NOM SI.KIN-KA-older.sibling
‘we (younger siblings) with our older siblings’
(Adriani 1893:82, my glossing)

(238) i kami **sengka**pulung.
i kami seng-ka-pulung
1.PE.NOM SI.KIN-KA-grandchild
‘I (grandfather) with my grandchildren. [we, who are a grandfather and grandchildren]’
(Adriani 1893:82–3, my glossing)

Pamona

(239) kami **san**tua’i.
kami sa(N#)-tua’i
1.PE SI.KIN-sibling
‘we (older siblings) with the younger brothers and sisters’
(Adriani 1931:262, my glossing)

(240) tau **sam**akumpu
tau sa(N#)-makumpu
person SI.KIN-grandchild[112]
‘a grandfather (grandmother) with grandchild(ren)’
(Adriani 1931:262, my glossing)

(241) tau setu **sa**ana
tau setu sa(N#)-ana
person DIST SI.KIN-child
‘they with their child(ren); He with his child(ren), she with her child(ren)’
[lit. those people are child(ren) and parent(s)]
(Adriani 1931:262, my glossing)

The *si- converse kinship formations also have some formal issues: 1) the prefix *ka-* in Cebuano *ma-sig-ka-* and Sangir *seng-ka-*, 2) unexpected phonemes after *si-* (i.e., /g/ in Cebuano *ma-sig-*

112Stokhof ed. (1984:90).

ka-, /ng/ in Sangir *seng-ka-*, /n/ in Bontok *sin-* and Pangasinan *san-*, and /N#/ in Pamona *sa(N#)-*), and 3) vowel changes (i.e., *i to /a/ in Pangasinan *san-* and Pamona *sa(N#)-* and *i to *e* /ə/ in Sangir *seng-ka-*). Some of these issues are also observed in *maR-si- formations with action-denoting roots in the same languages. For example, the unexpected phoneme /g/ after *si- is also attested in Cebuan*o (mag-)(i-)sig-*. However, as observed in the Pangasinan pair *san-* and *mag-si-* and the Sangir pair *seng-ka-* and *ma-siq-([paN-])*, the formal issues that are attested in a converse kinship pair formation are not observed in a form for action-denoting roots within a language.

In addition, it is noteworthy that the converse kinship pairs are also indicated with other formations, in contrast to the derivations from quality- and object-denoting roots discussed in the two preceding sections. One of these formations is the form *(mV(R/n)-)tVl(V)- with a disyllabic or monosyllabic segment *tVl(V)-, the first consonant of which is the voiceless alveolar stop /t/ and the second consonant of which is the voiced alveolar lateral /l/. Reflexes of the segment *tVl(V)- may or may not follow the reflexes of *maR- or *maN-, as in Hanunoo *mag-tal-* and Matigsalug Manobo *tala-*, respectively. Our available data suggest that the second vowel in *tVl(V)- is *i and the first vowel is either *a, *e (/ə/), or *o. For simplicity, we refer to the *(mV(R/n)-)tVl(V)- form by the form *maR-tali-, based on the Hanunoo form *mag-tal-* (< *maR-tal-) and the occurrence of the final *i in the forms in other languages, such as Uma *n-tali-* and Pendau *n-toli-*. Some *maR-tali- forms are as follows:

Hanunoo (Philippine, GCP, South Mindoro)

(242)	**mag-tal-**áma?	'father and child (son or daughter)'	áma? 'father'
	mag-tal-ína?	'mother and child (son or daughter)'	ína? 'mother'

(Conklin 1953)

Matigsalug Manobo (Philippine, GCP, Manoboic)

(243)	**tala-**ari 'siblings'	hari 'younger sibling'
	tala-anak 'parent and child	anak 'child'

(Wang et al. 2006:40)

Uma (Celebic, Kaili-Pamona)

(244)	**n-tali-**tina 'mother and child(ren)'	[tina 'mother']

(Esser 1964:22)

Pendau (Tomini-Tolitoli)

(245)	**n-toli** siama	'father and child' (lit. 'father counterpart')	[siama 'father']
	n-toli siina	'mother and child' (lit. 'mother counterpart')	[siina 'mother']
	n-toli mo'upu	'grandparent and grandchild'	[mo'upu 'grandchild']
	n-toli unga	'father and child, mother and child (lit. 'child counterpart')	[unga 'child']

(Quick 2003:186)

Another formation indicating the converse kinship pairs and groups is the prefix *maR-. This formation is identical to the actor voice *paR- class prefix *maR-. Consider the following examples in Tagalog and Ilocano:

Tagalog

(246) Ang **mag**amá ng Maryà at si Pédro
ang mag-amá -ng Maryà at si Pedro
NOM KIN-father -LK Maria and PN.NOM Pedro

ay naparoòn sa búkid.
ay na-paroon sa búkid.
INV AV.RLS-go.there LOC country
'Pedro and his daughter Maria have gone to the country.'
[The father and his child, who are Pedro and Maria, have gone to the country.']
(Bloomfield 1917:242, my glossing)

(247) Ang **mag**amá ni Ricardo'y mag(ka)síngdunong.
ang mag-ama ni Ricardo ay mag-ka-sing-dunong
NOM KIN-father PN.GEN Ricardo INV AV-KA-SI-wise
'Ricardo and his father are equally (both) wise.'
(Lopez 1977a [1937]:43, my glossing)

(248) **mag**-amà 'father and child' [amà 'father']
mag-inà 'mother and child' [inà 'mother']
mag-impò 'grandmother and grandchild' [impò 'grandmother']
mag-áli 'aunt and niece or nephew' [áli 'aunt']
(Bloomfield 1917:242)

Ilocano (cf. *maR- > *ag*-; *mag-si- > *ag-si*- 'actor plural')

(249) **ag**-ama 'father and son' ama 'father'
ag-ina 'mother and child' ina 'mother'
ag-uliteg 'uncle and nephew' uliteg 'uncle'
(Rubino 1997:87)

Some languages make a formal distinction between the converse kinship pair and group forms. For example, in Pangasinan, a converse kinship group word consists of the converse kinship pair marker *san*- and CV- reduplication:

Pangasinan

<u>Converse kinship pair (dual)</u>

(250) **san**-amá 'father and child' amá 'father'
san-iná 'mother and child' iná 'mother'
san-bái 'grandmother and grandchild' bái 'grandmother'
(Benton 1971:106)

<u>Converse kinship group (plural)</u>

(251) **san-a**-amá 'father and children' amá 'father'
san-i-iná 'mother and children' iná 'mother'
san-ba-bái 'grandmother and grandchild' bái 'grandmother'
(Benton 1971:106)

The distinction between converse kinship pair and groups are attested in languages that employ other formations without the use of *si- for indicating them. In Sarangani Manobo, converse kinship groups are expressed by the form *tel-tele*- (/təl-tələ-/), which comprises leftward CVC-reduplication of the converse kinship pair form *tele*- (/tələ-/):

Sarangani Manobo
Converse kinship pair (dual)
(252) **tele**-hazi ‘two siblings’ hazi ‘younger brother, sister or cousin’
tele-suled ‘two siblings or cousins’ suled [‘same-sex sibling or cousin’[113]]
(Elkins and Hendrickson 1985:170–1)

Converse kinship group (plural)
(253) **tel-tele**-hazi ‘three or more siblings’ hazi ‘younger brother, sister or cousin’
tel-tele-suled ‘three or more siblings or cousins’ suled [‘same-sex sibling or cousin’]
(Elkins and Hendrickson 1985:170–1)

In Tagalog, words indicating converse kinship groups are formed with the converse kinship pair marker *mag*- and CV- reduplication:

Tagalog
Converse kinship pair (dual)
(254) **mag**-amà ‘father and child’ amà ‘father’
(Bloomfield 1917:242)

Converse kinship group (plural)
(255) **mag-a**-amà ‘the group of a father with two or more of his children.’ amà ‘father’
(Bloomfield 1917:242)

As observed in the data, these Sarangani Manobo and Tagalog forms lack the segment *si-; however, reduplication is employed as a plural marker in the same manner as the Pangasinan pair, as well as Muna action-denoting *maR-si- actor dual form *si- -ha* and its non-dual plural *si-CVCV- -ha*.

Table 41: Formal distinctions between converse kinship pair and group forms in some Philippine languages

Language	Converse kinship pair (dual form)	Converse kinship group (plural form)
Pangasinan	san-	san-CV-
Sarangani Manobo	tele-	tel-tele-
Tagalog	mag-	mag-CV-

The *si- converse kinship pair and group forms resemble the *si- forms with object-denoting roots because both forms encode non-singular objects. However, the converse kinship pair meaning does not merely indicate the duality of the entity the root indicates; it also denotes a specific relationship between the two entities. In addition, the converse kinship pairs and groups are also indicated by *maR-tali- and other forms. The history of the converse kinship pair and group formations requires further study.

113Originally ‘male’s brother or male cousin, or female’s sister or female cousin’ (Elkins and Hendrickson 1985:171).

4.4 With location words

Some words denoting locations take the form *si- to derive words that indicate that two items are in the same locative relation denoted by the root toward each other, as in the following West Coast Bajau:

West Coast Bajau

(256)	**si**-bunda’	‘to face each other’	bunda’	‘in front of’
	si-dembila’	‘across from each other’	dembila’	‘across’
	si-lekat	‘to separate from each other; to divorce’	lekat	‘from’

(Miller 2007:251)

Miller does not provide any examples in sentence context. However, the meanings of the derived words suggest that they take a subject that encodes 1) dual participants or 2) a singular participant with a comitative argument. Despite this syntactic issue, however, we identify this *si-* as a possible reflex of the prefix *si- because the meanings of the derived words presuppose the presence of plural (or dual) participants, in the same way as the prefix *si- as an actor plural and reciprocal marker for action-denoting roots.

4.5 Discussion

Our data demonstrated the presence of the forms with some possible cognates of *si- with non-action-denoting roots and words in WISP languages. These forms are summarized in the following table:

Table 42: Forms with possible reflexes of *si- with non-action-denoting roots (i.e., with object- and quality-denoting roots, kinship terms, and location words)

Language	Functions	Forms	*maR-si- with action-denoting roots in the language
with quality-denoting roots			
Tagalog	Equal quality between dual participants	mag-(ka-)sing-	mag-si-([pag-/paN-]) ‘actor plural’
	Equal quality between plural participants	mag-ka-ka-sing-, mag-ka-sing-CV-	
	Associate sharing the same quality	(ka-)sing-	
Malay	Competitive	ber-si-	ber-si-, ber-se- ‘reciprocal’
	Associate sharing the same quality	se-	
Karo Batak	Different degrees of quality among plural participants	si-(RDP-) -en	si-(RDP) -en ‘reciprocal’ si-(RDP-)([N(#)-/er-]) =na ‘distributive actor plural’
with object-denoting roots			
Cebuano	Associate sharing the same quality of an object	i-sig-ka-	(mag-)(i)sig- ‘distributive actor plural’
	Plural or dual objects	ma-sig-ka-	
Aklanon	Associate sharing the same quality of an object	i-sig-ka- ka-siN-	mag-si(g)- ‘distributive actor plural’
	Associate sharing an object	i-sig-ka-	
Pamona	Associate sharing the quality of an object	ka-si-	N/A
	Associate sharing an object	ka-si-	
Malay	Plural participants sharing an object	se-	ber-si-, ber-se- ‘reciprocal’
	Associate sharing an object	se-	
Muna	Dual participants sharing an object	si-	si- -ha ‘actor dual’ si-CVCV- -ha ‘actor plural’
	Plural participants sharing an object	si-CVCV-	
with kinship terms			
Bontok	Converse kinship pair	sin-	N/A
Pangasinan	Converse kinship pair	san-	mag-si- ‘distributive actor

	Converse kinship group	san-CV-	plural' mag-si- -an 'successive distributive actor plural'
Sangir	Converse kinship pair	seng-ka-	ma-siq-([paN-]) 'distributive actor plural'
Pamona	Converse kinship pair Converse kinship group	sa(N#)-	N/A
with location words			
West Coast Bajau	Symmetrical relationship between dual/plural participants (in subject function or in subject and comitative arguments)	si-	si- 'reciprocal, comitative actor'

In contrast to the forms for action-denoting roots, we can observe that some forms for quality- and object-denoting roots and kinship terms contain the segment /ka/, as in Tagalog *mag-(ka-)sing-* and *(ka-)sing-*, Cebuano *ma-sig-ka-* and *i-sig-ka-*, Aklanon *i-sig-ka-* and *ka-siN-*, Pamona *ka-si-*, and Sangir *seng-ka-*. The form of the prefix *ka- and the meaning of the roots that appear with the prefix suggest that the source of the prefix is assumed to be the Proto-Austronesian prefix *ka- that is observed in various forms with quality-denoting roots. One such form is the causative form *pa-ka- with the causative prefix *pa-, as observed in Mayrinax Atayal *pa-ka-hauq* 'make [something] soft' with the root *hauq* 'soft' (Zeitoun and Huang 2000:404). The root appears in the underived form *ma-hauq* 'soft' with the 'stative' prefix *ma-*. This prefix originates from *k<um>a-, consisting of the actor voice infix *<um> and the quality-denoting class prefix *ka- (Blust 2003:404).

Sulawesi languages demonstrate that action-denoting lexical class prefixes may appear in the position of *ka- in causative forms with *pa-. For example, in Pamona, the causative form with the causative prefix *pa- for quality-denoting roots is *mam-pa-ka-*, as in *mam-pa-ka-rate* 'lengthen'. The root in this causative form is *rate* 'long', and its underived form is *ma-rate* 'long' with the prefix *ma-. On the other hand, those for *pang-* and *po-* class roots denoting actions are *mam-pa-pang-* and *mam-pa-po-*, as in *mam-pa-pang-inu* 'let drink' (root *inu* 'drink'; underived form in actor voice *mang-inu* 'drink' (< *p<um>ang-inu)) and *mam-pa-po-lonco* 'make run' (root *lonco* 'run'; underived form in actor voice *mo-lonco* 'run') (data from Adriani 1931). Similarly, Sangir also displays instances concerning causative forms for action-denoting roots. For example, the causative form of the word *mang-alaq* 'will take', historically consisting of the Proto-Malayo-Polynesian actor voice infix *<um>, lexical class prefix *pang-*, and the action-denoting root *alaq* 'take' (< Pre-Sangir *p<um>ang-alaq) is *ma-pang-alaq* 'will cause to take' (< *p<um>a-pang-alaq), in which the lexical class prefix *pang-* is placed before the action-denoting root. On the other hand, the causative form of the *peq-/pen-* class root *men-diko* 'will hear' is *ma-pen-diko* 'will cause to hear' (< *p<um>a-pen-diko) (data from Adriani 1893). If we assume that Pamona *po-* and Sangir *peq-/pen-* originate from the prefix *paR-, we can reconstruct the Proto-Malayo-Polynesian causative forms *(maN#-)pa-paR- and *(maN#-)pa-paN- for action-denoting roots and *(maN#-)pa-ka- for quality-denoting roots, respectively.

As discussed in Section 3.2.3.1, it is observed that some *si- forms for action-denoting roots contain lexical class prefixes, such as Tagalog *paN-* in the form *mag-si-paN-* for *paN-* class roots. Given that a lexical class prefix is placed after a derivational affix in *maR-si- forms in a similar manner to *pa- causative forms, we can assume that the *si- form for quality-denoting roots is

*maR-si-ka- with the quality-denoting lexical class prefix *ka-. In other words, as the causative forms *pa-paN- and *pa-ka- are used for action- and quality-denoting roots, respectively, it is assumed that the presence of the form *maR-si-paN- for action-denoting roots suggests that there is *maR-si-ka- for quality-denoting roots. However, it should be noted that no language that reflects the hypothesized form without any formal changes has been attested. Among the *si- forms for quality-denoting roots, the closest one to *maR-si-ka- is Tagalog *mag-ka-sing-* with the re-ordering of reflexes of the affixes *ka- and *si- and an additional consonant after *si-.

It should also be pointed out that some forms with the possible reflexes of *si- also have the formal distinction to indicate the semantic distinction between the dual and plural participants in subject function. It is remarkable to note that, in all instances, the non-dual plural form is marked with a reduplicative pattern, such as CV- and CVCV-. Such dual and plural formal distinctions are also attested in converse kinship pair and group forms without the form /si-/, as in the Tagalog converse kinship pair form *mag-* and the converse kinship group form *mag-CV-*, as well as the Sarangani Manobo converse kinship pair form *tele-* and converse kinship group form *tel-tele-*.

Table 43: Dual and plural distinction of *maR-si- forms with action-denoting roots and that of forms with possible reflexes of *si- with quality- and object-denoting roots

	Dual form	Plural form	Meaning
with action-denoting roots			
Muna	**si-** -ha	**si-**CVCV- -ha	Actor dual and plural (dual or plural participants in subject function perform an action/actions not to each other)
with quality-denoting roots			
Tagalog	mag-(ka-)**sing-**	mag-ka-ka-**sing-**, mag-ka-**sing**-CV-	Equal quality between dual and plural participants
with object-denoting roots			
Muna	**si-**	**si-**CVCV-	Dual and plural participants sharing an object
with kingship terms			
Pangasinan	**san-**	**san**-CV-	Converse kinship pair (dual) and group (plural)

5 Possible attestations of *si- formations in Formosan and Oceanic languages

This chapter introduces possible Formosan and Oceanic cognates of the *si- formations attested in the WISP languages mentioned above. Formosan languages refer to the Austronesian languages in Taiwan (except for Yami, which belongs to the Batanic subgroup of the Malayo-Polynesian branch). They are claimed to be the primary branches of the Austronesian family, in the same manner as the Malayo-Polynesian branch (Blust 2013:30). On the other hand, Oceanic languages belong to the Eastern Malayo-Polynesian branch of the Central-Eastern Malayo-Polynesian branch of the Malayo-Polynesian branch (Blust 2013:33).

5.1 Formosan

The following subsections will identify possible cognates of the PMP form *maR-si- and its inflected forms in Formosan languages and discuss what they would signify for the reconstruction of their Proto-Austronesian forms.

5.1.1 Kavalan *sim-*

We can expect that the Formosan languages would reflect the actor voice *si- form *maR-si- as *maR-si- because, according to Blust (2013:600–601), Proto-Malayo-Polynesian reflects Proto-Austronesian *m, *a, *R, *s, and *i, as *m, *a, *R, *s, and *i, respectively. However, we cannot detect a reflex of the form *maR-si-, or a form comprising reflexes of the two components, namely, *maR- and *si-.

If we minimize our formal expectation and search for a form that 1) contains the segment /si/ or /sV/, and that 2) derives a reciprocal or actor plural word with an action-denoting root, we can only detect the Kavalan form *sim-* in our data. Compare the following examples with the root *pukun*. They demonstrate the actor plural/distributive actor plural, reciprocal uses of the word *sim-pukun*:

Kavalan
Underived
(257) p<um>ukun (ya) ci Buya ci Abas-an.
AV-hit NOM PN Buya PN Abas-LOC
'Buya hit Abas.'
(Sung and Shen 2006:243, glossing modified and personal names are capitalized)

Actor plural / Distributive actor plural
(258) **Sim**-pukun ci Abas atu ci Buya ci Utay-an.
REC-hit PN Abas and PN Buya PN Utay-LOC
'Abas and Buya (together) hit Utay.' [Actor plural reading]
'Abas hit Utay and Buya also hit Utay.' [Distributive actor plural reading]
(Sung and Shen 2006:250, glossing modified and personal names are capitalized)

Reciprocal

(259) **Sim**-pukun (ya) ci Buya atu ci Utay.
SIM-hit NOM PN Buya and PN Utay
'Buya and Utay hit each other.'
(Sung and Shen 2006:243, glossing modified and personal names are capitalized)

It should be noted that in the data provided in Sung and Shen (2006:242), the final nasal in the form *sim-* is not assimilated to the following roots or prefixes, as in *sim-kawit* ['hold hands each other'] (< *kawit* 'hold hands') and *sim-sanu* ['tell each other'] (< *sanu* 'tell'). However, as they also point out, the data employed by Li (1996:70, quoted in Sung and Shen 2006:242 fn.) show the assimilation, as in *sin-tayta* 'look at each other'.

Furthermore, the prefix *sim-* also appears with undergoer voice forms that are marked with *-an.* For example, when the root *bula* 'give' takes the prefix *sim-* and the undergoer voice suffix *-an*, the derived word denotes the situation in which plural actors give their own items to each other. Compare the underived and *sim-* forms with the root *bula* 'give':

Underived undergoer voice form

(260) Bula-an ni Buya ci Utay tu usiq peRasku Raq.
give-UV GEN Buya PN Utay OBL one bottle liquor
'Buya gave Utay a bottle of liquor.'
(Sung and Shen 2006:258, glossing modified and personal names are capitalized)

Sim- form

(261) **Sim**-bula-an-na ni Utay atu ci Buya ya Raq.
SIM-give-UV-3.SG.GEN GEN Utay and PN Buya NOM liquor
'Utay and Buya gave each other liquor.'
(Sung and Shen 2006:258, glossing modified and personal names are capitalized)

Another point that should be addressed is the form *sim-* in the *sim-* undergoer voice form *sim--an.* The presence of the undergoer voice forms with *si- is not problematic because the undergoer voice forms with reflexes of *si- are also attested in some WISP languages. However, the phoneme /m/ in the undergoer form *sim-* deserves further discussion. If it only occurs in actor voice constructions, it can be analyzed as a reflex of the Proto-Austronesian actor voice infix *<um>, although it is expected as **s<m>i- because, according to Li and Tsuchida (2006:15), the Kavalan actor voice infix <*um*> is realized as <*m*> when the stem-initial consonant is an obstruent, as in *s<m>um* 'urinate'. However, the fact that it also occurs in the undergoer voice form poses the question of whether this analysis is valid.

One possibility that we can suggest for the presence of the phoneme *m* is that the form *sim-* originates from one of the putative 'lexical prefixes', which are considered to be characteristic of some Formosan languages (cf. Adelaar 2004). However, the data of Kavalan prefixes provided by Li and Tsuchida (2006:14–23) cannot identify the 'lexical prefix' that served as the source of the phoneme /m/.

5.1.2 Other possible cognates

Formosan languages also have forms that derive converse kinship pair words from kinship terms. The forms vary from language to language in the same manner as the converse kinship pair forms in Philippine and Sulawesi languages. Of such forms, we can observe that Seediq forms resemble the WISP converse kinship pair constructions the most because 1) the initial consonant is /m/, and 2) the second (or non-initial) consonant is /s/ or /t/, in a similar manner to the WISP forms *maR-si- and *maR-tali- that have the initial /m/ and the non-initial consonants /s/ and /t/, respectively:

Seediq
Converse kinship pair construction with **ms-**

(262) **ms**-qadil ‘wife and husband’ qadil ‘wife’
ms-bubu ‘mother and daughter’ bubu ‘mother’
(Sung and Shen 2006:260, fn. 22)

Converse kinship pair construction with **mt-**

(263) **mt**-suwayi ‘brothers/sisters’ suwayi ‘younger brother/sister’
[cf. PAn *Suaji ‘younger sibling’
> PMP *huaji ‘ibid.’]
mt-laqi ‘mother and son, father and son’ laqi ‘child’
(Sung and Shen 2006:260, fn. 22)

It should be noticed, in passing, that these forms are different from the Seediq reciprocal marker *m-C-*.

Seediq
(264) Wada mccebu dheya.
wada m-C-cebu dheya
PRETERIT REC-REC-shoot 3.PL
‘They fought.’
(Holmer 1996:34, quoted in Sung and Shen 2006:260, fn. 22, glossing modified)

In addition, the presence of the phoneme /t/ in the Tsou converse kinship pair prefix *nat-* suggests that the prefix may be related to the Seediq form *mt-* and WISP *maR-tali-. In this language, converse kinship pairs are encoded with the prefix *nat-* or *na-*, which is different from the reciprocal marker *yupa-*:

Tsou
(265) **nat-**’ohaesa
KIN-younger.brother
‘brothers and sisters’
(Sung and Shen 2006:261, glossing modified)

(266) **na-**’ina
KIN-mother
‘mother and daughter’
(Sung and Shen 2006:261, glossing modified)

It should be noted in passing, that other Formosan languages encode converse kinship pair constructions with other markers that do not include the forms *s(i)- or *t. For example, the

converse kinship pair forms in Amis and (Budai) Rukai are *maka- -ay* and *la-ma-*, respectively:

Amis

(267) **Mala**-kaka-**ay** ci Kacaw a ci Ofad.
REC-elder.sibling-AY PN Kacaw and PN Ofad
'Kacaw and Ofad are brothers.'

(Sung and Shen 2006:261, glossing modified. Personal names are capitalized in the Amis sentence)

(Budai) Rukai

(268) la-**ma**-taka
PL-REC-elder.brother/sister
'brothers and sisters (to each other)'

(Sung and Shen 2006:261, glossing modified)

5.1.3 Discussion

The regular sound correspondences between Proto-Malayo-Polynesian and Proto-Austronesian suggest that the Proto-Malayo-Polynesian form *maR-si- is a descendant of the Proto-Austronesian form *maR-si-. However, no reflex of this expected proto-form has been attested in our data. The only form that includes the form *si- and has some of the functions that some reflexes of the Proto-Malayo-Polynesian form *maR-si- have is the Kavalan form *sim-* in actor voice construction. The phonemes /si/ in this prefix suggests the Proto-Austronesian phonemes *si because this language reflects Proto-Austronesian phonemes *si as /si/, as in *siku* 'elbow'[114] (< PAn *siku 'elbow'; PMP *siku 'ibid.') and perhaps also *assim*[115] (PAn *asiN 'salty', with unexpected gemination of /s/ and unexpected reflex of *N (expected **asin).

In addition to the Kavalan forms, if *si- forms that denote converse kinship pairs are cognates of *si- for action-denoting roots in WISP languages, it is plausible to consider the segment *s* in the Seediq form *ms-* to be a cognate of Proto-Malayo-Polynesian *si-.

5.2 Oceanic *paRi-

This section examines the functions of reflexes of Proto-Oceanic *paRi- primarily in the data provided by Pawley (1973), Lichtenberk (2000), and Bril (2005), and explores how the Oceanic functions are expressed in WISP languages.[116]

In contrast to the Formosan data, our data show that no Oceanic languages have reciprocal, actor plural, or reflexive constructions with the form that contains a reflex of *si-. However, it is known that some Oceanic languages have reflexes of the 'reciprocal' *paRi- (Pawley 1973; Lichtenberk 2000; Brill 2005). Reflexes of *paRi- have reciprocal, actor plural (including 'distributive actor plural'), and reflexive functions. Some of these functions are also marked by reflexes of the Proto-Malayo-Polynesian *maR-si-, as in the following examples.

It should be noted that throughout this section, the formatives representing reflexes of *paRi- are glossed as PARI-. In some instances, these formatives are only monosyllabic or diverge otherwise very clearly from the assumed proto-form. The current exposition here follows the judgment of the specialists in Oceanic languages without further investigation of the presumed historical

114Data from Li and Tsuchida (2006:393).

115Data from Li and Tsuchida (2006:70).

116It should be noted that our listing of the functions is not exhaustive because we could not consult individual grammars of Oceanic languages.

relatedness:

Reciprocal

Mota

(269) Rar we **var-**vus.
they:DU ASPECT PARI-beat
'They two [i.e., the two of them] are beating each other.'
(Pawley 1973:150, glossing modified)

Nelemwa

(270) Hla **pe-**xua-i.
3.PL PARI-bite-I
'They bit one another.'
(Bril 2005:43, glossing modified)

WISP: Pendau

(271) Jimo **nosi**baro.
Jimo no-si-baro
3.PL.NOM AV.RLS-SI-argue
'They argued.'
(Quick 2003:298, glossing modified)

WISP: Cotabato Manobo

(272) **Eg-se-**tinudu-**ay** da='t belad.
AV-SI-point.at-EP 3.PL.NOM=LK hands
'They point at each other.'
(Kerr 1988:97, glossing modified)

WISP: Toba Batak

(273) **mar-si-**gulut 'to quarrel with each other' [gulut (mar-gulut) 'argue about something'[117]]
(274) (van der Tuuk 1971:140)

(275) **mar-si-**adu 'run in a race; be in a competition' adu 'chase; overtake'
(Nababan 1981:99)

Actor plural

East Futunan

(276) Kua **fe-**somo-'aki a 'ufi.
PERFECTIVE PARI-grow-APPL ABS yam
'Yams all grew at the same time.'
(Moyse-Faurie 2007:1530, glossing modified)

Nelemwa (New Caledonian)

(277) Hla **pe-**khuwo.
3.PL PARI-eat
'They eat together.'
(Bril 2005:46, glossing modified)

117Warneck (1977 [1906]:97).

WISP: Tagalog

(278) Silá'y **nagsi**alis.
sila ay nag-si-alis
3.PL.NOM INV AV.RLS-SI-leave
'They went away.'

(Lopez 1977a [1937]:91, my glossing)

(Distributive actor plural)

Standard Fijian

(279) Era **vei**-cici-yaki.
3.PL PARI-run-APPL
'They ran in all directions.'

(Schütz 1985:208, quoted in Lichtenberk 2000:39, my glossing)

WISP: Pangasinan

(280) **Magsi**kánta kayó.
mag-si-kánta kayó
AV-SI-sing 2.P.NOM
'Each of you sing.'

(Benton 1971:132, my glossing)

WISP: Aklanon

(281) **Nagsi**bakáł sanda it łambong.
nag-si-bakáł sanda it łambong
AV.RLS-SI-buy 3.PL.NOM GEN.INDEF shirt
'They each bought a different shirt.'

(Zorc 1968:135, my glossing)

Reflexive

East Futunan

(282) E kau **fe**-'umo pe kau moemiti fakatotonu pe le'ai.
NON-SPECIFIC 1.SG PARI-pinch or 1.SG dream really or not
'I pinch myself to know if I am dreaming or not.'

(Moyse-Faurie [2007:1522], quoted in Lichtenberk 2000:46, glossing modified)

Ajie (New Caledonian)

(283) Na **vi**-jiwé.
he PARI-kill
'He kills himself'

(Aramiou et al. 2001, quoted in Bril 2005:39, my glossing)

WISP: Salako

(284) … ià **sin**-soor ka tanàh, nangis.
… ià sin-soor ka tanàh, N-tangis
… 3 AV.SI-throw LOC floor AV-cry
'(If he asked permission to go to the earth and he did not get it,)
he would throw himself to the ground and cry.'

(Adelaar 2005c:88, glossing modified)

WISP: Toba Batak

(285) **mar-si-**jongjong ‘to raise oneself up’ jongjong ‘to stand’
(van der Tuuk 1971:140)

(286) **mar-si-**ájar ‘study [teach oneself]’ ájar ‘teach ([transitive])’
(Nababan 1981:99)

In addition, we can observe that the reflex of *paRi- also encodes the undergoer plural situation, namely, a situation with plural undergoers:

Undergoer plural

Nelemwa (New Caledonian)

(287) I **pe-**khi dooviu mahliili.
3.SG PARI-hit iron those.two
‘He hit the two metal pieces against each other.’
[? ‘The two metal pieces are made to hit against each other.’]
(Bril 2005:48, glossing modified)

Bwatoo (New Caledonian)

(288) A **ve-**cina xoot.
3.SG PARI-join rope
‘He joins the two ends of the rope.’
[? ‘The two ends of the ropes are joined by him.’]
(Rivierre et al. 2006, quoted in Bril 2005:48, glossing modified)

This use corresponds to one of the functions of the Proto-Malayo-Polynesian undergoer voice forms with *paR-si-, as in the Cotabato Manobo sentence below:

WISP: Cotabato Manobo

Undergoer plural

(289) **Eg-se-**tepeng**-en** ku sa elé.
UV-SI-measure-PV 1.SG.GEN NOM kunai.grass
‘I will even up the *kunai (cogon)* grass
(the bundles of grass cut for roofing are jostled to bring the stem ends uniformly together)’
[The bundles of *kunai* grass will be evened up by me.]
(Kerr 1988:99, glossing modified)

As such, Oceanic *paRi- forms functionally correspond to the two types of *si- forms in WISP languages. One is the actor voice *si- form *maR-si- (for the reciprocal, actor plural, and reflexive functions). The other is the unmarked voice *si- form *paR-si-.

5.2.1 With other types of roots and words

Some reflexes of the Proto-Oceanic *paRi- also appear with quality- and object-denoting roots, kinship terms, and location roots as well, in the same manner as the possible reflexes of *maR-si-. This section will provide the functions of such reflexes and compare them to the WISP counterparts.

5.2.1.1 With quality-denoting roots

In Nelemwa, the prefix *pe-* appears with a quality-denoting root to derive a word that indicates that the plural participants in subject function have the same quality denoted by the root:

Nelemwa (New Caledonian)
(290) **Pe-**ura-hli.
PARI-length-3.DU.POSS
'They (two) [i.e., the two of them] have the same length.'
(Bril 2005:50, glossing modified)

(291) **Pe-**rala-hla.
PARI-width-3.PL.POSS
'They are of the same width.'
(Bril 2007:1508, my glossing)

This function is also attested in Tagalog. However, while Tagalog distinguishes dual and plural formations, Nelemwa does not.

5.2.1.2 With object-denoting roots

When a reflex of the Proto-Oceanic *paRi- appears with an object-denoting root and is employed as a predicate, one of the functions of the derived words is to encode the situation where plural participants share the object indicated by the root:

Nelemwa (New Caledonian)
(292) **Pe-**kau=hli.
PARI-year=3.DU.POSS
'They (two) are the same age.'
(Bril 2005:50, glossing modified)

Cemuhi (New Caledonian)
(293) **Pi-**wödè-lé.
PARI-generation-3.PL.POSS
'They belong to the same generation.'
(Rivierre 1980:261, quoted in Bril 2005:50, glossing modified)

Paici (New Caledonian)
(294) **Pi-**nêê-ru mâ wë Kaapo.
PARI-name-3.DU.POSS with ART.PN Kaapo
'He has the same name as Kaapo.'
[? The two of them have the same name, namely, Kaapo.]
(Moyse-Faurie 2008:118, glossing modified)

Another use of the derived words is to encode the situation that the plural participants in subject function have the same quality of the object indicated by the root:

Nelemwa (New Caledonian)
(295) **Pe**-bala-hla.
PARI-partner-3.PL.POSS
'They are partners / in the same team.'
(Bril 2005:46, glossing modified)

Bwatoo (New Caledonian)
(296) ni tûûn a **ve**-bee-le
ART lineage REL PARI-ally-3.PL.POSS
'the lineages that are allies' (*bee-n* 'sibling, ally, friend')
(Bril 2005:48, glossing modified)

In addition, it appears that the Nelemwa *paRi- prefix *pe-* may also encode a singular associate of *bale* 'companion' (i.e., *pe-bale-t* 'an associate of companion, a "co-companion"') in the following sentence:

Nelemwa (New Caledonian)
(297) Kia **pe**-bale-t.
there.is.not PARI-companion-of.it
'The other (one of a pair, couple) is missing.'
(*pe-bale-t* here is the nominal argument of the verb *kia* ['there is not']).
(Bril 2007:1508, glossing modified)

When a *paRi- word with an object-denoting root is used as an argument or not as a predicate, it denotes the plurality of the object indicated by the root. Lichtenberk (2000:44) states that it indicates 'a group' or 'collective plural' of the item:

Standard Fijian
(298) **vei-**niu 'coconut grove or plantation' [niu 'coconut' (PMP *niuR 'coconut'), YK]
vei-vale 'group of houses' [vale 'house' (PMP *balai 'communal house'), YK]
(Lichtenberk 2000:44–5)

Nelemwa (New Caledonian)
(299) **pe**-hmawa-t 'patchwork' hmawa-t 'piece of [patchwork]'
(Bril 2007:1508, glossing modified)

As such, our data demonstrate that reflexes of the Oceanic *paRi- have some functions that are also indicated by the WISP formations with possible reflexes of *si-, namely, 1) sharing of an object by plural participants, 2) an associate sharing the quality of the object, and 3) the plurality of objects denoted by the root. Furthermore, the Nelemwa reflex of *paRi- also indicates the equality of the quality of the object among plural participants in subject function, as in the sentence *Pe-bala-hla* 'They are partners' (*bala* 'partner'). This meaning is not denoted by any *si- forms in WISP languages.

5.2.1.3 With kinship terms

Reflexes of *paRi- may appear with kinship terms, such as those indicating 'father', and indicate the converse kinship pair or group of the entity denoted by the term. For example, in Standard Fijian, the converse kinship pair words are formed with *vei- -ni*, as in the following example:

Converse kinship pair
Standard Fijian
(300) Erāū **vei-**tama-ni.[118]
3.DU PARI-father-NI
'They are father-and-child.'
(Schütz 1985:206, quoted in Lichtenberk 2000:44, glossing modified)

The Tobabaqita form *wai- -na* derives words indicating converse kinship pairs, such as 'mother and child', and converse kinship groups, such as 'mother and children', when it attaches to the root *thaina* 'mother':

Converse kinship pair and group
Toqabaqita
(301) **wai-**thaina-na
PARI-mother-NA
'mother's group, i.e., mother and her child/children (but not the father/husband)'
(Lichtenberk 2000:46, glossing modified)

Both converse kinship pair and group functions are indicated by *si- forms in WISP languages. However, the formal distinction between the converse kinship pair and group is not present in Oceanic *paRi- formations, in contrast to Pangasinan, which distinguishes the two forms. In addition to *si- forms, there are other formations, notably *maR-tali- formations, to denote these functions in WISP languages.

5.2.1.4 With location roots

Some reflexes of *paRi- appear with location roots and derive words that denote the symmetrical locations between dual participants in subject function with regard to the locational relationship indicated by the roots, as in the following Nelemwa example:

Nelemwa (New Caledonian)
(302) **Pe-**jeuk awôlô ma(h)leena.
PARI-be.near dwelling these
'These dwellings are close to one another.'
(Bril 2005:49, glossing modified)

It should be noted, in passing, that when the *paRi- prefix *pe-* is absent in this sentence, as in *Jeuk awôlô ma(h)leena,* the sentence signifies that 'these dwellings are close (to another reference point)' (Bril 2005:49).

The following examples are with the root that denotes 'back':

Iaai (New Caledonian)
(303) Ödru-mwe **ü-**hotuu-dru köu.
3.DU.PRES PARI-back-3.DU together
'They are back to back.'
(Ozanne-Rivierre 1984, quoted in Bril 2005:37, glossing modified)

118Schütz (1985) writes *vei-* as *vēī-* with macrons on the vowels. However, for simplicity, we employ the spelling *vei-* that is also used in Milner (1972).

Nengone (New Caledonian)

(304)	Buic	ci	**i**-co	jeu.
	3.PL	DURATIVE	PARI-back	together

'They are sitting back to back.' or: 'They are hostile to each other.'

(Dubois n.d., quoted in Bril 2005:37, glossing modified)

In addition, the Nengone symmetrical location verb *e-tada* ['be face to face'] with the prefix *e-*, an allomorph of *i-* that reflects *paRi-, allows the two-argument construction comprising a subject that encodes a singular participant and an argument that encodes the singular comitative participant of the singular participant in the subject:

Nengone (New Caledonian)

(305)	Bone	ci	**e**-tada	jeu	ne	Pua.
	3.SG	DURATIVE	PARI-front	together	COMITATIVE	Pua

'He's standing face to face with Pua.'

(Dubois n.d., quoted in Bril 2005:37, glossing modified)

The meaning of the symmetrical locative relationship indicated by *paRi- words corresponds to that indicated by the West Coast Bajau *si- words consisting of the prefix *si-* and location words, such as *si-bunda'* 'to face each other' (< *bunda'* 'in front of'). However, further data are required to investigate whether the following two syntactic functions are also possible for the West Coast Bajau prefix *si-*: 1) the function of taking plural participants in subject function and 2) that of taking a singular participant in subject function with a singular comitative participant argument.

5.2.2 Functions that are not attested in WISP languages

Some reflexes of *paRi- have functions that are not attested in *si- forms in WISP languages. The following subsections list such functions and their examples in Oceanic languages.

5.2.2.1 'Converse situation'

The term 'converse situation' refers to what Lichtenberk (2000:37) labels '[situation] where the participants are in a converse relation to each other'. In this situation, 'the roles of the relevant participants are not identical; rather, stand in a converse relation to each other' (Lichtenberk ibid.). In other words, a converse situation refers to a situation in which A is related in a certain way to B, and B is conversely related to A.

Lichtenberk (ibid.) observes that the Boumaa Fijian reflex *vei-* of the Proto-Oceanic prefix *paRi- is employed to indicate a 'converse situation':

Boumaa Fijian

(306)	Erau	sa	**vei**-'oti	ti'o	
	3.DU	ASPECT	PARI-cut	CONTINUATIVE	
	o	Sepo	vata	'ei	Elia.
	ART	Sepo	together	with	Elia

'Sepo and Elia are involved in an activity of (hair) cutting.'

(Dixon 1988:177, quoted in Lichtenberk 2000:37, glossing modified)

This example encodes, according to Lichtenberk (ibid.), 'the situation where the two people are in a

converse relation to each other: one is the agent in an action in which the other is the patient', in other words, a situation in which one is the cutter, and the other is 'the one whose hair is being cut'. This situation is neither reciprocal, such as two people cutting each other's hair, nor actor plural in an actor voice construction, such as two people cutting someone else's hair (Lichtenberk ibid.).

Cross-linguistically, this type of situations is also observed to be marked by a reciprocal marker in some languages (cf. Majid et al. 2011:42–43). For example, Hurst and Nordlinger (2011:83) report that the English sentence *two guys are chasing each other down a hall* is employed to describe the situation in which one person is chasing the other.

5.2.2.2 Repeated actions

Lichtenberk (2000:41) observes that some reflexes of *paRi- derive words indicating repeated actions, as in the following examples:

Samoan

(307) Na	**fe-**a'a	e	le	tama	le	polo.
PAST	PARI-kick	ERG	ART	boy	ART	ball

'The boy kicked the ball again and again.'
(Cook 1988:53, quoted in Mosel and Hovdhaugen 1992:180, quoted in Lichtenberk 2000:41)

Nelemwa

(308) I	**pe-**thalic.
3.SG	PARI-stumble

'She stumbles over and over again.'
(Bril 2005:53, glossing modified)

East Futunan

(309) Kua	**fe-**loi-saki	le	toe	ki	lona	tamana
PERFECTIVE	PARI-lie-APPL	ART	child	OBL	his	father

i	lona	ī.
OBL	his	fear

'The child lied over and over again to his father because he feared him.'
(Moyse-Faurie 2007:1532, glossing modified)

5.2.2.3 Fast actions (Intensive)

Bril (2005) observes that some reflexes of *paRi- derive words that indicate 'intensive' actions that are interpreted as fast actions. Consider the following examples:

Nelemwa (New Caledonian)

(310) Na	**pe-**diya	me	toven.
3.SG	PARI-do	depend	finish

'I did it fast to finish it.'
(Bril 2005:53, glossing modified)

(311) Kio	**pe-**top	ciic	hleny.
NEG	PARI-rot	wood	this

'This wood does not rot fast.'
(Bril 2005:53, glossing modified)

5.2.2.4 Intransitivization

The intransitivizing function here refers to the syntactic function of what Lichtenberk (2000:42) terms 'depatientive' function, which can be analyzed as a combination of two syntactic and semantic functions. The syntactic function is intransitivization, and this interpretation seems legitimate if we consider the examples provided by Lichtenberk. On the other hand, Lichtenberk (ibid.) states that in his 'depatientive' function encodes 'habitual or general' situations. This semantic function also seems to be present if we investigate the Oceanic data. However, it may also appear to be an epiphenomenon triggered by the use of the intransitivized words and the absence of their specific undergoers (in particular, patients and themes) because in Tagalog and some Central Philippine languages, at least, intransitive or intransitivizing formations do not necessarily indicate repeated, habitual, or generic actions. For this reason, this section deals with the intransitivizing function of *paRi- words in Oceanic languages.

The transitivity contrast between a root and a *paRi- word can be illustrated in the following Toqabaqita examples. In this language, the base *laba-ta'i* 'harm-TRANSITIVE' is transitive and takes an object. In the following example, it takes the object *nau ma maka nau* '[lit.] me and my father':

Toqabaqita

(312) Wane	e	laba-ta'i	nau	ma	maka	nau.
man	3.SG.NONFUTURE	harm-TRANSITIVE	1.SG	and	father	1.SG

'The man harmed me and my father.'

(Lichtenberk 2008:543)

However, when the prefix *kwai-*, (a doublet of) a reflex *wai-* of the Proto-Oceanic form *paRi-, appears with the base, it does not take an object. In addition, the derived word *kwai-laba-ta'i* indicates 'harm (people), spoil, damage (things)' that signifies that the action is not performed in contrast to the sentence with the underived base *laba-ta'i* in the previous example:

Toqabaqita

(313) Roo	wane	kero	**kwai**-laba-ta'i.
two	man	3.DU.NONFUTURE	PARI-affect.negatively-TRANSITIVE

'The two men harm (people), spoil, damage (things) etc.'

(Lichtenberk 2000:43, glossing modified)

In this construction, the actor can be singular:

Toqabaqita

(314) Wane	e	**kwai**-abingi.
man	3.SG.NONFUTURE	PARI-mistreat

'The man mistreats (others).'

(Lichtenberk 2000:43, glossing modified)

The Standard Fijian *paRi- form *vei-* has a function similar to the Toqabaqita intransitivizing function. In this language, the root *caqe* 'kick' appears as *caqe-t* 'kick-TRANSITIVE' when it takes a pronominal object suffix. In the following example, the third singular object pronoun *-a* in the predicate *caqe-t-a* 'kick something' cross-refers to the object *na polo* 'the ball':

Standard Fijian

(315) Āū ā caqe-t-a na polo.
1.SG PAST kick-TRANSITIVE-3.SG.OBJECT ART ball
'I kicked the ball.'

(Schütz 1985:322, glossing modified)

When the prefix *vei-* is attached to the root *caqe* 'kick' and the derived word *vei-caqe* 'kick (something)' modifies a noun, according to Schütz (1985:209), the habitual marker *dāū* is necessary:

Standard Fijian

(316) E tomata dāū **vei**-caqe.
3.SG person HABITUAL PARI-kick
'He is the person who plays a football.'

(Schütz 1985:209, quoted in Lichtenberk 2000:43, glossing modified)

5.2.2.5 Diminutive actions 'do a little'

The term 'diminutive action' refers to the action that is performed to a lesser degree than what its root would expect, or that is 'done a little'. It is termed 'attenuative meaning' in Bril (2005). In the following sentence, the action indicated by the root *ini* 'dig' is expected to be performed a little (cf. Bril 2000:52):

Cemuhi (New Caledonian)

(317) Go **pi**-ini a pwö-n!
2.SG PARI-dig ART hole-3.SG.POSS
'Dig the hole a little!'

(Rivierre 1980:259, quoted in Bril 2005:52, glossing modified)

5.2.2.6 Casual actions 'do casually'

The term 'casual action' refers to an action that is characterized by 'tentative undertaking done without any clear project, without any effort or specific intention, listlessly, with various degrees of success' (Bril 2007:1504). Consider the following examples in Nelemwa:

Nelemwa (New Caledonian)

(318) I u **pe**-kâlap mwamaidu.
3.SG PERFECTIVE PARI-lie.down over.there.down
'He lay down.' (having nothing else to do)

(Bril 2007:1504, my glossing)

(319) Va **pe**-diya fagau mwa.
1.PE PARI-make body house
'We have tried to build the wall of the house.'
(as might be, without any preconceived plan or idea)

(Bril 2007:1504, my glossing)

5.2.2.7 Solipsistic actions 'do alone'

In Nelemwa, the *paRi- form *pe-* also encodes 'solipsistic actions'. The term 'solipsistic action' refers to a situation in which an actor of an action performs the action for the actor's self, or on the actor's own (Bril 2005:55). Consider:

Nelemwa (New Caledonian)
(320) I **pe-**vhaa hada.
3.SG PARI-speak alone
'He speaks for himself, on his behalf, i.e., in his own name, not expressing a consensus.'
(Bril 2005:55, my glossing)

Bril (ibid.) states that the following sentence *i pe-vhaa* without *hada* 'alone' is also possible for indicating the solipsistic situation. However, it may also have intensive or durative readings:

Nelemwa
(321) I **pe-**vhaa.
3.SG PARI-speak
'He speaks and speaks.' or
'He speaks for himself.'
(Bril 2005:55, my glossing)

5.2.2.8 Spontaneous events

Some reflexes of *paRi- also derive spontaneous events of an inanimate participant in subject function, as observed by Lichtenberk (2000) and Bril (2005). In the following Oceanic examples, the participants in the subjects are inanimate and the *paRi- words encode the events that occurred on the inanimate participants:

Standard Fijian
(322) Sā **vei**-sau na draki.
ASPECT PARI-repay ART weather
'The weather has changed.'
(Milner 1972:111, quoted in Lichtenberk 2000:48, glossing modified)

Nelemwa (New Caledonian)
(323) **Pe-**nuk du bwa doo pwâ-mago.
PARI-fall DIRECTIONAL on earth fruit-mango
'Mangoes are falling (because they are ripe).'
(Bril 2005:51, glossing modified)

Cemuhi (New Caledonian)
(324) È **pi-**pènem a-li pomwa.
3.SG PARI-move ART.DEF door
'The door moves /works loose.'
(Rivierre 1980:259, quoted in Bril 2005:51, glossing modified)

5.2.2.9 Augmentative nouns

Bril (2005:37) reports that the Drehu *paRi- prefix *i-* derives augmentative nouns from object-denoting roots:

Drehu (New Caledonian)

(325)	**i-**puaka	‘big pig’	puaka	‘pig’
	i-laulau	‘large table’	laulau	‘table’
	i-hnaine	‘long hair’	hnaine	‘hair’

(Moyse-Faurie 1983:131–132; Sam 1995:84–88, quoted in Bril 2005:37)

5.2.2.10 Derived kinship terms

Lichtenberk (2000:46) observes that Toqabaqita has some kinship terms that optionally take the *paRi- prefix *wai-*. Consider:

Toqabaqita

(326) (**wai-**)funga
PARI-parent.in.law
‘father-in-law, mother-in-law’

(Lichtenberk 2000:46)

(327) (**wai-**)di’i nau
PARI-cross-cousin 1.SG
‘my cross-cousin’

(Lichtenberk 2007:1566, glossing modified)

Lichtenberk considers the root and the derived word to be synonymous. However, the presence of the prefix suggests that it may be an indication of a difference between the root and the derived word that existed in the past in the history of this language.

The Toqabaqita form *wai-* also has the converse kinship pair and group functions, as in *wai-thaina-na* ‘mother and her child/children’ (< *thaina* ‘mother’). This fact assumes that the entity indicated by this formation is a singular entity in a converse kinship pair, such as ‘child’ in the pair ‘mother and child’.

Such a word pair with a kinship term is attested in Hanunoo with the root *áriʔ* ‘any sibling younger than the speaker, i.e., younger brother or younger sister [i.e., younger sibling]’[119]:

Hanunoo (Philippine, GCP, South Mindoro)

Converse kinship pair

(328) **mag-tar**-áriʔ ‘(two) siblings’

(Conklin 1953:182)

Member of a converse kinship pair

(329)	**pa-tar**-áriʔ	‘sibling, i.e., sister or brother’
	pa-tar-áriʔ laláki	‘brother’ [< ‘male sibling’, laláki ‘male’]

(Conklin 1953:217)

The converse kinship pair form *mag-tar-áriʔ* ‘(two) siblings’ does not indicate the relative age distinction that is denoted by the root *árі* ‘younger sibling’, i.e., ‘younger’ in ‘younger sibling’. Likewise, no relative age distinction is indicated by the word *pa-tar-árіʔ* ‘sibling’. For this reason, we consider this function to be related to the converse kinship pair formation and to indicate a member of a converse kinship pair.

119Data from Conklin (1953:34).

5.2.2.11 'Together': with pronouns and demonstratives

Bril (2005:46) observes that the Nelemwa prefix *pe-* appears with dual or plural pronouns and demonstratives. The derived word indicates the unity of the plural objects indicated by the pronoun or demonstrative:

Nelemwa (New Caledonian)
(330) **Pe-**hî (hada).
PARI-1.DU.INCLUSIVE (only)
'We (two) together only.'

(Bril 2005:46, glossing modified)

(331) **Pe-**hleena, **pe-**hlaaleny, **pe-**hlaîdu, **pe-**yava.
PARI-these PARI-these PARI-these.down.there PARI-1.PL
'These together, these together, these down there together, we together.'

(Bril 2007:1508, glossing modified)

5.2.2.12 Grouping numeral 'two by two, two at a time'

The term 'grouping numeral' refers to numeral expressions such as 'two by two, two at a time'. This numeral is also referred to by the term 'distributive'; however, it is not employed here in order to avoid confusion with the term 'distributive' as in 'distributive actor plural construction'. In Bwatoo, the grouping numerals are expressed with the *paRi- prefix *ve-* (Bril 2005:46):

Bwatoo (New Caledonian)
(332) Nyi **ve-**lu.
put PARI-two
'Put two by two.'

(Rivierre et al. 2006, quoted in Bril 2005:46, glossing modified)

5.2.2.13 'Middle point' and 'medium size'

When the word meaning 'center, middle' appears with a reflex of *paRi-, the derived word indicates 'a symmetrical, middle point or space between two landmarks or objects' (Bril 2005:50):

Cemuhi (New Caledonian)
(333) a **pi-**wième-n
ART PARI-middle-3.SG.POSS
'(It is) the one in-between.'

(Rivierre 1980:262, quoted in Bril 2005:50, glossing modified.)

Nelemwa (New Caledonian)
(334) ni **pe-**wooxa-t
in PARI-middle
'(It is) in-between.'

(Bril 2005:50, glossing modified)

In addition, the Bwatoo *paRi-MIDDLE word *ve-daboo* indicates the brother or child in the middle with regard to the birth order among brothers or children, respectively.

Bwatoo (New Caledonian)
(335) **ve**-daboo-n
PARI-middle-3.SG.POSS
'(he is) the (brother/child) in the middle'
(Rivierre et al. 2006, quoted in Bril 2005:50, glossing modified)

The use of *paRi-middle words is supposed to be somehow related to the actor plural or reciprocal use of *paRi- forms with action-denoting roots or the use of *paRi- forms with locative words or quality-denoting roots because the idea of 'middle' and 'center' presupposes the presence of two or more locational or conceptual points.

Along these lines, Bril (2005:50) observes that the Bwatoo word *ve-goo-n* 'be of medium size', as in *balo ve-goo-n* 'a medium sized ball' [*balo* 'ball'] consists of the *paRi-prefix *ve-* and the quality-denoting root *goo-n* 'size'.

5.2.3 *paRi- and reduplication

Reflexes of *paRi- also appear with reduplication to derive words from action-denoting roots. For example, the Boumaa Fijian root *siivi(-ta)* 'pass, exceed' appears as the *paRi- form *vei-siivi-ti* or *vei-sivi* ['pass each other (once)'] when the subject encodes plural actors. Compare:

Boumaa Fijian
Underived

(336)	E	aa	siivi-ti	au	o	Jone	mai	Viidana.
	3.SG	PAST	pass-I	1.SG	ART	John	at	Viidana

'John passed me at Viidana.'
(one person passing another on the road, both going in the same direction)
(Dixon 1988:178, my glossing)

*paRi- form (*vei-*)

(337)	'eirau	aa	**vei**-siivi-ti/vei-sivi	(o	yau)	'ei	Jone
	3.DU	PAST	PARI-pass-I/PARI-pass	ART	1.SG	with	John

mai Viidana.
at Viidana
'John and I passed each other at Viidana.'
(two people traveling in opposite directions meet and pass on the road)
(Dixon 1988:178, my glossing)

The reduplicated *vei-* word *vei-sii-sivi*[120] ['pass each other (many times)'] is employed 'to describe two people passing each other alternately, each going in the same direction' (Dixon 1988:178). This use of reduplication suggests the plurality of events because the action of 'passing each other' occurs many times.

In addition, it should also be mentioned that the formation with a reflex of *paRi- and reduplication is also observed to appear with object-denoting roots. In Boumaa Fijian, the object-denoting root *vale* 'house' can take both the formations *vei-* and *vei-* with full root reduplication. However, Dixon (1988:176, 198) glosses both as 'every house', and we cannot observe their difference even in sentence context. Consider:

120The long vowel /ii/ in *siivi* 'is shortened if it falls in a penultimate syllable' (Dixon 1988:178).

Boumaa Fijian
*paRi- (*vei-*) form

(338)	’eitou	aa	’ana	i	na	**vei**-vale
	1.PAUCAL.EXCL	PAST	eat	in	art	PARI-house
	ta’uco’o	sara	i	na	’ora	yai.
	all	very	in	ART	village	this

‘We have eaten in every house in a village.’

(Dixon 1988:99, spelling and glossing modified)

*paRi- (*vei-*) with reduplication

(339)	Eratou	dui	i	na	**vei-vale**-vale.
	3.PAUCAL	respectively	in	ART	PARI-RDP-house

‘Each of them lives in a different house.’

(Dixon 1988:100, spelling and glossing modified)

5.2.4 Discussion

This section discusses the forms and functions of the putative reflexes of *paRi- in relation to the *si- formations in WISP languages.

Some authors, who explore the functions of reflexes of *paRi-, consider some forms in other Austronesian languages to be cognates of the Proto-Oceanic prefix *paRi-. For example, Pawley (1973) regards the Proto-Malayo-Polynesian *maR- as its cognate. On the other hand, Ross suggests that it is the Proto-Malayo-Polynesian *maki- that is the cognate of *paRi-. On the other hand, Lichtenberk (2000:58) assumes Proto-Southern-Tsouic *pari- to be a cognate in Formosan languages.

The current study suggests the possibility that some reflexes of particular functions may reflect Proto-Malayo-Polynesian *maR-si-, and, to a lesser degree, *paR-si-, due to the phonological and semantic correspondences between reflexes of the Proto-Oceanic and Proto-Malayo-Polynesian forms. However, our hypothesis has some formal and functional issues that need to be addressed.

5.2.4.1 Formal issues

In principle, the demonstration of the cognacy of given two forms in different languages requires evidence for the presence of the regular phonological correspondences between the two languages in the two forms, as discussed in Chapter 1. From a phonological point of view, however, Proto-Oceanic *paRi- is not an expected reflex of Proto-Malayo-Polynesian *maR-si-. For this reason, this section primarily points to a few issues for which such a hypothesis would need to provide plausible accounts. These issues include 1) the unexpected reflection of the Proto-Malayo-Polynesian *si- forms, primarily *maR-si- forms, 2) the irregular reflexes of the segment*aRi- in the proto-form *paRi-, and 3) irregular reflexes of *p in *paRi-.

The expected Proto-Oceanic reflex of the Proto-Malayo-Polynesian actor voice *si- form *maR-si- in Proto-Oceanic would be **ma-si- because of 1) the regular phonological correspondences between the Proto-Malayo-Polynesian and Proto-Oceanic languages and 2) the fact that the Proto-Malayo-Polynesian phoneme cluster *Rs is reflected as the Proto-Oceanic phoneme *s, as in PMP *beRsay ‘canoe paddle’ being reflected as Proto-Oceanic *pose ‘ibid’.

However, it is also plausible to assume that the reflex is **pa-si- because Oceanic languages lack

productive voice alternations (Ross 2004:507). The form *maR-si- contains the segment /m/ that can be traced back to the actor voice infix *<um>. Accordingly, the expected Oceanic reflex **ma-si- (< *p<um>a-si-) also contains the voice-marking segment /m/. Such m-initial voice-marked forms are unlikely in Oceanic languages.

In contrast to most WISP languages, the action-denoting roots lack voice markers, such as the actor voice *paR- class prefix *maR- and its patient voice counterpart *paR- -en, when they are used as predicates in sentences. Instead of the voice markers, however, the action-denoting roots have inherent lexical properties that determine whether the roots take an actor or undergoer subject. For example, the Longgu root *soko* 'chop' is an actor-subject verb and causes its subject *sa makaria* 'the boy' to have the actor role in the following intransitive and transitive sentences:

Actor-subject verb
Longgu (Oceanic)

(340) Soko sa makariva.
chop ART.SG boy
'The boy chopped.'
(Davis 1997:§5.2.5, quoted in Evans 2003:29, glossing modified)

(341) Soko-a sa makariva sa qato.
chop-TRANSITIVE:3.SG ART.SG boy ART.SG tree
'The boy chopped the tree.'
(Davis 1997:§5.2.5, quoted in Evans 2003:29, glossing modified)

On the contrary, the root *tuke* 'be thrown away' is an undergoer-subject verb and causes its subject *sa leboko* 'the bush knife' to be the undergoer (theme) in the following intransitive and transitive sentences:

Undergoer-subject verb
Longgu (Oceanic)

(342) Tuke sa leboto.
be.thrown.away ART.SG bush.knife
'The bush knife was thrown away.'
(Davis 1997:§5.2.2, quoted in Evans 2003:30, glossing modified)

(343) Tuke-a rao sa leboto.
be.thrown.away-TRANSITIVE:3.SG.OBJECT 1.SG ART.SG bush.knife
'I threw the bush knife away.'
['The bush knife was thrown away by me.']
(Davis 1997:§5.2.2, quoted in Evans 2003:30, glossing modified)

Even when a derivational prefix is attached to an action-denoting root, the derived word does not have a voice marker in Oceanic languages. For example, in the majority of WISP languages, the voice of the causative words with reflexes of the Proto-Austronesian causative prefix *pa- is marked by some formations, such as the actor voice causative form *maR-pa- and the thematic voice causative form *i-paR-pa-. In Tagalog, causative formations are marked with a voice marker in the same manner as non-causative formations. Compare the following examples. Both the non-causative word *nag-labas* 'brought' and the causative word *nag-pa-labas* 'had someone bring' are marked with the actor voice realis mood *pag-* class prefix *nag-*, which originates from the earlier form *p<um><in>ag- that consists of the action-denoting lexical class prefix *pag-*, the actor voice

infix <*um*>, and the realis mood infix <*in*>:

Tagalog

(344)	Si	Juan	ay	**nag**labás	nang	katre	sa	silíd.
	si	Juan	ay	nag-labas	nang	katre	sa	silid
	PN.NOM	Juan	INV	AV.RLS-bring	GEN	bed	LOC	room

'Juan brought (took) a bed out of the room.'

(Lopez 1977a [1937]:54, my glossing)

(345)	Si	Juan	ay	**nagpa**labás	nang	katre	sa	silid.
	si	Juan	ay	nag-pa-labas	nang	katre	sa	silid
	PN.NOM	Juan	INV	AV.RLS-CAUSATIVE-bring	GEN	bed	LOC	room

'Juan ordered brought (taken) out a bed from the room.'

(Lopez 1977a [1937]:54, my glossing)

By contrast, Tinrin, an Oceanic language, does not have such actor voice marking; however, it does show the actor- and undergoer-subject verb distinction. For example, the root *soghe* 'stab' is an actor-subject verb, and takes an actor as a subject. In the following example, the subject is *nrâ mwié* 'a woman (SUBJECT)'

Tinrin

Underived

(346)	Nrâ	soghe	Toni	nrâ	mwié.
	3.SG	stab	Tony	SUBJECT	woman

'A woman stabbed Tony.'

(Osumi 1995:117, quoted in Evans 2003:244, spelling and glossing modified. The personal name is capitalized.)

When the causative prefix *fa-* appears with the root *soghe*, the derived word *fa-soghe* 'have someone stab' takes the actor of the causative event as a subject. In the following example, the subject is *nrâ Saarri* 'Charlie (SUBJECT)':

Tinrin

Causative

(347)	Nrâ	**fa**-soghe	Toni	nrâ	Saarri	nri	treanrti.
	3.SG	CAUSATIVE-stab	Tony	SUBJECT	Charlie	by	person

'Charlie got someone to stab Tony.'

(Osumi 1995:117, quoted in Evans 2003:244, spelling and glossing modified. Personal names are capitalized.)

Likewise, the initial phoneme of a reflex of *paRi- appears to be *p* even if the reflex encodes a reciprocal or actor plural construction, which is usually indicated by an *m*-initial form in WISP languages. Compare the following Nelemwa examples. The *paRi- reciprocal word *pe-yage-i* 'help each other' in the second example is not marked with an actor voice affix:

Nelemwa

Underived with *yagei* 'help'

(348)	Hli	yagei-hli	a	hliili	meewu.
	3.DU	help.TRANSITIVE-3.DU	AGENT	these.two	brother

'These two brothers help them.'

(Bril 2007:1487, glossing modified)

*paRi- as a reciprocal marker: equivalent to *maR-si-

(349) Hli **pe**-yage-i hliili meewu.
3.DU PARI-help-I these.two brother
'These brothers help each other.'

(Bril 2007:1487, glossing modified)

As such, when we compare the actor voice forms in WISP and Oceanic languages, we can find the absence of the initial *m- that functions as an actor voice indicator. For this reason, we can also expect the Proto-Oceanic form **pa-si- for the forms that correspond to actor voice form *maR-si- in WISP languages.

In addition, given that the Proto-Malayo-Polynesian word-medial consonant cluster *Rs is reflected as Proto-Oceanic *s, the expected Proto-Oceanic reflex of Proto-Malayo-Polynesian *maR-si- should be **pa-si-, not *paRi- (or *paR-i-). However, since no reflex of *paRi- in modern Oceanic languages contains reflexes of Proto-Oceanic *s, we cannot change the proposal to reconstruct *paRi-. At this point, we do not have a solution to offer for the irregularity regarding reflexes of the Proto-Malayo-Polynesian consonant cluster in Oceanic languages.

Furthermore, as mentioned previously, Proto-Oceanic *paRi- also exhibits some phonological irregularities in its reflexes in its daughter languages. For example, Blust and Trussel (ongoing) observe that some reflexes of *paRi- appear to have immediate proto-forms *pai-, *pe-, and *pa- without the phoneme *R. In addition, we can observe the absence of the initial *p in some forms, as in Drehu *i-,* Nengone *e-,* and Iaai *ü-* (data from Bril 2005).

While other authors regard these various formations as unexpected reflexes of *paRi-, Ross considers that *paRi- has the doublet *pa(k)i- and attributes some forms to the doublet. He points out that the Proto-Oceanic prefix *paRi- became the Proto-New-Ireland prefix *var- with the following data (Ross 1988:284):

Evidence for Proto-New Ireland *var-

(350) Minigir, Tolai, Bilur, Nalik *var-*, Parpatar, Label *har-*, Lihir *her-*, Sursurunga, Siar *ar-*
Tangga *fa-*, Konomala *fə-*, Lavonggai *aŋ-*

(Ross 1988:284)

Some of the *paRi- words in these languages are as follows:

Tolai

(351) **var**-mari 'be in love' mari 'love (s.o.)'

(Ross 1988:284)

(352) **var**-ubu 'hit each other' ubu 'hit'

(Mosel 1984:146)

Patpatar

(353) **har**-kata 'spear each other' kata 'spear (s.t.)'
har-ubu 'fight with each other' ubu 'hit'

(Ross 1988:284)

Lihir

(354) **her**-cume1 'quarrel' cumer- 'be cross with (s.o.)'

(Ross 1988:284)

The expected form of *paRi- in this proto-form is **vari-, which is supported within the Meso-Melanesian cluster by Vitu, Uruava, and Roviana *vari-*, Torau *ari-,* and Maringe *fari-*. Ross identifies this phenomenon as innovatory loss of the final vowel /i/ in *paRi-. Furthermore, he also identifies Tangga *fa-* and Konomana *fə-* as reflexes of Proto-New-Ireland *var- because these languages reflect the Proto-New-Ireland phoneme *r (< Proto-Oceanic *R) as zero.

Furthermore, Ross points out the absence of *R in some putative reflexes of *paRi-, as in the following examples:

(355) Absence of *R in *paRi- in some Oceanic languages

New Ireland languages:	Duke of York *vai-*, Madak, Tabar, and Mandak *ve-*, Kandas *ai-*, Lamasong, Barok, and Tigak *e-*
Western Melanesian languages	
Meso-Melanesian cluster:	Nakanai *vai-*, Solos *he-*, Petats, Halia (Haku, Hanahan) *hi-*, Banoni, Piva *vai-*
North New Guinea cluster:	Manam *e-*, Kairiru *i-*
Papuan Tip cluster:	Dobu, Sewa Bay *e-*, Duau, Suau (Sariba) *he-*, Sinagoro, Keapara *ve-*, Motu *he-*, Lala *vi-*, Kuni *bai-*

(Ross 1988:284)

In contrast to his suggestion concerning the loss of the phoneme *i of POc *paRi- for the Proto-New-Ireland form *var-, however, he does not attribute the absence of the phoneme *R in some forms, such as *vai-, to the loss of the phoneme *R in POc *paRi-. Instead, he suggests the Proto-New-Ireland doublet *vai-, which can be traced back to POc *pa(k)i-.

In addition, it should be noted that POc *paRi- may appear as *vaRi- and *faRi- in its daughter languages, in the same manner as the phoneme *p appears as *v or *f in some reflexes of the POc *p- initial lexical items. These *p-*v and *p-*f correspondences are due to the phonological change that Ross (1988) terms ‘lenition’. This phonological change is not predictable from the regular phonological correspondences between Proto-Malayo-Polynesian and Proto-Oceanic because, according to Ross (1988:48), it ‘occurred independently at different times and places after the break-up of [Proto-Oceanic]’.

Another problem pertains to the morphemes that often co-occur in *maR-si- formations in WISP languages. Most notably, no *paRi- formations have reflexes of the PMP event plural suffixes *-an, *-ay, and *-anay. In place of the suffixes, for example, the plurality of events is indicated by CV(V)- reduplication in Boumaa Fijian.

Finally, we must note the general absence of lexical class prefixes in the Oceanic *paRi- words. The regular phonological correspondences between Proto-Malayo-Polynesian and Proto-Oceanic expect Proto-Malayo-Polynesian prefix *paR- to be Proto-Oceanic *paR-. However, lexical class prefixes that reflect *paR- and are placed after reflexes of *paRi- have not been observed. On the other hand, the homorganic nasal substitution is not productive in Oceanic languages (Dempwolff 1937, quoted in Blust 2004:75), although it is observed fossilized in some lexical items, such as Proto-Oceanic *panako ‘thief’ (< Proto-Malayo-Polynesian *paŋ-takaw [*paN-takaw] (*panakaw) ‘thief, steal’). In *paRi- forms, no reflexes that reflect *paN- after the reflexes of *paRi- have been attested.

5.2.4.2 Functional issues

As demonstrated above, reflexes of the Proto-Oceanic form *paRi- have various functions. Some of

these functions are identical to those indicated by *si- formations in WISP languages; however, others are not. The following table compares the functions of POc *paRi- and WISP *si- forms. It should be noted that the functions only attested in New Caledonian languages are indicated in the table because of their unique functions. It should also be noted that, unless otherwise mentioned, the corresponding WISP forms are in actor voice constructions.

Table 44: Functions of reflexes of Proto-Oceanic *paRi- in Oceanic languages and *si- forms in WISP languages

	Oceanic examples	Examples of corresponding *si- forms forms in WISP languages
with action-denoting roots		
Reciprocal	Mota *var-* Nelemwa *pe-*	Malay *ber-si-, ber-se-* Toba Batak *mar-si-* Cotabato Manobo *eg-se- -ay*
Actor plural	Nelemwa *pe-*	Tagalog *mag-si-([pag-/paN-])* Muna *si-CVCV -ha*
Actor dual	N/A	Muna si- -ha
Distributive actor plural	Standard Fijian *vei-*	Cebuano *(mag-)(i)sig-([pa-])*
Reflexive	East Futunan *fe-*	Toba Batak *mar-si-*
Undergoer plural	Nelemwa pe-	Cotabato Manobo *eg-se- -en* (patient voice)
'Converse' situation	Boumaa Fijian *vei-*	N/A
Repeated actions	Samoan *fe-* Nelemwa *pe-* East Futunan *fe-*	N/A
Fast actions (intensive) (New Caledonian)	Nelemwa *pe-*	N/A
Intransitivization	Toqabaqita *kwai-* Standard Fijian *vei-*	N/A
Diminutive actions 'do a little' (New Caledonian)	Cemuhi *pi-*	N/A
Casual actions 'do casually' (New Caledonian)	Nelemwa *pe-*	N/A
Solipsistic actions 'do on my own' (New Caledonian)	Nelemwa *pe-*	N/A
Spontaneous events	Standard Fijian *vei-* Nelemwa *pe-*	N/A
with quality-denoting roots		
Equal quality between plural participants (New Caledonian)	Nelemwa *pe-*	Tagalog *mag-ka-ka-sing-*, *mag-ka-sing-CV-*
Equal quality between dual participants (New Caledonian)	Nelemwa *pe-*	Tagalog *mag-ka-sing-*

Different degrees of quality between plural participants	N/A	Karo Batak *si-(RDP-) -en* Toba Batak *mar-si- -i*
Competitive	N/A	Malay *ber-si-*
with object-denoting roots		
Plural participants sharing an object (New Caledonian); e.g., '(plural participants) belong to the same class.'	Nelemwa *pe-*	Muna *si-CVCV-*
Dual participants sharing an object; e.g., '(dual participants) belong to the same class.'	N/A	Muna *si-*
Plural participants sharing the quality of an object; e.g., '(plural participants) are teachers.'	Nelemwa *pe-*	N/A
Associate sharing the quality of the object	Nelemwa *pe-*	Cebuano *i-sig-ka-* Pamona *ka-si-*
Associate sharing an object	N/A	Aklanon *i-sig-ka-* Pamona *ka-si-* Malay *se-*
Plural items	Standard Fijian *vei-* Nelemwa *pe-*	Cebuano *ma-sig-ka-*
Dual items	N/A	(Cebuano *ma-sig-ka-*)
Augmentative nouns (New Caledonian (Loyality Island))	Drehu *i-*	N/A
with kinship terms		
Converse kinship pair (dual)	Standard Fijian *vei- -ni*	Cebuano *ma-sig-ka-* Pangasinan *san-*
Converse kinship group (plural)	Toqabaqita *wai- -na*	Pangasinan *san-CV-*
Derived kin term	Toqabaqita wai-	N/A
with location words/roots		
Symmetrical relationship between two participants encoded in subject function (New Caledonian)	Nelemwa *pe-* Nengone *e-*	West Coast Bajau *si-* (no syntactic information)
Symmetrical relationship between two participants encoded in subject and comitative functions (New Caledonian)	Nengone *e-*	
with numerals		
Grouping 'X by X, in X's'	Bwatoo *ve-*	N/A

(New Caledonian)		
with pronouns and demonstratives		
‘Together’ (New Caledonian)	Nelemwa *pe-*	N/A
with specific words		
Mediam size with the root ‘size’ (New Caledonian)	Bwatoo *vi-*	N/A
Locative middle point with the root ‘middle’ (New Caledonian)	Nelemwa *pe-*	N/A

Some functions indicated by Oceanic *paRi- forms are not denoted by WISP *si- forms. Among such functions, it is worth mentioning that there are a variety of functions that have been attested only in New Caledonian languages. Bril (2005) attributes the broad polysemy of *paRi- forms in Northern New Caledonian languages (e.g., Nelemwa, Bwatoo, Cemuhi, and Paici) to the loss of the forms that combined with reflexes of *paRi-, and to the subsequent transfer of their functions to the reflexes of *paRi-. She assumes that the forms that were lost are 1) reduplication in *paRi- with reduplication formations and 2) reflexes of the applicative suffixes *-i and *-akini in the *paRi- -i and *paRi- -akini formations. This argument is based on the following observation concerning the correlations between various *paRi- formations and their functions in Oceanic languages:

Table 45:Some of the functions indicated by different types of *paRi- formations in some Oceanic languages (Bril 2005:28; terminology and order modified)

	*paRi-	*paRi- -i	*paRi- -akini	*paRi- with root reduplication
with action-denoting roots				
Actor plural (‘Collective or associated actors, multiple to an action’)	+	+		
Reciprocal	+	+	(+, but infrequent combination)	
Undergoer plural	+			
‘Distributive, dispersive’ Distributive actor plural (with multiple locations), casual action			+	+
Intransivization	+	+		+
Repeated or intensive actions	+	+		+
‘Reference to states and properties (often derived from frequentative or habitual actions)’ [syntactically another use of the functions of repeated	+			+

actions or intransitivization][121]				
with roots not indicating actions				
'Collective/plural entities, grouping, augmentative' Plural items, augmentative (with object-denoting roots)	+			+
'Comparison and symmetry' Situation in which plural actors share an object or the same quality (with object- and quality-denoting roots)	+			

It appears plausible to consider some reduplicative patterns to be among the sources of the meanings indicated by the *paRi- reflexes in the Northern New Caledonian languages. According to Bril (2005:60), '[r]eduplication generally indicates plurality of relationship, collective actions, non-completed, iterative or habitual actions, stative properties, tentative, distributive, separative, random actions, but it also expresses intensity, augmentation or diminution/attenuation, comparison of similarity [in Oceanic or Austronesian languages (?)]. It also has intransitivizing and nominalizing functions.' She continues that 'such functions are partly similar' to those indicated by the *paRi- reflexes in Northern New Caledonian languages without productive reduplication. She does not adduce any examples of reduplication in support of her statement concerning the functions of reduplication. However, it is observed that some functions of New Caledonian *paRi- forms are indicated by reduplicative patterns in WISP languages, such as Tagalog diminutive *mag-RDP-* (e.g., *mag-lakad-lakad* 'do a little walking' (< *l<um>akad* 'walk')[122], Malay diminutive *ber-RDP-* (e.g., *ber-jalan-jalan* 'walk about, go for a walk' (< *jalan* 'walk')[123].

However, it is unlikely that some of the functions originate ultimately from the reflexes of the applicative suffixes *-i and *-akini because their primary function is to introduce a new argument of a specific semantic role. Compare the following examples in Longgu. When the root *ango* 'crawl' takes the suffix *-vi* (< *-i), the derived word *ango-vi* 'crawl to, crawl for' takes the goal argument *vanga ngaia* 'its food'. When, on the other hand, the root takes the suffix *-tai'ni* (< *-akini), the derived word *ango-ta'ini* 'crawl with' takes the comitative argument *gale ngaia-gi* 'its babies':

Longgu
Root *ango* 'crawl' only
(356) Mwaa-i e ango.
snake-SG 3.SG crawl
'The snake is crawling.'
(Hill 1992:58, quoted in Evans 2003:122, spelling and glossing modified)

With *-i: Adding the goal argument *vanga ngaia* 'its food'
(357) Mwaa-i e ango-**vi**-ra vanga ngaia.
snake-SG 3.SG crawl-I-3.PL food 3.SG
'The snake crawled to/for its food.'
(Hill 1992:59, quoted in Evans 2003:173, spelling and glossing modified)

121This category apparently refers to the generic reading of the function indicating repeated actions or the intransitivizing function, as in the Standard Fijian sentence *E dau vei-vuke*. 'He often helps (as a general property).' 3.SG HABITUAL PARI-help) (Schütz 1985:208, quoted in Bril 2005:57).
122Data from Schachter and Otanes (1972:340).
123Data from Sneddon (1996:20).

With *-akini: Adding the comitative argument *gale ngaia-gi* 'its babies'

(358) Mwaa-i e ango-**ta'ini**-ra gale ngaia-gi.
snake-SG 3.SG crawl-AKINI-3.PL baby 3.SG-PL
'The snake crawled with its babies (on its back).'
(Hill 1992:59, quoted in Evans 2003:173, spelling and glossing modified)

The Oceanic applicative suffixes also appear with reflexes of *paRi-. In some languages and with some root types, the three types of *paRi- forms, namely, *paRi-, *paRi- -i, and *paRi- -akini, do not exhibit semantic differences among themselves. For example, East Futunan has the formations *fe-*, *fe- -Ci*[124], and *fe- -Caki,* which can formally be traced back to POc *paRi-, *paRi- -i, and *paRi- -akini, respectively: All the formations encode actor plural situations, as in the following examples:

East Futunan
*paRi-: Actor plural

(359) E **fe**-taki lalā kete
IMPERFECTIVE PARI-carry.in.hand their bag

e Setefano mo Samino.
ERG Setefano and Samino
'Setefano and Samino are carrying their bag together.'
(Moyse-Faurie 2007:1521, glossing modified)

*paRi- -i: Actor plural

(360) E **fe**-sola-**ki** a toe
IMPERFECTIVE PARI-run.away-I ABS child

o nono i le salatamu.
in.order.to hide OBL ART policeman
'The children are running [away] together […] to hide from the policeman.'
(Moyse-Faurie 2007:1524, glossing modified)

*paRi- -akini: Actor plural

(361) Kua **fe**-somo-**'aki** a 'ufi.
PERFECTIVE PARI-grow-AKINI ABS yam
'Yams all grew at the same time.'
(Moyse-Faurie 2007:1530, glossing modified)

In Nelemwa, reciprocal words are formed with either *pe-* or *pe- -i*, depending on the types of verbs and those of arguments (Bril 2005:42):

124The symbol C in the applicative suffixes refers to the final consonant that was etymologically present in the root to which the suffixes attach. See Evans (2003) for further information.

Table 46:Nelemwa reciprocal formations (cf. Bril 2005:42)

	One nominal argument	One pronominal argument	Two coreferential subject/object pronouns
direct transitive verb	pe- -i	pe- -i	pe-
indirect transitive verb	pe- (-i)	pe-	pe-
intransitive verb	pe-	pe-	N/A

Nelemwa
pe- reciprocal construction with two pronominal coreferential arguments *hla* '3.PL' and *-hla* '3.PL'

(362) Hla **pe**-xua-hla.
3.PL PARI-bite-3.PL
'They bit one another.'

(Bril 2005:43, glossing modified)

pe- -i reciprocal construction with one nominal argument (*tavia 'dog'*)

(363) Hla **pe**-xua-i tavia.
3.PL PARI-bite-I dog
'The dogs bit one another.'

(Bril 2005:43, glossing modified)

pe- -i reciprocal construction with one pronominal argument *hla* '3.PL'

(364) Hla **pe**-xua-i.
3.PL PARI-bite-I
'They bit one another.'

(Bril 2005:43, glossing modified)

However, it is also true that the applicative suffixes are employed to differentiate some *paRi- functions from other functions. For example, Nelemwa employs the form *pe-* to derive the actor plural forms and the form *pe- -i* to derive reciprocal forms from intransitive roots:

Nelemwa
With *pe-* (< *paRi-): Actor plural
(365) Hla **pe**-kulux.
3.PL PARI-hide
'They play hide and seek.'

(Bril 2005:44, glossing modified)

With *pe- -i* (< *paRi- -i): Reciprocal
(366) Hla **pe**-kulux-i hla.
hla pe-kuluk-i hla
3.PL PARI-hide-I 3.PL
'They hide from each other.'

(Bril 2005:44, glossing modified)

Other formal differences of Nelemwa formations with *pe-* and their conditions are summarized

in the following table:

Table 47:Distribution of the Nelemwa forms *pe-* and *pe- -i* (cf. Bril 2005:43)

	Reciprocal (‘reciprocal subject’)	Undergoer plural (‘reciprocal object’)	Actor plural (‘associative, collective’)	Other meanings (dispersive, intensive, spontaneous)
pe- + transitive verb	pe- -i	pe-	pe-	pe-
pe- + intransitive verb	pe-	N/A	pe-	pe-

To summarize, it may be the case that some of the diversified functions encoded by reflexes of Proto-Oceanic *paRi- originate from other additional morphological items, such as reduplication and affixes, that appear or appeared with the reflexes, but not the reflexes themselves. The current study does not precisely identify which functions are due to such functional transfer from other affixes. However, considering that the New Caledonian functions are innovations, we regard the remaining Oceanic functions for action-denoting roots as the primary diagnostic functions that will be employed in order to identify a Proto-Malayo-Polynesian cognate of the Proto-Oceanic prefix *paRi- in the following section.

5.2.5 Possible cognates suggested by other authors

This section discusses the two Proto-Malayo-Polynesian forms that are suggested or regarded as the earlier form of the Proto-Oceanic form *paRi- by other authors. One is PMP *maki- and the other is PMP *maR-.

*5.2.5.1 PMP *maki-*

As mentioned previously, Ross (1988) posits that the Oceanic *paRi- has a doublet *pa(k)i-, based on the attestations of the *R-less forms, such as *vai-, *ve-, and *vi- (which may reflect POc **pai-, **pe-, and **pi-) in New Ireland and Western Melanesian languages. Furthermore, offering some examples of the cognates of the putative doublet *pa(k)i- in some Philippine languages and Toba Batak (a Western Indonesian language), he points out the formal and semantic similarities between Proto-Malayo-Polynesian *maki- and its inflected forms, and considers them to be the source of his Oceanic doublet *pa(k)i-. The following examples are the Philippine and Toba Batak examples that he adduces:

Tagalog
(367) **maki**-kain ‘eat with (s.o.)’
(Schachter and Otanes 1972:333–334, quoted in Ross 1988:285)

Ilocano
(368) **maki**-sarita ‘talk with (s.o.)’
(Vanoverbergh 1955:139–140, quoted in Ross 1988:285)

Western Bukidnon Manobo
(369) **peki**-tavang ‘participate in helping’
(Elkins 1967:111, quoted in Ross 1988:285)

Cebuano
(370) **paki**-g-qaway ‘fight with (s.o.)’
(Zorc 1977, quoted in Ross 1988:284)

Hiligaynon
(371) **paki**-g-lutu-an ‘will be cooked with’
(Wolfenden 1971:132, quoted in Ross 1988:285)

Toba Batak
(372) **mahi**-solat ‘conceal oneself with (s.o.)’
(van der Tuuk 1971:133, quoted in Ross 1988:285)

A study concerning *maki- forms in WISP languages has been conducted by Liao (2011). She provides more detailed information regarding reflexes of the *maki- forms in some Philippine, northern and central Sulawesi, and northern Borneo languages.

The data provided by Liao demonstrate that the form *maki- is an actor voice and mood- and aspect-neutral form in the same manner as the actor voice forms *<um>, *maR-, and *paR- and is inflected for voice, mood, and aspect, such as *naki- ‘realis mood form of *maki-’), *paki- ‘subjunctive form of *maki-’), and *paki- -en ‘patient voice form of *maki-’. The inflectional paradigms suggested by Ross (2002) and the data quoted by Liao suggest the following paradigm for Proto-Malayo-Polynesian *maki-:

Table 48:A tentative reconstruction of *maki- in Proto-Malayo-Polynesian

	Actor voice	Patient voice	Locative voice	Thematic voice
Neutral	*maki-	*paki- -en	*paki- -an	*i-paki-
Neutral progressive	*maki-CV-	*paki-CV- -en	*paki-CV- -an	*i-paki-CV-
Realis	*naki-	*p<in>aki-	*p<in>aki- -an	*i-p<in>aki-
Realis progressive	*naki-CV-	*p<in>aki-CV-	*p<in>aki-CV- -an	*i-p<in>aki-CV-
Subjunctive	*paki-	*paki- -a	*paki- -i	*paki- -an

According to Liao (2011), the *maki- forms are polysemous. Mostly, however, their polysemies are observed across languages. Liao identifies five functions of *maki- and its inflected forms, namely, 1) ‘comitative’[125] or ‘social’ function, which encodes ‘actions performed in the company of other people’ (Liao 2011:215) [i.e., ‘(singular actor) do(es) something with (someone)’, ‘(singular actor) join(s) in doing something (with someone)’], 2) ‘request of comitative (social)’ [i.e., ‘(singular actor) ask(s) to do something with’, ‘(singular actor) ask(s) to join oneself in doing something’], 3) ‘polite request, polite imperative’ [(only *paki- subjunctive/imperative forms;

125It should be noted that the function described by her term ‘comitative’ seems to correspond to what we refer to by our term ‘comitative actor’ in the sense that a singular actor does something together with another participant encoded as comitative; however, we employ her term for the description of the functions of *maki-. This treatment is due to the fact that the *maki- ‘comitative’ function also encodes the meaning of ‘joining in an action that has started earlier’ with some roots, and that our term ‘comitative actor’ does not have this meaning.

'please do something (for me (i.e., the speaker)); ask you (i.e., the addressee) to do something')], 4) 'requestive' [i.e., 'ask someone to do something'], and 5) 'causative' ['have someone to do something']:

'Comitative'
Tagalog (Greater Central Philippine)

(373) **Nakiki**kain=siya	ng	hapunan	sa	Nanay.	
naki-CV-kain=siya	nang	hapunan	sa	Nanay	
AV.RLS.KI-PROG-eat=3.SG.NOM	GEN	supper	LOC	mother	

'He is eating supper with Mother.'
(Schachter 1976:510, quoted in Liao 2011:212, glossing modified)

'Permissive Comitative'
Ilocano (Northern Luzon)

(374) **Maki**inúm=ak	man?
maki-inum=ak	man
AV.KI.drink=1.SG.NOM	please

'May I have a drink (with you)?'
(Rubino 2000:lxxi) quoted in Liao 2011:212, glossing modified)

'Polite request, polite imperative'
Tagalog (Greater Central Philippine)

(375) **Paki**abot=mo=nga	ang	asin.
paki-abot=mo=nga	ang	asin
AV.SUBJUNCTIVE.KI-pass=2.SG.GEN=please	NOM	salt

'Please pass the salt.'
(Ramos 1985:134, quoted in Liao 2011:216, glossing modified)

'Requestive'
Hiligaynon (Greater Central Philippine)

(376) **Pakig**dala=ko	sang	basket	kay	Mr. Cruz.
pakig-dala=ko	sang	basket	kay	Mr. Cruz
AV.SUBJUNCTIVE.KI-send=1.SG.GEN	OBL	basket	LOC	Mr. Cruz

'I will request that the basket be sent to Mr. Cruz.'
(Wolfenden 1975:93, quoted in Liao 2011:218, glossing modified)

'Causative'
Maranao (Greater Central Philippine)

(377) **Paki**tabasen	o	mama	ko	wataq	so	karatas.
paki-tabas-en	o	mama	ko	wataq	so	karatas
UV.KI-cut-PV	GEN	man	OBL	child	NOM	paper

'The man will have the child cut the paper.'
(McKaughan and Macaraya 1996:6, quoted in Liao 2011:219, glossing modified)

As such, the functions of *maki- and *paki- that Liao identifies and those that Ross (1988) considers are different. In particular, while Liao considers the comitative function (i.e., syntactic function to encode plural actors in subject and comitative functions) to be primary, Ross does not seem to consider this syntactic function. Instead, he compares the semantic and syntactic functions

of the Oceanic *paRi- words solely with the semantics of the *maki- and *paki- words in WISP languages. Reflexes of both *paRi- and *maki-/paki- encode the actor plural situations; however, a closer observation between the two types of forms suggest their syntactic differences. While the majority of the reflexes of *maki- form the comitative actor two-argument constructions, the reflexes of *paRi- forms actor plural one-argument constructions when they denote actor plural situations.

Ross, however, considers *pa(k)i- to be a 'doublet' of *paRi-. Given that his use of 'doublet' refers to different forms that have the same syntactic and semantic properties, his Proto-Oceanic form *pa(k)i- should have the same syntactic property as that of *paRi- (i.e., encoding plural actors in subject function). As detailed above, because the syntactic properties of Proto-Oceanic *paRi- and Proto-Malayo-Polynesian *maki-/paki- are different, it would appear to be implausible to assume that the Proto-Oceanic *pa(k)i- as a doublet of *paRi- reflects Proto-Malayo-Polynesian *maki- or *paki-.

*5.2.5.2 PMP *maR-*

Recognizing that reflexes of the Oceanic proto-form *paRi- denote reciprocal, actor plural, undergoer plural, and repetitive situations, Pawley (1973:152–153) assumes that the cognate of *paRi- in WISP languages is *maR- or *paR-. He adduces the following Toba Batak examples with formations *mar- -i* and *mar- -an*:

Toba Batak

(378) **mar**-bunga-**i** 'to flower, of many trees, plants'
mar-harowan-**i** 'to hold feasts, with many people'
mar-gots-**i** 'to make music, of many people'

(Pawley 1973:152)

(379) **mar**-habang-**an** 'to fly, of many birds'
mar-hembang-**an** 'to be spread out, of mats'

(Pawley 1973:153)

Similarly, Bril (2005:59–60) compares Tagalog *mag-* and Malay *ber-*, which are the reflexes of PMP *maR-, to Proto-Oceanic*paRi-.

The Tagalog form *mag-* indeed derives reciprocal words from the <*um*>-class action-denoting roots. For example, the actor voice form of the root *bati* 'greet' is *b*<*um*>*ati* 'to greet someone' (realis *b*<*um*>*ati*), which includes the actor voice prefix <um> but no lexical class prefix. When this root takes the 'reciprocal' prefix *mag-* (realis *nag-*), the derived word *mag-bati* indicates the reciprocal meaning 'to greet each other':

Tagalog
Underived actor voice form

(380)	B-um-ati	ang	dalaga	sa	bisita.
	<um>bati	ang	dalaga	sa	bisita
	AV-greet	NOM	girl	LOC	visitor

'The girl greeted the visitor.'

(Shkarban and Rachkov 2007:898, glossing modified)

Reciprocal

(381)	**Nag-**bati	ang	dalaga	at	ang	bisita.
	nag-bati	ang	dalaga	at	ang	bisita
	AV-greet	NOM	girl	and	NOM	visitor

'The girl and the visitor greeted each other.'

(Shkarban and Rachkov 2007:898, glossing modified)

Some other roots that belong to this formal type are as follows:

(382)	**mag**-usap	'to talk to each other'	um-usap	'to talk to [s.o.]'
	mag-yakap	'to embrace each other'	y<um>akap	'to embrace [s.o.]'
	mag-babag	'to fight with each other'	b<um>abag	'to figh with [s.o.]'

(Shkarban and Rachkov 2007:900)

In addition, the prefix *mag-* also derives reflexive words from some action-denoting roots:

Reflexive

(383)	**mag**-ahit 'to shave oneself'	um-ahit 'to shave [s.o.]'
	mag-bigti 'to kill oneself by hanging'	b<um>igti 'to kill [s.o.] by hanging'

(Shkarban and Rachkov 2007:903)

In Malay, on the other hand, the prefix *ber-* (< *maR-) also derives reciprocal words from some action-denoting roots:

Malay

Reciprocal

(384)	**ber**-tengkar(-an)	'to quarrel'

(Ogloblin and Nedjalkov 2007:1447)

(385)	**ber**-bicara	'to converse'	mem-bicara-kan 'to discuss [i.e., to speak about]'
	ber-tengkar	'to quarrel'	mem-per-tengkar-kan 'to quarrel because of'

(Ogloblin and Nedjalkov 2007:1449)

Furthermore, the prefix *ber-* also derives reflexive verbs from the word for 'shaving', in the same manner as the Tagalog word pair, *um-ahit* 'to shave [s.o.]' and *mag-ahit* 'to shave oneself':

Reflexive

(41)	**ber-**cukur	'to shave oneself'	men-cukur 'to shave [s.o.]'

(Ogloblin and Nedjalkov 2007:1448)

Because reflexes of both Proto-Oceanic *paRi- and Proto-Malayo-Polynesian *maR- derive reciprocal and reflexive words, one may consider the claim that Proto-Oceanic *paRi- reflects Proto-Malayo-Polynesian *maR- and its inflected forms with *paR- to be valid. However, we argue against this claim because of the unexpected presence of the final *i in the Proto-Oceanic form *paRi-. The Proto-Malayo-Polynesian prefix *maR- is expected to be reflected as Proto-Oceanic *maR-. The Oceanic reflex would more accurately be *paR- without the actor voice segment *m- if we consider the non-productivity of voice affixes in Oceanic languages. However, either form cannot explain the reason for the presence of the final *i in *paRi-.

5.2.5.3 Summary and discussion

The previous subsections examined the two Proto-Malayo-Polynesian forms, namely, *maki- and *maR-, that have been suggested as cognates of Proto-Oceanic *paRi-. Because no suggested Proto-Malayo-Polynesian forms exhibit the expected phonological correspondences with the Proto-Oceanic form *paRi-, we must also consider their functions in order to identify the Proto-Malayo-Polynesian form of the Proto-Oceanic form *paRi-.

Table 49: Possible proto-forms of Proto-Oceanic *paRi-

PMP form	Expected POc form	Functions in WISP languages that are also found in Oceanic languages	Problems
*maki-	**paki-	N/A	– No attestation of 'permissive'-related functions, such as 'ask to join in' and 'ask s.o. to do s.t.', in Oceanic languages.
*maR-	**paR-	Reciprocal, Reflexive	– The presence of *i in the Proto-Oceanic form *paRi- is not explainable.
*maR-si-	**pa-si-	Actor plural, Reciprocal, Reflexive, Comitative actor; Undergoer plural	– Some functions of Proto-Oceanic *paRi- are not attested in reflexes of *maR-si- in WISP languages. – Sound correspondences are not sufficiently accounted for (expected Proto-Oceanic /s/ instead of /R/)

Of the three suggested proto-forms, we can first eliminate the claim that the Proto-Oceanic form *paRi- reflects Proto-Malayo-Polynesian *maR- and *paR- because of the lack of an account for the final phoneme *i in *paRi-. If we consider what the source of this phoneme is, we can assume that a Proto-Malayo-Polynesian form that has a phoneme *i somewhere after *maR- or *paR- is a more plausible source.

As far as our available data suggest, there are only two derivational forms, namely, *maki- and *maR-si-, that end with the vowel *i and that are widely attested in WISP and therefore are plausible candidates for being cognates of *paRi-. Between these forms, the form *maR-si- and its inflected forms, such as the *paR-si- undergoer voice forms, are more likely to be one set of the sources of the Proto-Oceanic prefix *paRi- due to the fact that these formations show much better functional correspondences to the Oceanic forms in question.

Finally, it should be noted that we may be able to suggest as a possibility that Proto-Oceanic *paRi- is an innovation in the Proto-Oceanic language and does not sufficiently correspond to any formatives attested in WISP languages. The formal and functional correspondences between *paRi- and *maR-si- (and *paR-si-) are suggestive; however, given the formal and functional differences, any of the suggested cognate relations might be considered less convincingly established.

6 Concluding remarks

In the current study, we explored the forms and functions of reflexes of the prefix *si- (in many languages expanded to *maR-si-) in the Western Indonesian, Sulawesi, and Philippine languages (referred to as 'WISP languages' throughout the study). In particular, we investigated the *maR-si- forms with action-denoting roots in actor voice constructions; however, we also examined possible cognates of *si- forms of roots and words other than action-denoting roots, such as 1) quality-denoting roots, 2) object-denoting roots, 3) kinship terms, and 4) location words. Furthermore, we assessed possible cognate forms of *si- forms in other Austronesian languages and revealed some formal and functional commonalities between some of these forms and *si- forms in WISP languages.

Actor voice *si- forms exhibit various formal variations. For example, some reflexes of proto-phonemes in *maR-si- are not identical to those found in lexical items. Some *si- forms do not contain the initial consonant /m/ or the initial actor voice prefix *maR-, as in Ilocano *ag-si-* and Makassarese *si-*. By contrast, other forms include some additional phonemes, such as Aklanon *mag-si(g)-*.

Furthermore, our data suggest that there are *si- forms with lexical class prefixes that denote morphological and lexical classes of roots. For example, Tagalog *mag-si-pag-* is used for *pag-* class roots, and *mag-si-paN-* is employed for *paN-* class roots. The functions of these forms with lexical class prefixes are identical to the Tagalog *(maR-)si- form *mag-si-* without a lexical class prefix.

In addition, there are a few forms that may change the meaning indicated by *si-. The event plural suffixes *-an and *-ay also appear with *si- forms, as in Pangasinan successive distributive plural *mag-si- -an*, which has a different meaning from the distributive form *mag-si-*. Similarly, reduplicative forms may also add a different meaning. The Muna form *si-CVCV- -ha* with CVCV-reduplication denotes actor plural situations, while the form *si- -ha* without the reduplication indicates actor dual situations.

Forms with *si- have various meanings. In actor voice constructions, our data show that si-forms denote the following situations and constructions: 1) actor plural (e.g., *they hit a ball*), 2) distributive actor plural (e.g., *each of them hit a ball*), 3) reciprocal (e.g., *they hit each other* with the root *hit*), 4) reflexive (e.g., *he hits himself*), 5) successive distributive actor plural (e.g., *they hit a ball in turn*), 6) actor dual (e.g., *the two of them hit a ball*), and 7) comitative actor (e.g., *he hits a ball with her*).

In undergoer voice constructions, it is observed that *si- forms indicate 1) undergoer plural constructions (i.e., the plurality of undergoers in subject function; e.g., *the balls are hit by him*), 2) actor plural constructions (i.e., the plurality of actors in subject function; e.g., *the ball is hit by them*), and 3) comitative undergoer constructions (i.e., the plurality of undergoers in both subject and comitative arguments; e.g., *the ball, along with another ball, is hit by him*).

For quality-denoting roots, possible *si- forms denote 1) competitive situations with regard to the quality indicated by the root (e.g., (*they) compete in speed* with the root *fast*), 2) situations in which plural participants are of different degrees of quality (e.g., *they are of different height* with the root *tall*), 3) situations in which plural or dual participants are of the same degree of quality (e.g., *they/the two of them are of the same height* with the root *tall*), and 4) singular associate entities that are of the same degree of quality as someone else is (e.g., *he is as tall as she* is with the root *tall*). Of these functions, the first three are marked with reflexes of *maR-, as in Malay *ber-si-* 'FUNCTION (1)', Toba Batak *mar-si- -i* 'FUNCTION (2)', and Tagalog *mag-(ka-)sing-* 'FUNCTION (3)'. In addition, some forms exhibit reflexes of the prefix *ka-, as in Tagalog *mag-(ka-)sing-* 'FUNCTION (3)' and *(ka-)sing-* 'FUNCTION (4)'.

With object-denoting roots, possible *si- forms indicate 1) situations in which plural or dual participants share an object (e.g., *they/the two of them have the same teacher* with the root *teacher*), 2) singular associate entities that share an object with someone else (e.g., *he is her classmate*, i.e., *he is in the same class as her,* with the root *class*), 3) singular associate entities that have the same quality of an object as someone else does (e.g., *he is a teacher as she is* with the root *teacher*), and 4) plurality or duality of the object denoted by the root (e.g., *eyes* derived from the root *eye*). In a similar manner to *si- forms with quality-denoting roots, some forms include reflexes of *ka-, as in Cebuano *i-sig-ka-* 'FUNCTION (3)', Pamona *ka-si-* 'FUNCTIONS (2) and (3)', and Cebuano *ma-sig-ka-* 'FUNCTION (4)'. In addition, the prefix *i-* is found in some associative forms, as in Cebuano *i-sig-ka-* 'FUNCTION (3)' and Aklanon *i-sig-ka-* 'FUNCTIONS (2) and (3)'. This prefix is identical in form to the thematic voice marker *i-* in these languages; however, the relationship between the two forms warrants further study.

With kinship terms, possible *si- forms signify 1) converse kinship pairs (e.g., *a mother and a child* with the root *mother*) and 2) converse kinship groups (e.g., *a mother and children* with the root *mother*). Some of these forms also exhibit the presence of reflexes of *ka-, as in Sangir *seng-ka-*, in the same manner as *si- forms with quality- and object-denoting roots. This fact alone may suggest that *si- forms with object-denoting roots and those with kinship terms could be different realizations of the same single ancestral form denoting the plurality or duality of the object indicated by the root. However, this assumption requires further investigation because these kinship categories are also marked with other forms, such as *maR-tali- (or *(mV(R/n)-)tVl(V)-, (e.g., Hanunoo *mag-tal-*, Uma *n-tali-*, Pendau *n-toli-*, and Matigsalug Manobo *tala-*)) and *maR- (e.g., Tagalog *mag-*).

With words indicating locative relations, possible *si- forms denote the symmetrical relationship between two localities, such as *face each other* with the root meaning 'in front of'. The syntactic behavior of this function, namely, whether the localities are encoded in subject function or subject and comitative arguments, merits further analysis.

Possible cognates of WISP *si- forms are also found in some Formosan languages. In Kavalan, the prefix *sim-* attaches to action-denoting roots and denotes reciprocal and actor plural situations. This language also has the undergoer voice *si- form *sim- -an*, which signifies actor plural situations. In addition to the Kavalan forms, it is also observed that the Seediq forms *ms-* and *mt-* denote converse kinship relationships, in a similar manner to *si- forms and *maR-tali- forms in WISP languages.

On the other hand, our available data on Oceanic languages do not suggest the presence of a form that contains the segment *si and indicates reciprocal, actor plural or reflexive situations. However, it is widely observed that some reflexes of the Proto-Oceanic prefix *paRi- derive words denoting such situations from action-denoting roots. In addition, other reflexes of *paRi- also appear with quality- and object-denoting roots, kinship terms, and location words and indicate situations and functions that are almost identical to those denoted by *si- forms in WISP languages. Based on the comparison between Oceanic *paRi- forms and WISP *si- forms for action-denoting roots, the current study suggests the possibility that some *paRi- forms are cognates of *maR-si- and *paR-si- forms in WISP languages.

The exploration of *si- forms in actor voice constructions with action-denoting roots could decrease the original functions of the prefix *si- to three, only two of which play a role in individual languages: 1) actor plural, 2) reflexive, and 3) reciprocal. In other words, our investigation suggests that the prefix *si- 1) only possesses the semantic feature 'actor plural', 2) only possesses the semantic feature 'actor-cum-undergoer role' (i.e., reflexive), or 3) possesses both of these semantic features (i.e., reciprocal). These original functions and their semantic components suggest the following development paths: 1) actor plural → reciprocal → reflexive, 2) reflexive → reciprocal

→ actor plural, and 3) (a) reciprocal → actor plural and (b) reciprocal → reflexive.

Further identification of the original function requires two types of exploration. One type is further research on the relationship between *si- forms with action-denoting roots and those with other types of roots and words. The other type involves research on the relationship between the *si- forms and other derivational forms, including 1) other expressions of reciprocal, actor plural, and reflexive constructions and 2) combinations of *si- forms and other derivational affixes, such as the Proto-Malayo-Polynesian causative prefix *pa-, as in Cotabato Manobo *eg-pe-se-* and *eg-se-pe--ay*[126] (cf. *pa- > *pe-*, YK). In addition to these topics, further comparative research on voice- and case-marking systems in WISP languages would also contribute to a better understanding of the history of the prefix *si-.

126Data from Kerr (1988:102).

Bibliography

Adelaar, K. Alexander. 1984. Some Proto-Malayic affixes. *Bijdragen tot de Taal-, Land- en Volkenkunde* 140:402–421.

———. 1992a. The Relevance of Salako for Proto-Malayic and for Old Malay Epigraphy. *Bijdragen tot de Taal-, Land- en Volkenkunde* 148(3/4):381–408.

———. 1992b. *Proto-Malayic: the reconstruction of its morphology and parts of its lexicon and morphology*. Canberra: Pacific Linguistics.

———. 2004. The Coming and Going of 'Lexical Prefixes' in Siraya. *Language and Linguistics* 5(2): 333–361.

———. 2005a. Much ado about *di-*. *Bijdragen tot de Taal-, Land- en Volkenkunde* 161(1):127–142.

———. 2005b. The Austronesian languages of Asia and Madagascar: a historical perspective. In *The Austronesian languages of Asia and Madagascar*, ed. by K. Alexander Adelaar and Nikolaus. P. Himmelmann, 1–42. London and New York: Routledge.

———. 2005c. *Salako or Badameà: sketch grammar, texts and lexicon of a Kanayatn dialect in West Borneo*. Frankfurter Forschungen zu Südostasien 2. Wiesbaden: Harrassowitz.

Adriani, Nicolaus. 1893. *Sangireesche spraakkunst* [Sangirese grammar]. Leiden: A. H. Adriani.

———. 1931. *Spraakkunst der Bare'e-taal* [Grammar of the Bare'e language]. Bandoeng: Nix.

———. and M. L. Adriani-Gunning. 1908. *Hoofdstukken uit de spraakkunst van het Tontemboansch naar aanteekeningen en mededeelingen van den heer J. Alb. T. Schwarz en anderen* [Chapters from the grammar of Tontemboan according to the notes and information of Mr. J. Alb. T. Schwarz and others]. 's Gravenhage: Martinus Nijhoff.

Alpatov, Vladimir M., Anna Bugaeva and Vladimir P. Nedjalkov. 2007. Reciprocals and sociatives in Ainu. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1751–1822. Amsterdam: John Benjamins.

Aramiou, Sylvain et al. 2001. *Dictionnaire ajië–français* [Ajie-French dictionary]. Nouméa: Fédération de l'Enseignement Libre Protestant.

Barr, Donald. F. 1988a. Da'a Verbal Affixes and Clitics. In *Papers in Western Austronesian Linguistics No. 4*, ed. by Hein Steinhauer, 11–49. Canberra: Pacific Linguistics.

———. 1988b. Focus and Mood in Da'a Discourse. In *Papers in Western Austronesian Linguistics No. 4*, ed. by Hein Steinhauer, 77–129. Canberra: Pacific Linguistics.

Barr, Donald. F. and Barr, Sharon. G. 1988. Phonology of Da'a, Central Sulawesi. In *Papers in Western Austronesian Linguistics No. 4*, ed. by Hein Steinhauer. 131–151. Canberra: Pacific Linguistics.

Benton, Richard A. 1971a. *Pangasinan dictionary.* PALI Language Texts: Philippines. Honolulu: University of Hawaii Press.

———. 1971b. *Pangasinan reference grammar.* PALI Language Texts: Philippines. Honolulu: University of Hawaii Press.

Bloomfield, Leonard. 1917. *Tagalog Texts with Grammatical Analysis.* 3 vols. Urbana: The University of Illinois.

Blust, Robert. 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30(2):73–129.

———. 2003a. Three notes on Early Austronesian morphology. *Oceanic Linguistics* 42:438–478.

———. 2003b. *A short morphology, phonology and vocabulary of Kiput, Sarawak*. Canberra: Pacific Linguistics.

———. 2004. Austronesian nasal substitution: a survey. *Oceanic Linguistics* 43(1):73–148.

———. 2007. The linguistic position of Sama-Bajaw. *Studies in Philippine Languages and Cultures* 15:73–114.

———. 2010. The Greater North Borneo hypothesis. *Oceanic Linguistics* 49:44–118.

———. 2013. *The Austronesian languages*. Revised edition. Canberra: Asia-Pacific Linguistics.

———. 2015. The Case-Markers of Proto-Austronesian. *Oceanic Linguistics* 54(2):436–491.

——— and Stephen Trussel. ongoing. *The Austronesian Comparative Dictionary*. http://www.trussel2.com/acd/.

Bril, Isabelle. 2005. Semantic and functional diversification of reciprocal and middle prefixes in New Caledonian and other Austronesian languages. *Linguistic Typology* 9:25–76.

———. 2007. Polysemy of the reciprocal marker in Nêlêmwa. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1479–1509. Amsterdam: John Benjamins.

Cabredo Hofherr, Patricia and Brenda Laca eds. 2012. *Verbal Plurality and Distributivity*. Berlin and Boston: De Gruyter.

Capus, P. A. 1898. Grammaire de Shisumbwa ou Lusumbwa [A grammar of Shisumbwa or Lusumbwa]. *Zeitschrift für afrikanische und oceanische Sprachen*. IV. Jahrgang, 1–123.

Conant, C. E. 1911. The RGH law in Philippine languages. *Journal of the American Oriental Society* 31:74–85.

Conklin, Harold C. 1953. *Hanunóo-English vocabulary*. Berkeley: University of California Press.

Cook, Kenneth. 1988. *A cognitive analysis of grammatical relations, case, and transitivity in Samoan*. Ph.D. dissertation. University of California, San Diego.

Coolsma S. 1873. *Handleiding bij de Beoefening der Soendaneesche Taal* [Manual for the Practice of the Sundanese Language]. Batavia: H. M. van Dorp & Co.

Davis, Karen. 1997. *A grammar of the Hoava language, western Solomons*. Ph.D. dissertation. University of Auckland.

De Guzman, Videa P. 2000. Some remarks on the grammatical functions of the nonabsolutive agent in Tagalog. In *Grammatical analysis: morphology, syntax, and semantics: studies in honor of Stanley Starosta*, ed. by Videa P. De Guzman and Byron Bender,. 224–240. Canberra: Pacific Linguistics.

Dempwolff, Otto. 1920. *Die Lautentsprechungen der indonesischen Lippenlaute in einigen anderen austronesischen Südseesprachen* [The sound correspondences of the Indonesian labials in some other Austronesian languages]. *Zeitschrift für Eingeborenen-Sprachen*, Supplement 2. Berlin: Reimer.

———. 1937. *Vergleichende Lautlehre des austronesischen Wortschatzes* [Comparative phonology of the Austronesian vocabulary]. *Vol. 2 Deduktive Anwendung des Urindonesischen auf austronesische Einzelsprachen* [Deductive application of Proto-Indonesian to individual languages of the Austronesian family]. *Zeitschrift für Eingeborenen-Sprachen*. Supplement 17. Berlin: Reimer.

Dixon, R. M. W. 1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

Donohue, Mark. 2005. The pretenders to the Muna-Buton group. In *Papers in Austronesian subgrouping and dialectology*, ed. by John Bowden and Nikolaus Himmelmann, 21–36. Canberra: Pacific Linguistics.

DuBois, Carl D. 1976. *Sarangani Manobo: An Introductory Guide*. Manila: Linguistic Society of the Philippines.

Dubois, M. J. n.d. *Dictionnaire nengone–français* [Nengone-French dictionary]. Unpublished manuscript.

Dunnebier, W. 1929. Spraakkunst van het Bolaang Mongondowsch [Grammar of Bolaang Mongondow]. Volume 1. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch Indië* 85: 297–468.

Echols, John M. and Hassan Shadily. 1994. *Kamus Indonesia-Inggris* [Indonesian-English dictionary]. 3rd edition. Jakarta: Gramedia Pustaka Utama.

Elkins, Richard E. 1967. *Major grammatical patterns of Western Bukidnon Manobo*. Ph.D. dissertation. University of Hawaii.

——— and Gail R Hendrickson. 1985. *A sampling of Philippine kinship patterns*. Metro Manila: SIL.

English, Leo James. 1986. *Tagalog-English dictionary.* Manila: Congregation of the Most Holy Redeemer.
Esser, S. J. 1964. *De Uma-taal (West Midden-Celebes): spraakkunstige schets en teksten* [The Uma language (West Central-Sulawesi): grammatical sketch and texts]. s'-Gravenhage: Martinus Nijhoff.
Evans, Bethwyn. 2003. *A study of valency-changing devices in Proto Oceanic*. Canberra: Pacific Linguistics.
Evans, Nicholas. 2003 *An interesting couple: The semantic development of dyad morphemes*. Arbeitspapier 47. Cologne: Institut für Sprachwissenschaft, Universität zu Köln.
———, Alice Gaby, Stephen C. Levinson, and Asifa Majid eds. 2011. *Reciprocals and Semantic Typology.* Amsterdam: John Benjamins.
Ferreirinho, Naomi. 1993. *Selected Topics in the Grammar of Limos Kalinga, The Philippines.* Canberra: Pacific Linguistics.
Forster, Jannette. 1983. Use of dialogue in Dibabawon narrative discourse. *Philippine Journal of Linguistics* 14(1):45–60.
Fortescue, Michael. 1984. *West Greenlandic*. London: Croom Helm.
Frajzyngier, Zygmunt and Traci S. Curl eds. 2000a. *Reciprocals: Forms and Functions.* Amsterdam: John Benjamins.
———. eds. 2000b. *Reflexives: Forms and Functions.* Amsterdam: John Benjamins.
Gaby, Alice. 2008. Distinguishing reciprocals from reflexives in Kuuk Thaayorre. In *Reciprocals and Reflexives*, ed. by Ekkehard König and Volker Gast, 259–288. Berlin and Boston: De Gruyter Mouton.
Geniušienė, Emma. 1987. *The Typology of Reflexives.* Berlin: Mouton de Gruyter.
Gil, David. 2013. Riau Indonesian: a language without nouns and verbs. In *Flexible Word Classes: Typological studies of underspecified parts of speech*, ed. by Jan Rijkhoff and Eva van Lier. Oxford: Oxford University Press.
Guentchéva, Zlatka and Nicole Rivière. 2007. Reciprocal and reflexive constructions in French. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 561–607. Amsterdam: John Benjamins.
Healey, P. M. 1960. *An Agta grammar.* Manila: Bureau of Printing.
Heine, Bernd and Hiroyuki Miyashita. 2008. The intersection between reflexives and reciprocals: A grammaticalization perspective. In *Reciprocals and Reflexives*, ed. by Ekkehard König and Volker Gast, 169–224. Berlin and Boston: De Gruyter Mouton.
Hill, Deborah. 1992. *Longgu Grammar.* Ph.D. dissertation. Australian National University.
Himmelmann, Nikolaus, P. 1996. Person marking and grammatical relations in Sulawesi. In *Papers in Austronesian Linguistics No. 3, ed. by Hein Steinhauer,* 115–136. Canberra: Pacific Linguistics.
———. 2001. *Sourcebook on Tomini-Tolitoli languages. General information and wordlists*. Canberra: Pacific Linguistics.
———. 2002. Voice in western Austronesian: an update. In *The history and typology of western Austronesian voice systems,* ed. by Fay Wouk and Malcolm D. Ross, 7–16. Canberra: Pacific Linguistics.
———. 2008. Lexical categories and voice in Tagalog. In *Voice and grammatical relations in Austronesian languages*, ed. by Simon Musgrave and Peter Austin, 247–293. Stanford: CSLI Publications.
———. 2017. Word Classes. In *Oxford Bibliographies in Linguistics*, http://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0159.xml.
Holmer, Arthur J. 1996. *A Parametric Grammar of Seediq*. Lund: Lund University Press.
Hurst, Peter and Rachel Nordlinger. 2011. Reciprocal constructions in English: Each other and

beyond. In *Reciprocals and Semantic Typology,* ed. by Nicholas Evans, Alice Gaby, Stephen C. Levinson, and Asifa Majid, 75–90. Amsterdam: John Benjamins.
Jukes, Anthony. 2006. *Makassarese (basa Mangkasara'): a description of an Austronesian language of South Sulawesi*. Ph.D. sissertation. Melbourne: Department of Linguistics and Applied Linguistics, University of Melbourne.
Kemmer, Suzanne E. 1993. *The Middle Voice*. Amsterdam: John Benjamins.
———. 1997. Collective and distributive marking, or: where unity meets multiplicity. In *The Twenty-third LACUS Forum 1996*, ed. by A. K. Melby, 231–249. Chapel Hill, NC: Linguistic Association of Canada and the United States.
Kern, H. 1899. Bijdragen tot de Spraakkunst van het Oudjavaansch [Contribution to the Old Javanese grammar]. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië*, Deel 50, 3de/4de Afl. [6e Volgreeks, 6e Deel]: 401–404.
———. 1902. Bijdragen tot de Spraakkunst van het Oudjavaansch [Contribution to the Old Javanese grammar]. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië*, Deel 54, 1ste/2de Afl., [6e Volgreeks, 10e Deel], 173–194.
Kerr, Harland B. 1988. Cotabato Manobo grammar. *Studies in Philippine Linguistics* 7(1):1–123.
Kitada, Yuko. 2015. The Etymology of the Sociative-Progressive Circumfix in Suwawa (Gorontalo-Mongondowic). In *Language Change in Austronesian Languages*, ed. by Malcolm D. Ross and I Wayan Arka, 79–92. Canberra: Asia-Pacific Linguistics.
Koch, Harold. 2003. Morphological reconstruction as an etymological method. In *Historical Linguistics 2001: Selected papers from the 15th International Conference on Historical Linguistics, 13-17 August 2001*, ed. by Barry J. Blake and Kate Burridge, 271–291. Amsterdam: John Benjamins.
———. 2009. On reconstructing pronominal proto-paradigms: methodological considerations from the Pama-Nyungan language family of Australia. In *Discovering history through language: papers in honour of Malcolm Ross*, ed. by Bethwyn Evans,. 317–344. Canberra: Pacific Linguistics.
———. 2013. The etymology of a paradigm: the Pama-Nyungan 3SgF reconsidered. In *Lexical and Structural Etymology: Beyond Word Histories*, ed. by Robert Mailhammer, 53–82. Berlin and Boston: Walter de Gruyter.
König, Ekkehard and Volker Gast. 2008. Reciprocity and reflexivity – description, typology and theory. In *Reciprocals and Reflexives*, ed. by Ekkehard König and Volker Gast, 1–32. Berlin and Boston: De Gruyter Mouton.
———. eds. 2008. *Reciprocals and Reflexives. Theoretical and Typological Explorations*. Berlin and Boston: De Gruyter Mouton.
Kroeger, Paul. 2005. Kimaragang. In *The Austronesian languages of Asia and Madagascar*, ed. by K. Alexander Adelaar and Nikolaus. P. Himmelmann, 397–428. London and New York: Routledge.
Lehmann, Christian. 2008. Roots, stems and word classes. *Studies in Language* 32(3):546–567.
Li, Paul Jen-kuei. 1996. *The Formosan Tribes and Languages in I-Lan*. I-Lan: I-Lan County Government. (in Chinese).
——— and Shigeru Tsuchida. 2006. *Kavalan dictionary.* Taipei: Institute of Linguistics, Academia Sinica.
Liao, Hsiu-chuan. 2011. On the Development of Comitative Verbs in Philippine Languages. *Language and Linguistics* 12(1):205–237.
Lichtenberk, Frantisek. 1985. Multiple uses of reciprocal constructions. *Australian Journal of Linguistics* 5:19–41.
———. 2000. Reciprocals without reflexives. In *Reciprocals: Forms and Functions,* ed. by Frajzyngier Zygmunt and Traci S. Curl, 30–62. Amsterdam: John Benjamins.
———. 2007. Reciprocals and related meanings in To'aba'ita. In *Typology of Reciprocal*

Categories and Constructions, ed. by Vladimir P. Nedjalkov, 1547–1571. Amsterdam: John Benjamins.
———. 2008. *A grammar of Toqabaqita*. Berlin: Mouton de Gruyter.
Lobel, Jason William. 2013. *Philippine and North Bornean languages: Issues in description, subgrouping, and reconstruction.* Ph.D. dissertation. University of Hawai'i.
Lopez, Cecilio 1977a [1937] Preliminary Study of the Affixes in Tagalog. In *Selected writings of Cecilio Lopez in Philippine linguistics,* ed. by Ernesto Constantino, 28–104. Quezon City: Archives of Philippine Languages and Dialects, University of the Philippines.
———.1977b [1949]. Studies of Sugbuanon Affixes. In *Selected writings of Cecilio Lopez in Philippine linguistics, ed. by Ernesto Constantino,* 123–185. Quezon City: Archives of Philippine Languages and Dialects, University of the Philippines.
Majid, Asifa, Nicholas Evans, Alice Gaby and Stephen C. Levinson. 2011. The semantics of reciprocal constructions across languages: An extensional approach. In *Reciprocals and Semantic Typology,* ed. by Nicholas Evans, Alice Gaby, Stephen C. Levinson, and Asifa Majid, 29–60. Amsterdam: John Benjamins.
Martens, M. Notes on Uma Verbs. In *Papers in Western Austronesian Linguistics No. 4,* ed. by Hein Steinhauer, 167–237. Canberra: Pacific Linguistics.
Maslova, Elena. 2008. Reflexive encoding of reciprocity: Cross-linguistic and language-internal variation. In *Reciprocals and Reflexives*, ed. by Ekkehard König and Volker Gast, 225–258. Berlin and Boston: De Gruyter Mouton.
McKaughan, Howard P. and Batua A. Macaraya. 1996. *A Maranao Dictionary.* Manila: De La Salle University Press.
Miller, D. Gary 1993. *Complex Verb Formation.* Amsterdam: John Benjamins.
Miller, Mark T. 2007. *A grammar of West Coast Bajau*. Ph.D. dissertation. The University of Texas at Arlington.
Milner, G. B. 1966. *Samoan Dictionary*. Oxford: Oxford University Press.
———. 1972. *Fijian Grammar.* Fiji: Government Press Suva.
Mead, David. 2002. Proto Celebic focus revisited. In *The history and typology of western Austronesian voice systems,* ed. by Fay Wouk and Malcolm D. Ross, 143–177. Canberra: Pacific Linguistics.
Mosel, Ulrike. 1984. *Tolai Syntax and its Historical Development*. Canberra: Pacific Linguistics.
———. and Even Hovdhaugen. 1992. *Samoan Reference Grammar.* Oslo: Scandinavian University Press.
Moyse-Faurie, Claire. 1983. *Le drehu, langue de Lifou (Iles Loyautés)* [Drehu, a language of Lifou (the Loyalty Islands)]. (Langues et Cultures du Pacifique, 3.) Paris: SELAF.
———. 2007. Reciprocal, sociative, reflexive, and iterative constructions in East Futunan. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1511–1543. Amsterdam: John Benjamins.
———. 2008. Constructions expressing middle, reflexive and reciprocal situations in some Oceanic languages. In *Reciprocals and Reflexives*, ed. by Ekkehard König and Volker Gast, 105–168. Berlin and Boston: De Gruyter Mouton.
Nababan, P. W. J. 1981. *A grammar of Toba-Batak.* Canberra: Pacific Linguistics.
Nedjalkov, Igor V. and Vladimir P. Nedjalkov. 2007a. Reciprocals, sociatives, comitatives, and assistives in Yakut. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1095–1161. Amsterdam: John Benjamins.
Nedjalkov Igor V. and Vladimir P. Nedjalkov 2007b. Reciprocal and sociative constructions in Evenki (with an appendix on Manchu). In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1593–1642. Amsterdam: John Benjamins.
Nedjalkov, Vladimir P. 2007a. Overview of the research. Definitions of terms, framework, and related issues. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P.

Nedjalkov, 3–114. Amsterdam: John Benjamins.
———. 2007b. Polysemy of reciprocal markers. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 231–333. Amsterdam: John Benjamins.
———. 2007c. Reciprocal derivation involving non-verbals. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 353–377. Amsterdam: John Benjamins.
———. 2007d. Reciprocals, assistives, and plural in Kirghiz. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1231–1279. Amsterdam: John Benjamins.
Nedjalkov, Vladimir P. ed. 2007. *Typology of Reciprocal Categories and Constructions*. Amsterdam: John Benjamins.
Nedjalkov, Vladimir P. and Igor V. Nedjalkov. 2007. Reciprocal, sociative and competitive constructions in Karachay-Balkar. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 969–1019. Amsterdam: John Benjamins.
Nedjalkov, Vladimir P., Elena K. Skribnik, Evgenij A. Kuzmenkov and Natalija S. Yakhontova. 2007. Reciprocal, sociative, and assistive constructions in Buryat and Khalkha-Mongol. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1281–1348. Amsterdam: John Benjamins
Nichols, Johanna. 1982. Ingush transitivization and detransitivization. *Berkeley Linguistics Society* 8: 445–462.
———. 1986. Head-marking and dependent-marking grammar. *Language* 66: 56–119.
———. 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
———. 2007. What, if anything, is typology?. *Linguistic Typology*. 11(1): 231–238.
———. 2016. Morphology in Typology. In *The Cambridge Handbook of Morphology*, ed. by Andrew Hippisley and Gregory Stump, 710–42. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press.
Ogloblin, Aleksandr K. and Vladimir P. Nedjalkov. 2007. Reciprocal constructions in Indonesian. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1437–1476. Amsterdam: John Benjamins.
Osumi, Midori. 1995. *Tinrin Grammar (New Caledonia)*. (Oceanic Linguistics Special Publication, 25.) Honolulu: University of Hawaii Press.
Ozanne-Rivierre, Françoise. 1984. *Dictionnaire iaai-français (Ouvéa, Nouvelle-Calédonie)* [Iaai-French dictionary (Ouvéa, New Caledonia)] (Langues et Cultures du Pacifique, 6.) Paris: SELAF.
Palmer, F. R. 1995. Reviewed Work: *The Middle Voice* by Suzanne Kemmer. *Journal of Linguistics,* 31(2): 478–480.
Pawley, Andrew. 1973. Some problems in Proto-Oceanic grammar. *Oceanic Linguistics* 12:103–188.
Polinsky, Maria and Robert Kluender. 2007. Linguistic typology and theory construction: Common challenges ahead. *Linguistic Typology* 11:273–283.
Post, Ursula and Mary Jane Gardner. 1992. Binukid dictionary. *Studies in Philippine Linguistics* 9(2). Manila: Linguistic Society of the Philippines and Summer Institute of Linguistics
Quick, Philip A. 2003. *A Grammar of the Pendau Language*. Ph.D. dissertation. The Australian National University.
Ramos, Teresita V. 1985. *Conversational Tagalog: A Functional-Situational Approach*. Honolulu: University of Hawaii Press.
Rapold, Christian J. 2011. Semantics of Khoekhoe reciprocal constructions. In *Reciprocals and Semantic Typology,* ed. by Nicholas Evans, Alice Gaby, Stephen C. Levinson, and Asifa Majid, 61–74. Amsterdam: John Benjamins.
Reid, Lawrence A. 1976. *Bontok-English Dictionary*. Canberra: Pacific Linguistics.
———. 1992. On the development of the aspect system in some Philippine languages. *Oceanic Linguistics* 31:65–91.

———. 2009. Challenging the status-quo: Drift, direct inheritance and reconstruction. In *Linguistic substrata in the Tibeto-Burman area*, ed. by Yasuhiko Nagano, 1–10, *Senri Ethnological Studies* 75. Osaka: National Museum of Ethnology.
Reid, Lawrence A. and Hsiu-chuan Liao. 2004. A brief syntactic typology of Philippine languages. *Language and Linguistics* 5(2):433–490.
Ringe, Don and Joseph F. Eska. 2013. *Historical Linguistics: Toward a Twenty-First Century Reintegration*. Cambridge: Cambridge University Press.
Rivierre, Jean-Claude. 1980. *La langue de Touho: Phonologie et grammaire du cèmuhî* [The language of Touho: Phonology and grammar of Cèmuhî]. (Langues et Civilisations à Tradition Orale, 30.) Paris: SELAF.
Rivierre, Jean-Claude, Sabine Ehrhart, and Raymond Diéla. 2006. *Le Bwatoo et les dialectes de la région de Koné (Nouvelle-Calédonie)* [The Bwatoo and the dialects of the Koné region (New Caledonia)]. Leuven: Peeters.
Robins, R. H. 1959. Nominal and verbal derivation in Sundanese. *Lingua* 8: 337–369.
Robinson, Laura C. 2008. D*upaningan Agta: grammar, vocabulary, and texts*. Ph.D. dissertation. University of Hawai'i.
Roolvink, R. 1965. The passive-active per-/ber-//per-/memper- correspondence in Malay. *Lingua* 15:310–337.
Ross, Malcolm D. 1988. *Proto Oceanic and the Austronesian languages of western Melanesia*. Canberra: Pacific Linguistics.
———. 1995. Reconstructing Proto-Austronesian verbal morphology: evidence from Taiwan. In *Austronesian Studies Relating to Taiwan*, ed. by Paul Jen-kuei Li, Tsang Cheng-hwa, Ying-kuei Huang, Dah-an Ho, and Chiu-yu Tseng, 727–792. Symposium Series of the Institute of History and Philology, No. 3. Taipei: Academia Sinica.
———. 2002. The history and transitivity of Western Austronesian voice and voice-marking. In *The History and Typology of Western Austronesian Voice Systems,* ed. by Fay Wouk and Malcolm D. Ross, 17–62. Canberra: Pacific Linguistics.
———. 2004. The Morphosyntactic Typology of Oceanic Languages. *Language and Linguistics* 5(2):491–541.
———. 2009. Proto Austronesian verbal morphology: a reappraisal. In *Austronesian historical linguistics and culture history: a festschrift for Robert Blust*, ed. by Alexander Adelaar and Andrew Pawley, 295–326. Canberra: Pacific Linguistics.
———. 2015. Reconstructing Proto-Austronesian verb classes. *Language and Linguistics* 16(3):279–345.
Rubino, Carl R. Galvez. 1997. *A reference grammar of Ilocano*. Ph.D. dissertation. University of California, Santa Barbara.
———. 2000. *Ilocano Dictionary and Grammar.* Honolulu: University of Hawai'i Press.
Sam, Drilë Léonard. 1995. *Dictionnaire drehu–français (Lifou, Nouvelle-Calédonie).* [Drehu-French dictionary (Lifou, New Caledonia)]. Nouméa: Centre territorial de recherche et de documentation pédagogique de Nouvelle-Calédonie.
Sapir, Edward. 1930. Southern Paiute, a Shoshonean language. *Proceedings of the American Academy of Art and Sciences*. 65(1).
Schachter, Paul. 1976. The Subject in Philippine Languages, Topic, Actor, Actor-Topic or None of the Above. In *Subject and Topic*, ed. by Li C. N., 491–518. New York: Academic Press.
———. 1984, Semantic-role-based syntax in Toba Batak. In *Studies in the structure of Toba Batak*, ed. by Paul Schachter, 122–149. Los Angeles: UCLA.
——— and Fe T. Otanes 1972. *Tagalog Reference Grammar.* Berkeley: University of California Press.
Schütz, Albert J. 1985. *The Fijian Language*. Honolulu: University of Hawaii Press.
Shkarban, Lina I. and Gennadij E. Rachkov. 2007. Reciprocal, sociative, and comitative

constructions in Tagalog. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 887–931. Amsterdam: John Benjamins.
Smith, Alexander D. 2017. The Western Malayo-Polynesian Problem. *Oceanic Linguistics* 56(2):435–490.
Sneddon, J. N. 1984. *Proto-Sangiric and the Sangiric languages*. Canberra: Pacific Linguistics.
———.1993. The drift towards final open syllables in Sulawesi languages. *Oceanic Linguistics* 32:1–44.
———. 1996. *Indonesian: a comprehensive grammar.* London: Routledge.
Sung, Li-May and Chia-chi Shen. 2006. Reciprocals in Kavalan and a Typological Comparison. In *Streams converging into an ocean : Festschrift in honor of professor Paul Jen-Kuei Li on his 70th birthday*, ed. by Henry Y. Chang, Lilian M. Huang, and Dah-an Ho. 239–277. Taipei: Institute of Linguistics, Academia Sinica.
Tabulova, N. T. 1976. *Grammatika abazinskogo jazyka [A Grammar of Abaza]*. Cherkessk: Karachaevo-cherkesskoe otdelenie Stavropol'skogo knizhnogo izd-va.
Tuular, Klara B. 2007. Reciprocals, sociatives, comitatives, and assistives in Tuvan. In *Typology of Reciprocal Categories and Constructions*, ed. by Vladimir P. Nedjalkov, 1163–1229. Amsterdam: John Benjamins.
van den Berg, René. 1996. The demise of focus and the spread of conjugated verbs in Sulawesi. In *Papers in Austronesian Linguistics No. 3*, ed. by Hein Steinhauer, 89–114. Canberra: Pacific Linguistics.
———. 2013 [1989]. *A grammar of the Muna language*. Dordrecht: Foris Publications.
van der Tuuk, H. N. 1971. *A grammar of Toba Batak.* The Hague: Martinus Nijhoff.
van der Veen, H. 1940. *Tae' (Zuid-Toradjasch)-Nederlandsch woordenboek* [Tae' (South-Toraja)-Dutch dictionary]. The Hague: Martinus Nijhoff.
Vanoverbergh, Morice. 1955. *Ilokano grammar.* Baguio City: Catholic School Press.
Walker, Dale F. 1976. *A grammar of the Lampung language: the Pesisir dialect of Way Lima*. NUSA 2.
Wang, Peter, Robert Hunt, Jeff McGriff, and Richard E. Elkins. 2006. *The Grammar of Matigsalug Manobo.* SIL Philippines.
Warneck, J. 1977 [1906]. *Toba-Batak-Deutsches Wörterbuch* [A Toba-Batak-German Dictionary]. KITLV. The Hague: Martinus Nijhoff.
Watahomigie, Lucille J., Jorigine Bender, Philbert Watahomigie, Sr. and Akira Y. Yamamoto. 2001. *Hualapai Reference Grammar (Revised and Expanded Edition)*. Kyoto: Nakanishi Printing Co.
Wolfenden, Elmer P. 1971. *Hiligaynon reference grammar.* Honolulu: University of Hawaii Press.
———. 1975. *A Description of Hiligaynon Syntax*. Norman: SIL.
Wolff, John U. 1996. The development of the passive verb with pronominal prefix in western Austronesian languages. In *Reconstruction, classification, description: festschrift in honor of Isidore Dyen*, ed. by Bernd Nothofer, 15–40. Hamburg: Abera.
———. 2012 [1972]. *A dictionary of Cebuano Visayan*. Philippine Journal of Linguistics Special Monograph Issue no. 4. Manila: Linguistic Society of the Philippines. Electronic file provided by the Project Gutenberg.
Woollams, Geoff. 1996. *A grammar of Karo Batak, Sumatra*. Canberra: Pacific Linguistics.
Zeitoun, Elizabeth. 2002. *Reciprocals in the Formosan languages: A preliminary study*. Paper presented at the Ninth International Conference on Austronesian Languages (9-ICAL), Canberra, 8–11 January 2002.
——— and Lillian M. Huang. 2000. Concerning ka-, an overlooked marker of verbal derivation in Formosan languages. *Oceanic Linguistics* 39:391–414.
Zobel, Erik. 2005. Buol. In *The Austronesian languages of Asia and Madagascar*, ed. by K. Alexander Adelaar and Nikolaus. P. Himmelmann, 625–648. London and New York:

Routledge.
Zorc, R. David. 1968. *A Study of the Aklanon Dialect. Volume One: Grammar.* Kalibo, Aklan: Aklan Printing Center.
———. 1977. *The Bisayan Dialects of the Philippines: Subgrouping and Reconstruction*. Canberra: Pacific Linguistics.
———. 1986. The genetic relationships of Philippine languages. In *FOCAL 2: Papers from the Fourth International Conference on Austronesian Linguistics*, ed. by Paul Geraghty, Lois Carrington, and S. A. Wurm, 147–173. Canberra: Pacific Linguistics.

Curriculum Vitae

Yuko Kitada

Contact information
Address: Eigelstein 85, 50668 Köln, Germany
Email: yukokitada@outlook.com

Personal data
Date of birth: 26 July 1983
Nationality: Japanese
Gender: Female

Research interests
Austronesian linguistics (historical and comparative morphology and syntax of Western Indonesian, Sulawesi, and Philippine languages), historical linguistics (historical morphology), morphology, syntax, semantics, and linguistic typology.

Education

2010-	Doctor's program, University of Cologne, Germany
2006-2008	Master's program, Tokyo University of Foreign Studies, Tokyo, Japan
2002-2006	Bachelor's program, Department of Southeast Asian Studies, Tokyo University of Foreign Studies, Tokyo, Japan

Academic qualifications

2008	Master of Arts, Tokyo University of Foreign Studies Specialization: Linguistics Thesis: The Applicative in Madurese.
2006	Bachelor of Arts, Tokyo University of Foreign Studies Specialization: Language and Area Studies (Malaysian) Thesis: Mangarai-go no shoyū hyōgen [Possessive Expressions in Manggarai].

Publication

Kitada, Yuko. 2015. The Etymology of the Sociative-Progressive Circumfix in Suwawa (Gorontalo-Mongondowic). In *Language Change in Austronesian Languages*, ed. by Malcolm D. Ross and I Wayan Arka, 79–92. Canberra: Asia-Pacific Linguistics.

Yuko Kitada