

Studies in Philippine Languages and Cultures (formerly *Studies in Philippine Linguistics*) is a joint venture undertaken by the Linguistic Society of the Philippines and the Summer Institute of Linguistics devoted to the timely publication of papers of an empirical or theoretical nature which contribute to the study of language and communicative behavior in the Philippines. A certain proportion of the papers published will be penultimate versions of works of special interest to the Philippine linguistic community destined for further publication in more final form elsewhere.

Scholars engaged in the study and description of Philippine languages and cultures are hereby invited to submit contributions for inclusion in this series.

Volumes 15–19 of *Studies in Philippine Languages and Cultures* are a collection of selected papers originally presented at the Tenth International Conference on Austronesian Linguistics (10-ICAL), Puerto Princesa, Palawan, Philippines, 17-20 January, 2006.

SPLC 19 10-ICAL Borneo, Sulawesi & Philippines Papers, volume editors Phil Quick & Joanna Smith

For inquiries and submissions, please write to:

Academic Publications Manager
Summer Institute of Linguistics
PO Box 2270 CPO
1099 Manila
Philippines

E-mail: Acad-Pub_Philippines@sil.org

ISSN 0119-6456
0808-2.5C

© 2008 Linguistic Society of the Philippines and
Summer Institute of Linguistics

Long Terawan Berawan Phonology: Questions on Diphthongs and Syllabicity*

Jürgen M. Burkhardt

Johann Wolfgang von Goethe Universität Frankfurt a.M., Germany
Language Development SIL (M) Bhd., Malaysia

The paper takes a second look at some aspects of the phonology of Long Terawan Berawan (LTB), a language spoken in North-Sarawak, Malaysia. Of special interest in LTB is the ultimate syllable. Blust (1992) and Clayre (1996) identify a phonemic contrast for long versus short consonants in its onset and for long versus short vowels in its nucleus.

Building on their findings and providing new field data, the present paper examines ambivalent LTB sound combinations, especially the notion of diphthong with respect to syllabicity.

Diphthongs are identified in terms of stress patterns, i.e. a stressed vowel carrying the syllable peak followed by an unstressed non-syllabic semivowel reinterpreted as approximant. Vowel combinations that do not show this stress pattern are identified as constituting the peak of two different syllables. On this basis, some of Clayre's monosyllabic words are reinterpreted as disyllabic and a revised notion of the LTB syllable is proposed.

1. Introduction

The Long Terawan variety of Berawan (henceforth LTB) has received considerable attention in the last three decades. Blust's (1974) dissertation contains a 100-item wordlist of the language variety in which some of the items exhibit long consonants. Asmah (1983) provides a first preliminary phonological sketch of the language. In passing, she mentions consonant lengthening as a phonetic process. Her article does not record contrastive vowel length. Blust (1992:413) shows the phonemic status of long consonants and mentions a contrast of long versus short vowels. Clayre's (1996) phonological analysis of LTB provides a full inventory of LTB vowel and consonant phonemes and convincingly demonstrates the phonemicity of long vowels and long consonants based on acoustic phonetic evidence. In the same article, she provides a generalization about the nature of the LTB syllable. As an area that needs further investigation, she mentions the notion of diphthong which will be, in combination with the question of syllabicity, the focus of this investigation. García-Bellido & Clayre (1997) employ the concept of prosodic constraints to explain gaps in the combination

* I wish to thank Dr Philip Quick, Dr Calvin Rensch and Dr Kenneth Gregerson for their consultative and editorial input as well as my dear wife Jey for her editorial help. An earlier version of the paper was presented at the 10th International Conference on Austronesian Linguistics, 17–20 January 2006, Puerto Princessa, Palawan, Philippines. I would like to extend my thanks to Language Development SIL (M) Bhd., for providing funds for me to participate in that conference.

of segments in the Berawan word. The latter article is largely not relevant for the scope of our investigation since I do not employ a framework of prosodic weight distinctions.

Sections 2 and 3 give a summary of Clayre's findings with respect to consonant and vowel phonemes, their distributional restrictions, and presents a discussion on the phonemic status of schwa and the central vowel phone [ɐ]. Section 4 establishes unambiguous LTB syllable patterns and then introduces Clayre's notion of the LTB syllable. Section 5 investigates ambiguous vowel sequences and reinterprets them. Section 6 offers a revised notion of the LTB syllable and lists the different types of vowel-approximant sequences found in LTB. Section 7 concludes the paper.

2. Consonant Phonemes

Blust (1992:412–413, 1995:126) as well as Clayre (1996:218) list 19 consonant phonemes of which 14 appear in short as well as long form (see Table 1).

Table 1. LTB consonant phonemes¹

Obstruents	p(:) b(:)	t(:) d(:)	c(:) j(:)	k(:) g(:)	/
Fricatives		s			h
Nasals	m(:)	n(:)	ɲ(:)	ŋ(:)	
Vibrants		r(:)			
Laterals		l(:)			
Approximants	w		y		

All simple consonant phonemes appear word initially and word medially. The exceptions are the glottal stop, which is a phoneme word medially but a mere phonetic vowel onset word initially, and the glottal fricative /h/, whose occurrence is restricted to the word final position. The labio-velar approximant /w/, I recorded word initially only for proper names.

Word finally, the occurrence of simple consonants is restricted to voiceless plosives (including the glottal stop),² the glottal fricative /h/ and nasals (with the exception of the palatal nasal). As a result of her phonological analysis, Clayre does not posit a word final occurrence of the approximants y and w. In section 5, we will reexamine the notion of approximants with respect to the coda of the final syllable.

¹ For notational simplicity, I am using /c/ for the voiceless affricate [cç] and /j/ for the voiced affricate [jɟ]. Another difference between the above table and Clayre's is that /h/ appears as an approximant in her table, but as a fricative here. Analogously to Clayre, /y/ is used for the palatal approximant [j]. All other phonological consonant symbols used in this paper have the same phonetic value as that symbol in IPA. I have simplified Clayre's table by combining plosives and affricates into one category, i.e. obstruents, and by subsuming the two alveo-palatal affricates and the palatal nasal under the same place of articulation. All LTB language examples listed in this paper are elicited by the author himself if not referenced otherwise.

² Word final plosives are always unreleased, therefore, for notational convenience, the phonetic sign for unreleasedness is not indicated.

Long consonants, on the other hand, appear only in word medial position as in Table 2. They have phonemic status since they contrast with simple, that is short, consonants.³

Table 2. Contrasts between short and long consonants⁴

Contrast	LTB	English	LTB	English
p – p:	/napa:n/	to winnow	/nap:a:n/	to slap
b – b:	/labih/	dirty	/lab:eh/	end
t – t:	/lutoh/	soggy	/lut:o/	to float
d – d:	/adi:ŋ/	ear	/ad:iŋ/	earwax
c – c:	/dici:ŋ/	wall	/kac:i:ŋ/	button
j – j:	/paju/	to scold	/kaj:uh/	wood
r – r:	/mareh/	eight	/tar:eh/	younger sibling
l – l:	/kulah/	to turn	/kul:ah/	thin
k – k:	/pəlike/	horsefly	/pəlak:eh/	omen bird
g – g:	/lagu/	song	/ag:uk/	mumps
m – m:	/dimah/	rubbish	/dim:ah/	five
n – n:	/sanay/	sun heat	/san:ay/	insect species
ɲ – ɲ:	/paɲin/	commoners	/maɲ:in/	to drown
ŋ – ŋ:	/liŋan/	out of view	/liŋ:an/	self

3. Vowel Phonemes

Clayre (1996:223–225) identifies six vowels of which all except schwa appear long as well as short. She lists the vowel phonemes and their phonetic realizations⁵ as shown in Table 3.

³ The contrast is neutralized after penultimate schwa, where all consonants occur automatically long (Blust 1995:124). Thus, they are phonemically represented with a simple consonant symbol throughout this paper.

⁴ The LTB examples in Table 2 that show the contrasts /c/-/c:/, /r/-/r:/, /ɲ/-/ɲ:/ and /ŋ/-/ŋ:/ are taken from Blust 1992:413. All other examples, except for the one that shows the contrast /g/-/g:/, are taken from Clayre 1996:218–221.

⁵ Clayre uses the label *close* instead of *high* and *open* instead of *low*. The above label *mid* comprises Clayre's *mid-close* vowels ([e] and [o]), her *mid-open* vowels ([ɛ] and [ɔ]) and her *mid-central* vowel [ə], the neutral vowel in the system. I agree with Clayre that the long low (open) vowel [ɑ:] is considered a central vowel, even though it is pronounced further back. Structurally, it functions like a central vowel. The short low (open) central vowel [ɐ] is raised if it occurs before a word final nasal, as in /liŋan/ [liŋɛn], but not before word final h or glottal stop, as in /dimah/ [dimeh] or /silaʔ/ [silɛʔ]. In the latter environment, [ɐ] is pronounced somewhere halfway between the open-mid and open position, which is the place assigned to it in the current version of the IPA vowel chart.

Table 3. LTB vowel phonemes with phonetic realizations

	front	central	back
high	/i:/ [i:]		/u:/ [u:]
	/i/ [ɪ]		/u/ [ʊ]
mid	/e:/ [e:]	/ə/ [ə]	/o:/ [o:]
	/e/ [ɛ]		/o/ [ɔ]
low		/a/ [ɐ]	
		/a:/ [ɑ:]	

As we see in Table 4, there is phonemic contrast between short vowels and their long counterparts. This contrast only occurs in the ultima of an LTB word (Clayre 1996:212).

Table 4. Contrasts between short vowels and their long counterparts in LTB⁶

Contrast	LTB		English	LTB		English
i – i:	/usin/	[ʔusin]	money	/usi:n/	[ʔosi:n]	rain
e – e:	/mat:e?/	[mɛt:e?]	to throw	/kat:e:ʔ/	[kɛt:e:ʔ]	to throw away
a – a:	/nakan/	[nɛkɛn]	to climb	/naka:n/	[nɛka:n]	fed (perfect)
u – u:	/gadun/	[gɛdun]	green	/ladu:n/	[lɛdu:n]	to crash
o – o:	/nipo?/	[nɪpɔ?]	to surround	/nipo:ʔ/	[nɪpo:ʔ]	to put together

Clayre appears undecided on the number of LTB vowel phonemes. On the one hand, she lists eleven vowel phonemes (1996:223–225) as presented in Table 3. On the other hand, to account for the fact that long consonants and long vowels occur exclusively in the ultimate syllable, she does not analyze them as “full phonemes of the language, but rather as a feature, or prosody of the nuclear [ultimate] syllable” (1996:212). That would reduce the number of vowel phonemes to six and add a feature of phonemic length for the ultima. This difference in interpretation does not bear on the current investigation. So, for reasons of simplicity, I adopt the first approach and count eleven LTB vowel phonemes as presented in Table 3.

Clayre’s list of vowel phonemes in long and short forms contains three more items than Asmah’s (1983:575), who identified eight vowel phonemes, that is, in her notation, /i, ɛ, ê, e, a, o, ô, u/. Asmah does not note any systematic differences in vowel length. Blust reports contrastive vowel length for /i/, /e/ as well as /o/ and mentions that it “was often recorded as a qualitative difference”. He then points out that “if a qualitative analysis of vowel contrasts is adopted, the number of vowel phonemes will increase to nine” (1992:412).

⁶ The examples that show the contrasts /i-/i:/, /e-/e:/ and /a-/a:/ are from Clayre (1996:223–224).

3.1. Is [ɐ] in the ultima a phonetic realization of /a/ or of /ə/?

Clayre assigns [ɐ]⁷ to her short /a/ phoneme, the counterpart to /a:/ [ɑ]. Blust (1992:411–412) hints at assigning it to /a/ or /ə/, and while he does not commit himself explicitly to either interpretation, he seems to favor assigning it to /ə/, as is discernible from his phonological representation of LTB words, e.g. /dimeh/ ‘rubbish’ and /dimmeh/ ‘five’, and his counting of vowel phonemes, i.e. nine (He does not note a length contrast for /u/ – /u:/).

3.1.1. Assigning ultimate [ɐ] to /ə/ (and ultimate [ɑ:] to /a/)

If [ə] did not occur in the ultima, we could assign [ɐ] to /ə/ and get a symmetrical ten vowel phoneme system as in Table 5.

Table 5. Possible LTB ten vowel system if [ə] did not occur in the ultima

/i:/	[i:]			/u:/	[u:]
/i/	[ɪ]			/u/	[ʊ]
/e:/	[e:]	/ə/	[ə] non-ultima [ɐ] ultima	/o:/	[o:]
/e/	[ɛ]			/o/	[ɔ]
		/a/	[ɐ] non-ultima [ɑ:] ultima		

However, we have to take into account that LTB exhibits two environments in which [ə] does appear in the ultimate LTB syllable, namely, before [ɾ]#⁸ and [ʊʔ]#, as in (1).

- (1) [kʊtəɾʔ] ‘explode’
 [bək:əɾʔ] ‘heavy’
 [pɪtəʊʔ] ‘to hang’
 [sətəʊʔ] ‘touch’

Schwa contrasts with [ɐ] in the ultima before an approximant-glottal stop sequence as in (2).

⁷ In her 1996 article, she uses the symbol [ʌ] to denote short /a/, whereas in García-Bellido & Clayre (1997), [ɐ] is used instead. In the present investigation, the latter is used consistently as the phonetic realization of /a/.

⁸ Whereas Clayre (1996) does not report an occurrence of [ə] in the ultima, Blust’s (1974:303) LTB wordlist contains one record with this environment: /ɲəɾʔ/ [ɲəɾʔ] ‘sharp’, in his notation *ñəyq*. Furthermore, Blust (1992:420–421) contains two entries which exhibit schwa before /yʔ/[iʔ]: /bəkəyʔ/ [bək:əɾʔ] ‘heavy’ and /kəbəyʔ/ [kəb:əɾʔ] ‘long’ (in Blust’s notation *bekeiʔ* and *kebeiʔ*, whereby orthographic *e* denotes schwa).

- (2) [kʊtəɪʔ] ‘explode’ – [kʊləɪʔ] ‘skin’
 [pɪtəʊʔ] ‘to hang’ – [lɪtəʊʔ] ‘murky’

Furthermore, there is penultimate contrast of [ə] with [ɪ], [e] and [ʊ] as Table 6 shows.

Table 6. Evidence of contrast between schwa and /i/, /a/ and /u/⁹

Contrast	LTB		English	LTB		English
ə – i	/pətəwʔ/	[pət:əʊʔ]	massage	/pɪtəwʔ/	[pɪtəʊʔ]	to hang
ə – a	/məla:ʔ/	[məl:a:ʔ]	awake	/mala:ʔ/	[məla:ʔ]	to take
ə – u	/dəkih/	[dək:ih]	house post	/dukih/	[duk:ih]	thorn

Since [ə] occurs in the ultima before Aʔ# (A = approximant) and contrasts with [e] in that environment, it is only possible to assign ultimate [e] to /ə/ if ultimate [ə] is not assigned to the same phoneme.

On the other hand, [ə] could be assigned to /e/, since [ɛ] and [ə] are in complementary distribution in the ultima. Thus, it is possible to assign the two phones to the same phoneme:

/e/ [ə] doubly closed syllable, i.e. Aʔ#
 [ɛ] elsewhere

Then, [e] can be assigned to /ə/ for the ultima and to /a/ for other syllables. We would then still reap a system with 10 vowel phonemes as Table 7 shows.

Table 7. LTB vowel phoneme system with assigning ultimate [e] to /ə/

/i:/	[i:]			/u:/	[u:]	
/i/	[ɪ]			/u/	[ʊ]	
/e:/	[e:]		/ə/	[e] ultima	/o:/	[o:]
/e/	[ə] doubly closed syllables			[ə] elsewhere	/o/	[ɔ]
	[ɛ] elsewhere		/a/	[a:] ultima		
				[e] elsewhere		

The vowel system this analysis yields is somewhat awkward: Two allophones are assigned to the short mid-front vowel /e/, whereas all other front and back vowels have only a singular phonetic realization.

Furthermore, [ə] could be assigned to /o/ instead:

⁹ The example that shows the contrast /ə/-/u/ is from Clayre (1996:239).

/o/	[ə]	doubly closed syllable
	[ɔ]	elsewhere

If we assign [ə] to either /e/ or /o/, we need a strong reason to justify why we favor one over the other. Irrespective of whether we choose /e/ or /o/, we reap the same awkwardness in the LTB vowel phoneme system.

3.1.2. Assigning ultimate [ɐ] to /ə/ for closed syllables and to /a/ for doubly closed syllables

Taking the contrasts between [ə] and [ɐ] in the penult as well as in the doubly closed ultima into account, we could salvage the symmetry in the ten vowel phoneme system presented in Table 5 above by revising it, as Table 8 shows.

Table 8. Possible LTB ten vowel system if [ɐ] and [ɑ:] did not contrast before Aʔ#

/i:/	[i:]			/u:/	[u:]
/i/	[ɪ]			/u/	[ʊ]
/e:/	[e:]	/ə/	[ɐ] closed ultima	/o:/	[o:]
/e/	[ɛ]		[ə] ultima/_Aʔ# and non-ultima	/o/	[ɔ]
		/a/	[ɑ:] closed and open ultima		
			[ɐ] ultima/_Aʔ# and non-ultima		

The vowel system presented in Table 8 would be suitable if there wasn't any contrast between the short and the long low vowel in doubly closed syllables.

Blust 1992 and 1995 as well as his 1974-wordlist do not report LTB words showing the occurrence of [ɑ:] before Aʔ#.

On the other hand, Clayre (1996) and I recorded words with [ɑ:] in this environment as shown in example (3).¹⁰

- (3) [pa:ɪʔ] 'bitter' (Clayre 1996:217; Blust (1992:418) recorded [paiʔ])
- [la:ɪʔ] 'displeased' (Clayre 1996:239)
- [ma:ʊʔ] 'drunk' (Clayre 1996:217; Blust (1992:418) recorded [maoʔ])
- [la:ʊʔ] 'crow of the cock' (my recording)
- [tən:ɑ:ɪʔ] 'guts' (my recording; Blust (1974:277) recorded [tənaeʔ])

There is contrast between [ɐ] and [ɑ:] before Aʔ# as the examples in (4) show.

- (4) [pa:ɪʔ] 'bitter' – [kəpɛɪʔ] 'near'
- [la:ɪʔ] 'displeased' – [lɛɪʔ] 'upper arm'
- [ma:ʊʔ] 'drunk' – [gimeʊʔ] 'root'
- [la:ʊʔ] 'crow of the cock' – [mureʊʔ] 'to make'

¹⁰ The notion of doubly closed codas will be discussed in detail in section 5.4.1.

The vowel system listed in Table 8 does not account for this contrast. We therefore have to abandon the possibility of a ten vowel system and assume an eleven vowel system with three central vowel phonemes, that is /ə/, /a/ and /a:/.

3.1.3. Assigning ultimate [ɐ] to /a/ and ultimate [ɑ:] to /a:/ (as well as ultimate [ə] to /ə/)

The LTB vowel system presented in Table 3 assigns each of the three central phones to a different phoneme, that is [ə] to /ə/, [ɐ] to /a/ and [ɑ:] to /a:/. It is the most suitable description of the LTB vowel system for the following reasons:

First, [ə] is exclusively assigned to /ə/, which gives the schwa phoneme the same phonetic realization in all syllable types, that is [ə]. This has the advantage of simplifying the analysis.

Second, the same reasons hold for assigning ultimate [ɐ] to /a/; phoneme /a/ with phonetic realization [ɐ] already exists in the penult, so, linking /a/ and [ɐ] in the ultima too makes the analysis more consistent and simpler.

Third, assigning [ɐ] to /a/ gives /a:/ a short counterpart, which creates a beautiful symmetry as all other non-neutral vowels show this contrast. This also accounts for the fact that /a:/ [ɑ:] is by far more frequent than the other long vowels and it would be rather surprising if it did not exhibit a short counterpart which all the others have. The symmetry this approach yields is summarized in the following two points:

- phonemic symmetry : It yields /V/ – /V:/ correspondences for all non-neutral vowels.
- phonemic – phonetic symmetry: /V/ [V], that means, all vowel phonemes are assigned a singular phonetic realization.

Based on the above conclusions, all vowel phonemes occur in the ultima in general, that is /i/, /i:/, /e/, /e:/, /ə/, /a/, /a:/, /u/, /u:/, /o/ and /o:/.

In the open ultima, only five vowel phonemes occur. There is no contrast between short and long vowels in this environment, because all vowels are long word finally. For this reason, it is not surprising that schwa, which cannot be long, does not occur in this position.

The closed ultima is the only syllable which exhibits phonemic contrasts between short and long vowels. All vowel phonemes except schwa occur in this position, that is /i/, /i:/, /e/, /e:/, /a/, /a:/, /u/, /u:/, /o/ and /o:/. The contrast is neutralized before the glottal fricative h, where they are always short. The occurrence of schwa in the ultima is limited to doubly closed syllables.

In doubly closed ultimas, only the three central vowel phonemes occur, that is /ə/, /a/ and /a:/.

In the penultimate syllable, only four vowel phonemes occur which are all short, that is /i/ [ɪ], /ə/ [ə], /a/ [ɐ] and /u/ [ʊ]/ (Clayre 1996:214, 223–224).¹¹ In the

¹¹ Before a consonant, penultimate /i/ is phonetically realized as [ɪ], as in /tisu/ [tɪsu:] (Clayre 1996:216). Before an unlike vowel, however, there seems to be fluctuation between penultimate [i] and [ɪ]. For example, Clayre (1996:217) lists [bɪu:] ‘wind’ whereas I recorded [biu:].

antepenult, the vowel phoneme inventory is limited to schwa.¹² LTB does not seem to have words with more than three syllables.¹³ Table 9 summarizes the distribution of LTB vowel phonemes.

Table 9. Occurrence and number of vowel phonemes per syllable

syllable	antepenult	penult	open ultima	closed ultima	doubly closed ultima
type of vowel phonemes	—	i	i	i – i:	—
	—	—	e	e – e:	—
	ə	ə	—	—	ə
	—	a	a	a – a:	a – a:
	—	u	u	u – u:	—
	—	—	o	o – o:	—
number of vowel phonemes	1	4	5	10	3

4. The Syllable

4.1. Unambiguous syllable patterns

Clayre (1996:213) postulates that every LTB syllable has an obligatory onset and thus minimally the shape CV. This is true of trisyllabic words, which always begin with a consonant phoneme. Thus, antepenultimate syllables, only occur in the shape CV as in (5).

¹² Schwa is the sole vowel I recorded for the antepenult, with the exception of one Malay loan word. Blust (1992:413) and Clayre (1996:213) assume three vowel phonemes for the antepenult, that is /i/, /ə/ and /u/. However, the only antepenult vowel found in Blust's published data is /ə/. Clayre (1996:213) records three words with a vowel other than schwa in the antepenult: <tabip:é? > 'firefly', <bisika:n> 'wild bee' and <tupanai> 'recently'. The LTB entries for 'wild bee' and 'firefly' I recorded as /bəsika:n/ and /təbip:eh/ whereas my entry for 'recently' consists of two words, that is /tuh panay/, literally 'this recent'. The only word with an antepenultimate nucleus other than /ə/ I recorded is a Malay loan word, /sudagal/ 'rich person' (from Malay saudagar 'merchant, trader').

¹³ Neither Clayre (1996:212) nor I recorded any words that consist of four or more syllables. Blust's (1974:271) wordlist contains one 4-syllable word that is /pəŋ-kətu?oh/ 'right'. However, according to LTB native speaker Timok Belay Wan, this expression constitutes not one but two words, that are /pəŋ kətu?oh/ "on (the) right". Timok furthermore pointed out that /pəŋ/ is not a LTB word, but used in the neighboring Batu Belah Berawan variety. According to him, "on (the) right" is /tan kətu?oh/ in LTB.

- (5) /bəlira:ŋ/ 'monitor lizard'
 /kəbəlɪn/ 'hill'
 /kəlawaʔ/ 'spider'
 /təlana:ʔ/ 'soul'

On the other hand, disyllabic and monosyllabic LTB words can begin a vowel phoneme. In such cases, the onset is a non-phonemic glottal stop shown in (6).

- (6) /aka:ŋ/ [ʔəkɑ:ŋ] 'ghost'
 /ina:ʔ/ [ʔmɑ:ʔ] 'mother'
 /ul:oh/ [ʔul:ɔh] 'head'
 /e:ŋ/ [ʔe:ŋ] 'waist'
 /u:k/ [ʔu:k] 'loud shout'
 /u:n/ [ʔu:n] 'edible soft part of bamboo shoot'

Therefore, it is feasible to postulate a V syllable pattern for the penult and the ultima, making its syllable onset optional. This conclusion implies that monosyllables behave like the ultimate syllable of a polysyllabic word. This implication rests on the fact that monosyllables show the same vowel length contrasts that are exclusive to the ultima of polysyllabic words as shown in Table 10.

Table 10. Contrasts between short vowels and their long counterparts in LTB monosyllables¹⁴

Contrast	LTB		English	LTB		English
i – i:	/biʔ/	[biʔ]	at times	/bi:ʔ/	[bi:ʔ]	load
e – e:	/leʔ/	[leʔ]	only	/le:ʔ/	[le:ʔ]	to, in the direction of
a – a:	/tan/	[tɛn]	side	/ta:n/	[ta:n]	bearing something patiently
u – u:	/puʔ/	[puʔ]	hair	/bu:ʔ/	[bu:ʔ]	where
o – o:	/ŋoʔ/	[ŋɔʔ]	throat	/mo:ʔ/	[mo:ʔ]	supposing, given that

As there are monosyllables without an onset and monosyllables in general behave like the ultima of a polysyllabic word, it would be rather surprising if polysyllabic words with an onsetless ultima did not occur in LTB.

Thus, I posit V as a possible syllable type for the ultima in general. Sections 5.3, 5.4.2 and 5.6 provide LTB examples that are reinterpreted as words which lack an ultimate onset. In those sections, further reasons are given to show the optionality of the ultimate onset.

As outlined above, there is contrast between short and long vowels as well as short and long consonants in the closed ultimate syllable. Furthermore, the antepenultimate onset is obligatory, but the penultimate and ultimate onset is optional. This yields the following unambiguous phonemic syllable pattern for the closed ultima: (C(:))V(:)C. The

¹⁴ The example that shows the contrast /u/-/u:/ is from Clayre (1996:223–224).

nucleus of the final open syllable is always long. Therefore, length is merely phonetic in this syllable type, reaping a mere (C(:))V, not a (C(:))V(:) as phonemic syllable type.

The unambiguous syllable patterns on which our subsequent discussion will be based are shown in Table 11.

Table 11. Unambiguous syllable patterns in LTB

antepenult	penult	closed ultima	open ultima
CV	(C)V	(C(:))V(:)C	(C(:))V

4.2. Clayre's notion of the LTB syllable

Clayre distinguishes pre-nuclear from nuclear syllables. She defines the pre-nuclear syllable as a simple syllable “that contains an obligatory onset and rhyme. The onset consists of a single consonant, the rhyme of a short vowel, giving the pattern CV” (1996:213). Any syllable occurring before the ultimate (nuclear) syllable she considers pre-nuclear.

4.2.1. The nuclear syllable

Clayre defines the nuclear syllable (S) as a complex syllable that can only occur as the ultimate syllable of a Berawan word. According to that definition, its onset minimally consists of a short consonant (C) occupying one segment slot and maximally of a long consonant (C:) occupying two segment slots as in (7).

- (7) <labih>¹⁵ ‘dirty’ – CV.CVC pattern
 <lab:eh> ‘end’ – CV.CCVC pattern

The rhyme of the nuclear syllable she postulates as minimally binary and maximally ternary wherein X below can be occupied either by a consonant or a vowel. In her approach, rhyme structures consisting of two segments or prosodic positions are labeled as binary and the ones consisting of three segments ternary (Clayre 1996:215–217).

Table 12. LTB Binary and ternary rhymes according to Clayre 1996

Binary rhyme (two segments)	Ternary rhyme (three segments)
Rhyme	Rhyme
V X	V V X

Under a binary rhyme (VX), Clayre subsumes two patterns, i.e. VC and VV. As ternary rhymes, she lists VVC and VVV patterns. For rhymes that contain more than

¹⁵ I am using < > brackets to exclusively refer to Clayre's (1996) phonological representations of LTB words. For my own phonological interpretations, which are also used to cite Clayre or Blust data in sections 2,3, and 4.1, I use /./ (forward slash).

one V slot, I am using subscript indexes to indicate whether V slots are occupied by a long vowel (same index) or different vowels (different indexes). As we can deduce from Clayre's conceptualization outlined above, the notions of nucleus and coda as an intermediate level between rhyme and segment slots are not employed.

Clayre's analysis yields the following nuclear rhyme patterns in (8) and (9).

(8) **Binary rhymes**

VC	as in <lum> 'while, in'
V _i V _j	as in <mai> 'rattan'
V _i V _i	as in <bi:> 'lip'

(9) **Ternary rhymes**

V _i V _j C	as in <gium> 'cloud' and <lai?> 'arm'
V _i V _i C	as in <ki:ŋ> 'downriver'
V _i V _j V _k	as in <bəliəu> 'shaman'
V _i V _i V _j	as in <sapa:u> 'roof'
V _i V _j V _j	as in <mui:> 'wash' and in <mai:> 'rapids'

She mentions that the notion of diphthong in LTB needs further investigation.

The patterns V_iV_jC as in <gium> or <lai?>, V_iV_iV_j as in <sapa:u> and V_iV_jV_j as in <mui:> look like potential candidates for diphthongs. The V_iV_jV_k pattern as in <bəliəu> appears as a likely candidate for triphthongs.

A further question to be examined is whether all of the segments of these vowel sequences are part of the ultimate syllable or whether they are vowel sequences across the syllable break that separates the ultima from the penult. Thus, all rhyme patterns listed above, except for VC and V_iV_i, are ambiguous and have to be checked against unambiguous ones.

Clayre's V_iV_i as in <bi:> 'lip' is unambiguous and with respect to the phonological framework of prosodic weight Clayre (1996) and Garcia-Bellido & Clayre (1997) are using, it is justified to assign two vowel slots for that purpose. For the scope of my investigation, however, it suffices to consider the rhyme of an open ultima phonologically a mere V, since word final vowel length, as noted above and indicated in Garcia-Bellido & Clayre (1997:23), is predictably long.

Prosodic weight is not relevant for this paper's investigation. Therefore, I am employing a rather simple approach to the notion of syllable, subdividing it into onset and rhyme and the rhyme into nucleus and coda. The onset and coda can, if not empty, only be represented by consonants and/or approximants. The nucleus can only contain vowels.¹⁶ When referring to Clayre's notion of LTB syllable, I am using her terminology as outlined above.

¹⁶ The nucleus can be represented by vowels or diphthongs, if an alternative approach is adopted (see section 5.4.1 and the Appendix).

5. Examination of ambiguous vowel sequences

5.1. Word Stress and Clayre's notion of the nuclear rhyme

Clayre (1996:212) states that stress in LTB falls on the ultimate syllable, an observation that I share. On the other hand, Clayre's notion of the nuclear rhyme does not account for differences in stress placement as shown in (10) (I am underlining stressed segment slots).

(10)	$\underline{V}_i \underline{V}_j V_k$	as in <sap <u>a</u> :u>	[səpɑ:ʊ]	'roof'
	$V_i \underline{V}_j \underline{V}_k$	as in <mu <u>i</u> :>	[mui:]	'to wash'
	$V_i \underline{V}_j C$	as in <g <u>i</u> um>	[gɹum]	'cloud'
	$\underline{V}_i \underline{V}_j C$	as in <la <u>i</u> ?>	[ləiʔ]	'arm'
	$V_i \underline{V}_j \underline{V}_k$	as in <bəli <u>a</u> >	[bəlɪəʊ]	'shaman'

Thus, her notion of nuclear syllable is underspecified for the prediction of word stress.

5.2. $V_i V_j V_k$ pattern reinterpreted as vowel-approximant (V:A#)

$\underline{V}_i \underline{V}_j V_k$ as in <sapa:u> [səpɑ:ʊ] 'roof' behaves like a base vowel [ɑ:] with an off-glide [ʊ]. If it were interpreted as a long diphthong, it would occupy three rhyme slots in her framework and not match any unambiguous LTB syllable pattern, whose nucleus is maximally represented by two vowel slots. If, on the other hand, we interpret it as a long vowel followed by an approximant, that is /sapa:w/,

- it matches the unambiguous ultimate syllable pattern CV:C
- it predicts stress placement correctly, that is on V_i , which solely forms the nucleus in the proposed reinterpretation
- LTB words that fit this pattern are typically retentions of PAN word final vowel-approximant sequences as shown in Table 13

Table 13. Proto forms of LTB -V:A#

My reinterpretation	English	Proto form	Source
/pa <u>t</u> a:y/ [pəta:ɪ]	corpse	PAN *patay ‘dead’	Wurm & Wilson (1983:56)
/a <u>t</u> a:y/ [ʔəta:ɪ]	liver	PAN *atay	Wurm & Wilson (1983:123)
/ma <u>n</u> a:y/ [məna:ɪ]	male (animals)	PAN *manay	Wurm & Wilson (1983:126)
/la <u>k</u> a:w/ [ləka:ʊ]	walk	PMP *lakaw	Blust (1992:418)
/ka <u>s</u> a:w/ [kəsa:ʊ]	rafters	PAN/PMP *kasaw	Blust (1992:420)
/keluba <u>w</u> / [kələba:ʊ]	water buffalo	PPH *kaRabaw	Wurm & Wilson (1983:27)

5.3. V_iV_jV_j pattern reinterpreted as a V.V# sequence

LTB rhymes that fit into Clayre’s V_iV_jV_j pattern as <mui> [mɔi:] ‘to wash’ or <biu> [bi:ʊ] ‘wind’ are reinterpreted as /V.V:/ [V.V:] sequences for the following reasons:

- The second (long) vowel behaves like a word final monophthong in an ultimate syllable pattern with onset, e.g. /mubi/ [mubi:] ‘often’ and /niru/ [niru:] ‘see’
- The reinterpretation above accounts for the placement of stress since it predictably falls on the entire nucleus of the ultima, that is on [i:] or [u:] in the above examples.
- The first vowel behaves like a penultimate nucleus, which is always short.
- Words like these are perceived as disyllabic by my LTB native speaker informants, who tend to write them with an intervening approximant, that is <<muwi>> or <<biyu>>.
- Proto forms associated with LTB words that conform to this pattern are typically disyllabic as shown in Table 14.

Table 14. Proto forms of LTB -V.V#

My reinterpretation	English	Proto form	Source
/mu. <u>i</u> / [mɔ.i:]	to wash	PPH *quRis	Wurm & Wilson (1983:36)
/ti. <u>u</u> / [ti.ʊ:]	egg	PAN *təluR	Wurm & Wilson (1983:66)
/bi. <u>u</u> / [bi.ʊ:]	wind	PMP *baRiw	Blust (2000:315)
/bəli. <u>o</u> / [bəli.o:]	rat	PMP *bəlabaw	Blust (2000:315)
/di. <u>o</u> / [di.o:]	far	PMP *zauq	Blust (1992:419)

5.4. $V_iV_jC\#$ pattern

5.4.1. Reinterpreted as a $VA\#$ pattern

V_iV_jC as in <kulai?> [kʊləɪ?] ‘skin’ or <murau?> [mʊrəʊ?] ‘to make’ represent the ultimate rhyme in the words above. The status of the off-glides [ɪ] and [ʊ] is potentially ambiguous since on the one hand, each one could be an approximant forming a CC pattern with the following glottal stop (/yʔ/ and /wʔ/). On the other hand, [ɪ] or [ʊ] could be part of the ultimate nucleus, an off-glide to the low base vowel /a/ forming a diphthong with the latter (/ai/ or /au/). LTB does not have any unambiguous nuclei that consist of a combination of unlike vowels nor does it have unambiguous codas with CC clusters. Thus, either interpretation is possible.

If we consider the high vowels [ɪ] and [ʊ] as approximants, we have to introduce word final A? clusters, a restricted form of a CC cluster which allows only for the combination of an approximant followed by a glottal stop. Ultimate syllables ending with this cluster are labeled doubly closed syllables in this paper. Examples like <lai?>, <kulai?> and <murau?> are then reinterpreted as in (11).

(11)	/layʔ/	[ləɪʔ]	‘arm’	CVA?
	/kulayʔ/	[kʊləɪʔ]	‘skin’	CVCVA?
	/pələyʔ/	[pəl:əɪʔ]	‘to widen’	CVCVA?
	/murawʔ /	[mʊrəʊʔ]	‘to make’	CVCVA?
	/kucəwʔ/	[kʊcəʊʔ]	‘to reverse direction’	CVCVA?

This interpretation, which I have already implied in my discussion of the phonemic status of schwa in section 3, has the advantage that it limits nuclei to monophthongs and does not require the notion of diphthong at all. Thus, it makes it possible to treat glides uniformly as approximants. This approach is adopted in the remaining sections of the paper.

In a diphthongal interpretation, on the other hand, [kʊləɪ?] ‘skin’ etc. would be interpreted as /kulaiʔ/ CVCD?. Whereas this alternative interpretation avoids the introduction of a doubly closed rhyme, it would, on the other hand, require the introduction of four short diphthong phones, that is [əɪ], [əʊ], [ɪɪ], [ɪʊ], as well as two long ones, that is [ɑ:ɪ] and [ɑ:ʊ] (the occurrence of the latter two is addressed in section 5.7). Their phonemic/allophonic status would then have to be examined. This would make the analysis more complex. However, the diphthongal approach may, at least with respect to short diphthongs, reflect phonetic reality more accurately than the VA? approach described above, since the phone sequences [ɪɪ], [ɪʊ], [əɪ] and [əʊ] in the examples above have about the same length as the monophthong /a/ [ɛ] in /sulaʔ/ [sulɛʔ] ‘to recover’, but are perceptually shorter than the long monophthong /a:/ [ɑ:] as in /kula:ʔ/ [kula:ʔ] ‘fungus’ or the vowel approximant sequence /aw/ [ɛʊ] (see section 5.5) as in /buraw/ [bʊrɛʊ] ‘partially sighted’. Furthermore, the diphthongal interpretation may reflect historical sound changes more accurately, for example the LTB diphthongization of the monophthongs *i and *u (PAN *i > LTB [aɪ]; PAN *u > LTB [aʊ], see Table 18). The Appendix will outline the diphthongal interpretation and the generalization about LTB syllable patterns this alternative approach yields.

5.4.2. Reinterpreted as a V.V(:)C# sequence

Analogously to section 5.3, the pattern V_iV_jC as in <gium> ‘cloud’ (I recorded /giu:m/ [giu:m]) or <buaŋ> [buəŋ] ‘beetle’ is reinterpreted as a disyllabic V.V(:)C# sequence for the following reasons:

- It accounts for the placement of stress since it falls predictably on the entire nucleus of the ultimate syllable, that is [u:] or [ɐ], respectively above.
- The first vowel in the sequence, that is [ɪ] or [ʊ], is short like a typical penultimate nucleus.
- The second vowel in the sequence, that is [u:] or [ɐ], exhibits the range of vowel length only found in the nucleus of the closed ultimate syllable, that is the occurrence of short as well as long vowels.
- Native speaker perception tends to be disyllabic with the tendency to insert an approximant, e.g. <<giyum>> or <<giyuum>> and <<buwang>>.
- Proto forms associated with LTB words that conform to this pattern are typically disyllabic as shown in Table 15.

Table 15. Proto forms of LTB – V.V(:)C#

My reinterpretation	English	Proto form	Source
/gi. <u>u</u> :m/ [gi. <u>u</u> :m]	clouds	PPH *GaD/qum ‘cloudy’	Wurm & Wilson (1983:38)
/mi. <u>a</u> ŋ/ [mi. <u>ɐ</u> ŋ]	steep	PAN *tʰi[dd]aŋ	Wurm & Wilson (1983:203)
/bu. <u>a</u> ŋ/ [b <u>u</u> . <u>ɐ</u> ŋ]	beetle	PAN *tabuh/an ‘bee’	Wurm & Wilson (1983:16)
/ti. <u>u</u> ŋ/ [ti. <u>ʊ</u> ŋ]	egg plant	PMP *təRuŋ	Wurm & Wilson (1983:66)
/di. <u>o</u> :n/ [di. <u>o</u> :n]	leaf	PAN *Dahun	Wurm & Wilson (1983:118)
/ta. <u>o</u> :n/ [tə. <u>o</u> :n]	needle	PAN *zaRum	Wurm & Wilson (1983:137)

There are also LTB words ending with a glottal stop that fit this V.V(:)C # pattern as in (12).

- (12) /si.aʔ/ [si.ɐʔ] ‘ceremonial skull’
 /si.a:ʔ/ [si.ɑ:ʔ] ‘to lean on’

5.5. V_iV_j Pattern reinterpreted as vowel-approximant (VA#)

V_iV_j as in <mai> ‘rattan’, <parai> ‘rice plant’ or <paya> ‘sambar deer’. This pattern is not very common in LTB. Whereas I recorded [ma:i], [pə:a] and

[pɛiɑ:ʊ] for the above examples, I nevertheless found a few short vowels followed by a word final approximant as in (13).

- (13) /pələway/ [pələʋɛɪ] ‘fishing method’
 /daway/ [dəʋɛɪ] ‘wire’ (Malay loan)
 /paluy/ [pəlɔɪ] ‘stupid’ (Brunei Malay loan¹⁷)
 /ngar:aw/ [ŋɛr:əʋ] ‘to disturb’
 /utaw/ [ʔʊtəʋ] ‘type of baby clothing’
 /buraw/ [bʊrəʋ] ‘partially sighted’
 /pariw/ [pəriʋ] ‘to stagger’

The endings of these words are reinterpreted as a short vowel with an approximant as off-glide.

This way, stress placement is predicted accurately, for the ultimate nucleus consist now of only one vowel.

The reinterpretation conforms to the unambiguous ultimate syllable pattern CVC#

Proto forms of LTB words that conform to this pattern may have ended on a vowel followed by an approximant or other consonant, but so far, I have only found two matching proto forms¹⁸ of which the second one below refers to a Malay loanword as shown in Table 16.

Table 16. Proto form of LTB –VA#

My reinterpretation	English	Proto form	Source
/buraw/ ¹⁹	partially sighted	PAN *bulaR ‘cataract’	Wurm & Wilson (1983:31)
/daway/	wire	PAN *daway	Wurm & Wilson (1983:241)

5.6. V_iV_jV_k pattern reinterpreted as a V.V:A# sequence

Clayre’s V_iV_jV_k as in <bəliiau> ‘shaman’, the singular example she recorded for this pattern, looks like a triphthong at first glance. Irrespective of the fact that I recorded [bəlɪɑ:ʊ] instead, the two LTB speakers I consulted perceive this word as a trisyllabic one and tend to insert an approximant if they attempt to write it, that is <<beliyaw>> or <<beliyaaw>>. I have found only four more examples that conform to this pattern, as shown in (14).

¹⁷ This is from Brunei-Malay *paloi* ‘stupid’ (orthographic spelling), Anon. (1991:55).

¹⁸ PAN **kacaw* ‘cause disturbance’ (Wurm & Wilson 1983:32) may be the proto form for LTB /ngar:aw/ ‘to disturb’.

¹⁹ Central Dusun has a cognate with this word: *bolou* ‘blind’ (Anon. 1995:112).

- (14) /di.ə:y/ ~ /ji.ə:y/ [d~jɪ.ɑ:ɪ] ‘face’
 /du.ə:y/ [dʊ.ɑ:ɪ] ‘in-laws of a sibling’
 /bu.ə:w/ [bʊ.ɑ:ʊ] ‘to migrate’
 /pu.ə:w/ [pʊ.ɑ:ʊ] ‘not having slept enough’

The reinterpretation of $V_iV_jV_k$ as a $V.V:C\#$ sequence can be summarized in these points:

- It rules out a triphthongal interpretation which wouldn't conform to any of the unambiguous syllable patterns listed in Table 11 above.
- It predicts word stress correctly, that is on the nucleus of the ultimate syllable.
- It conforms to unambiguous rhyme patterns established above, that is V for the penult and V:C for the ultima.
- The proto reconstructions that I have found for two of the above examples exhibit a $V.CVC\#$ sequence as shown in Table 17.

Table 17. Proto forms of LTB – $V.V:A\#$

My reinterpretation	English	Proto form	Source
/bəli.ə:w/ [bəli.ɑ:ʊ]	shaman	PAN *bali(y)an	Wurm & Wilson (1983:184)
/d~ji.ə:y/ [d~jɪ.ɑ:ɪ]	face	PMP *daqey	Wurm & Wilson (1983:71)

5.7. Clayre's notion of a floating glottal – reinterpreted as a word final consonant

Clayre mentions that a “ $V:V$ sequence can have an additional glottal stop, at the end of the rhyme” (1996:217). She lists the monosyllabic words [mɑ:ʊʔ] ‘drunk’ and [pɑ:ɪʔ] ‘bitter’ as examples. In García-Bellido & Clayre (1997:37), a glottal stop following a $V:V$ sequence is associated with the short vowel in the sequence, forming a complex segment with the latter: $V \rightarrow Vʔ$. Thus, García-Bellido & Clayre's interpretation of the two examples above would be $CV_i\underline{V}_iV_j$ (my underlining of stress).

This assumption avoids the introduction of a 4-place rhyme structure in the framework of the two authors, which would otherwise be necessary since the rhyme in words like <auʔ>²⁰ ‘anus’ and <laiʔ> ‘arm’ is already classified as ternary and assigned the rhyme structure VVC by Clayre (1996:216).

On the other hand, the introduction of the floating glottal notion results in two different interpretations of the word final glottal stop. While it is considered a consonant in <auʔ> and <laiʔ>, it is assumed to be a component of the last vowel in <ma:uʔ> and <pai:ɪʔ>. This variable treatment of [ʔ] requires further scrutiny.

The word final glottal stop in LTB is phonemic. It contrasts with other plosives in the same position, as in (15).

- (15) /buraʔ/ [bʊrəʔ] ‘wasteful’ – /burak/ [bʊrək] ‘in a boiled state’
 /tad:aʔ/ [təd:əʔ] ‘sign’ – /adat/ [ədət] ‘customary law’

²⁰ I recorded [ɑ:ʊʔ].

Moreover, vowels followed by glottal stop at the end of a word contrast with final open vowels, as in (16). Likewise in (17), VA(:)? sequences contrast with VA(:) sequences word finally.

- (16) /ləmayaʔ/ [ləmɐɐʔ] ‘weak’ – /ləmaya/ [ləmɐɐa:] ‘to pass’²¹
 /ka:ʔ/ [ka:ʔ] ‘future tense word’ – /ka/ [ka:] ‘raven’
 (17) /ta:yʔ/ [ta:ɪʔ] ‘faeces’ – /ta:y/ [ta:ɪ] ‘all’
 /mala:wʔ/ [mɛla:ʊʔ] ‘weather’ – /ŋəla:wʔ/ [ŋəla:ʊʔ] ‘to crow’

The phonemicity of the word final glottal stop suggests it should be treated as a consonant in its own right, not as a feature of a word final vowel or vowel-approximant sequence.

Additional support for this conclusion can be drawn from a diachronic perspective. Clayre (1996:217) mentions that the glottal stop in [ma:ʊʔ] and [pa:ɪʔ] reflects a stop in earlier forms of the language, that is PMP *ma-buhek and PMP *paqit. Not surprisingly, ultimate rhymes that conform to the patterns /-ayʔ/ and /-awʔ/, such as /kulayʔ/, /sakayʔ/ and /gimawʔ/ also correspond to proto forms ending on a plosive as shown in Table 18.

Table 18. Proto forms of LTB –VAʔ#

LTB entry	English	Proto form	Source
/kulayʔ/ [kʊləɪʔ]	skin	PAN *kulit	Wurm & Wilson (1983:190)
/sakayʔ/ [səkəɪʔ]	painful	PAN *sakit	Wurm & Wilson (1983:146)
/gimawʔ/ [gimɐʊʔ]	root	PMP *Ramut	Blust (2000:315)

Therefore, I opt for treating the word final glottal stop uniformly as a full-fledged consonant. Since I am interpreting a pattern like /-ayʔ/ in /layʔ/ as VAʔ, LTB entries like [pa:ɪʔ] and [ma:ʊʔ] would therefore be interpreted as CV:Aʔ, as in (18).

- (18) /pa:yʔ/ [pa:ɪʔ] ‘bitter’ CV:Aʔ
 /la:yʔ/ [la:ɪʔ] ‘displeased’ CV:Aʔ
 /ma:wʔ/ [ma:ʊʔ] ‘drunk’ CV:Aʔ
 /la:wʔ/ [la:ʊʔ] ‘crow of the cock’ CV:Aʔ
 /təna:yʔ/ [tən:aɪʔ] ‘guts’ CVCV:Aʔ
 /səra:yʔ/ [sər:aɪʔ] ‘bad smell of cooking oil’ CVCV:Aʔ

Since the nucleus is a mere V: in my interpretation, stress is entirely predictable, that is on the long ultimate low vowel (/pa:yʔ/, /ma:wʔ/, etc.). The problem of a four-place rhyme pattern does not arise in this reinterpretation.²²

²¹ See Clayre (1994:250).

²² It also wouldn’t arise in a diphthongal interpretation, in which [a:ɪ]/_ʔ# and [a:ʊ]/_ʔ# could simply be regarded as long diphthongs, yielding D:ʔ as the ultimate rhyme pattern.

6. Revising the notion of the LTB syllable

The goal of the investigation undertaken in the previous sections has been twofold. Firstly, we were aiming at a generalization for the LTB syllable that adequately accounts for a consistent syllabic representation of all possible word forms of the language and that makes accurate predictions about stress placement. To achieve that goal, we identified the LTB consonant and vowel phonemes (section 2 and 3). From there, we proceeded to the notion of the LTB syllable, identifying unambiguous syllable patterns first before exploring Clayre's notion of the LTB syllable which makes a distinction between nuclear and pre-nuclear syllables (section 4). In the course of section 5, we investigated Clayre's binary and ternary rhymes with respect to ambiguous vowel sequences.

In section 4.1, Table 11, we identified the unambiguous syllable patterns as shown in (19).

- (19) CV for the antepenult
 (C)V for the penult
 (C(:))V for the open ultima
 (C(:))V(:)C for the closed ultima

Furthermore, in section 5.4, we argued for the existence of doubly closed ultimate syllables. Thus, we are getting the syllable patterns shown in Table 19.

Table 19. LTB syllable patterns based on conclusions abovementioned

antepenult	penult	open ultima	closed ultima	doubly closed ultima
CV	(C)V	(C(:))V	(C(:))V(:)C	(C(:))V(:)A?

Table 20 shows that the vowel-approximant interpretations adopted in the previous sections match the unambiguous LTB ultimate rhyme patterns for the closed ultima that is presented in Table 19.

Table 20. LTB rhyme patterns for the closed and doubly closed ultima

unambiguous rhyme pattern	VC	/bura?/	wasteful
matching ambiguous rhyme patterns (with respect to the nucleus)	VA	/buraw/	partially sighted
	VA?	/kulay?/	skin
unambiguous rhyme pattern	V:C	/aka:n/	knowledge
matching ambiguous rhyme pattern	V:A	/laka:w/	to walk
matching ambiguous rhyme pattern (with respect to the nucleus)	V:A?	/ma:w?/	drunk

In this analysis, glottal stops are uniformly treated as consonants. Second, high front and back vowels are uniformly treated as approximants (A) if they appear in the onset or coda of a syllable (e.g. /w/ [ʊ] in /pa[?]way[?]/ ‘wing’, /mu[?]raw[?]/ ‘to make’, and /bu[?]raw/ ‘partially sighted’), but as vowels, if they appear in its nucleus. (e.g. /u/ [ʊ] in /pu[?]/ ‘hair’). Thus, we arrive at the phonological generalization for the LTB syllable as shown in Figure 1.

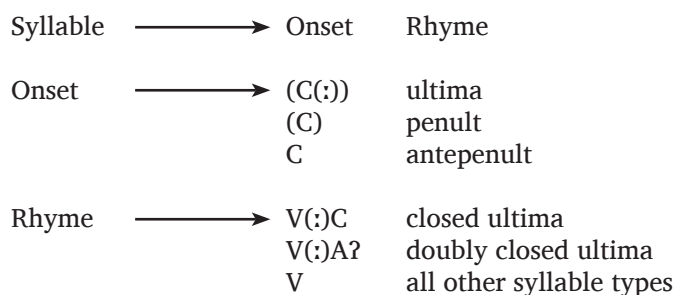


Figure 1. LTB syllable patterns²³

The generalization in Figure 1 makes stress placement in LTB entirely predictable. It always falls on the nucleus of the ultima, which is either a short vowel (V) or a long one (V:). Furthermore, it accurately predicts the range of rhyme patterns that follow from the generalization,²⁴ as shown in Table 21 and Table 22 for the ultima.

Table 21. Range of LTB syllable patterns for the open ultima

Onset	Rhyme ²⁵		
C	V	/niru/	to see
C:	V	/up:ɔ/	news
∅	V	/ti.u/	egg

²³ Principally, the length mark (:) is not needed for the generalization, since it can be deduced that a C position can be occupied either by a short consonant, that is C, or by a long one, C:. The same is true for vowels with respect to the V position. Nevertheless, the length mark has been added here to indicate that only the ultima permits a contrast of short vs. long segments. An alternative approach to the one employed in this paper would be to regard length merely as a phonological element added to simple segments, that is /C/ + /:/ and /V/ + /:/. That approach would limit the number of vowel phonemes to six and the number of consonant phonemes to nineteen.

²⁴ The generalization would predict an occurrence of the pattern C:V:A? for the ultima. However, no language data were found that exhibit this pattern.

²⁵ Word final vowels are always phonetically long.

Table 22. Range of LTB syllable patterns for the closed and doubly closed ultima

Onset	Rhyme					
	Rhyme with V nucleus					
C	VC	/katoh/	always	VA?	/kulay?/	skin
C:	VC	/bit:oh/	rock	VA?	/map:ay?/	to stop on a journey
∅	VC	/bu.əŋ/	beetle	VA?	/ji.əw?/	skewer
	Rhyme with V: nucleus					
C	V:C	/kuman/	to eat	V:A?	/pa:y?/	bitter
		/mite:n/	split		/səra:y?/	bad smell of cooking oil
C:	V:C	/mit:e:n/	stand	V:A?	—	—
∅	V:C	/gi.u:m/	clouds	V:A?	/a:w?/	anus
		/e:ŋ/	waist			

The generalization we arrived at mainly differs from Clayre's notion of the LTB syllable insofar as it

- makes stress placement predictable by narrowing it down to the nucleus of the ultimate syllable.
- makes the onset for the ultimate and penultimate syllable optional whereas Clayre's notion stipulates an obligatory onset.
- rules out the notion of triphthongs, which are principally possible (although not explicitly labeled so) in Clayre's approach.
- offers an interpretation to avoid the notion of diphthong.

To round up our discussion, let us look at the whole range of vowel-approximant sequences that LTB exhibits in the rhyme of the ultimate syllable.

6.1. Vowel-Approximant Sequences

Blust (1992:412) lists -uy, -oy, -ay, -iw, -éw²⁶ and -aw as the "diphthongs" occurring in LTB. The dash in front of them indicates that their position is meant to be word final. In my interpretation adopted above, these are all considered vowel-approximant sequences, just as implied in Blust's transcription.

6.1.1. Vowel-approximant as a V:A# sequence

All of the sequences in (20) occur as V:A# sequences in LTB.

²⁶ Blust's -éw is -ew in my notation.

- (20) /-u:y/ as in /kucu:y/ [kʊcʊ:ɪ] ‘to sit with stretched legs’
 /-o:y/ as in /talo:y/ [tɛlɔ:ɪ] ‘to stab with a spear’
 /-a:y/ as in /ala:y/ [ʔɛlɑ:ɪ] ‘normality’
 /-i:w/ as in /kiki:w/ [kiki:ʊ] ‘to scratch’
 /-e:w/ as in /male:w/ [mɛle:ʊ] ‘to change’
 /-a:w/ as in /mala:w/ [mɛlɑ:ʊ] ‘sky’

As the examples in (20) show, the whole spectrum of long vowels appears before word final approximants in general, just as they occur before unambiguous word final consonants. This is another indication that word final approximants behave like word final consonants. The central vowel /a:/ occurs before both approximants, that is before /y/ and /w/. The back vowels /u:/ and /o:/ occur before /y/, whereas the front vowels /i:/ and /e:/ occur before /w/.

6.1.2. Vowel-approximant as a VA# sequence

Only a limited set of patterns was found, as in (21), that exhibits a short vowel followed by a word final approximant.

- (21) /-uy/ as in /paluy/ [pɛlʊɪ] ‘stupid’ (Brunei Malay loan)
 /-ay/ as in /pəlaway/ [pɛləʊɛɪ] ‘fishing method’
 /-iw/ as in /pariw/ [pɛrɪʊ] ‘to stagger’
 /-aw/ as in /buraw/ [bʊrɛʊ] ‘partially sighted’

6.1.3. Vowel-approximant before word final glottal stop

The only vowels that appear before Aʔ in LTB are schwa as well as the short and long low vowel, as in (22).

- (22) /-əyʔ/ as in /pələyʔ/ [pɛl:əɪʔ] ‘to put on’
 /-ayʔ/ as in /kulayʔ/ [kʊləɪʔ] ‘skin’
 /-a:yʔ/ as in /pa:yʔ/ [pɑ:ɪʔ] ‘bitter’
 /-əwʔ/ as in /kucəwʔ/ [kʊcəʊʔ] ‘to turn around’
 /-awʔ/ as in /murawʔ/ [mʊrɛʊʔ] ‘to do, make’
 /-a:wʔ/ as in /ma:wʔ/ [mɑ:ʊʔ] ‘drunk’

In sum, LTB exhibits the vowel-approximant sequences in the rhyme of the ultima as shown in Table 23.

Table 23. LTB vowel-approximant sequences in the ultimate rhyme

u(:)y	i(:)w
əyʔ	əwʔ
o:y	e:w
a(:)y(?)	a(:)w(?)

7. Conclusion

This paper investigated the notions of diphthongs and syllabicity in Long Terawan Berawan. It finds that Clayre's notion of the ultimate syllable, which is the stressed one, does not make any predictions about stress placement. A reinterpretation was suggested to solve this problem of underspecification. Her diphthongs with the first vowel element stressed are reinterpreted as vowel-approximant sequences. If the second vowel element in the diphthong is stressed, her notion is reinterpreted as a sequence of monophthongs across a syllable break. What appears as a triphthong is reinterpreted as a vowel followed by a syllable break and a vowel-approximant sequence. These reinterpretations eliminate the notion of diphthong and triphthong, which leaves only monophthongs to fill the syllable nucleus. This allows for an accurate prediction of stress placement.

Symbols and Abbreviations

A	approximant
D	diphthong
∅	empty ultimate onset
PAN	Proto-Austronesian
PMP	Proto-Malayo-Polynesian
PPH	Proto-Philippine
[]	encloses phonetic data
//	encloses phonemic data in the author's notation
()	optional item
< >	encloses phonemic data in Clayre's (1996) notation
< < > >	encloses data in orthographic native speaker perception
:	indicates length in phonetic and phonemic data
.	indicates a syllable boundary
*	indicates a proto form

Appendix

Diphthongal interpretation of a vowel-off-glide sequence before final glottal stop

If we choose the diphthongal interpretation as briefly outlined in section 5.4.1, we need to clarify the phonemic/allophonic status of the six diphthong phones this alternative analysis yields: four short diphthongs, [əi], [əu], [ɐi], [ɐu], as in (1),

- | | | | |
|-----|-----------|------------------|-------|
| (1) | [kʊɐɪʔ] | ‘skin’ | CVCDʔ |
| | [pəl:əɪʔ] | ‘to put on’ | CVCDʔ |
| | [mʊɐʊʔ] | ‘to make’ | CVCDʔ |
| | [kʊcəʊʔ] | ‘to turn around’ | CVCDʔ |

and two long ones, [ɑ:i] and [ɑ:u], as in (2).

- | | | | |
|-----|------------|----------------------------|--------|
| (2) | [tən:ɑ:iʔ] | ‘guts’ | CVCD:ʔ |
| | [səɪ:ɑ:iʔ] | ‘bad smell of cooking oil’ | CVCD:ʔ |
| | [pɑ:iʔ] | ‘bitter’ | CD:ʔ |
| | [mɑ:uʔ] | ‘drunk’ | CD:ʔ |

Table (i) shows that all of the six diphthong phones have phonemic status in this alternative analysis since they are in contrast with each other as well as with monophthongs.

Table (i). Contrasts among LTB diphthongs and of diphthongs with monophthongs

Contrast	LTB		English	LTB		English
ai – əi	/sipaiʔ/	[sɪpɐɪʔ]	to reach opposite river bank	/nip:əiʔ/	[nɪp:ɐɪʔ]	being held up
	/lucaiʔ/	[lʊcɐɪʔ]	exit, go out	/pəlacəiʔ/	[pələcɐɪʔ]	to frighten
ai – a	/buraiʔ/	[bʊɐɪʔ]	mottle	/buraʔ/	[bʊɐʔ]	wasteful
ai – i	/map:aiʔ/	[mɐp:ɐɪʔ]	to stop by	/map:iʔ/	[mɐp:iʔ]	thick
əi – e	/pəlacəiʔ/	[pələcɐɪʔ]	to frighten	/laceʔ/	[ləcɛʔ]	to disappear
əi – i	/pəlacəiʔ/	[pələcɐɪʔ]	to frighten	/naciʔ/	[nɛciʔ]	to stick something in
au – əu	/litauʔ/	[lɪtəʊʔ]	murky	/pitəuʔ/	[pɪtəʊʔ]	to hang
	/nucəuʔ/	[nʊcəʊʔ]	to wash	/kucəuʔ/	[kʊcəʊʔ]	turn around
au – a	/məpauʔ/	[mɐp:əʊʔ]	lazy	/məpaʔ/	[mɐp:ɛʔ]	to cut

Contrast	LTB		English	LTB		English
au – u	/litauʔ/	[lɪtəʊʔ]	murky	/lisuʔ/	[lɪsʊʔ]	room
	/nurauʔ/	[nʊrəʊʔ]	made, done	/niruʔ/	[nɪrʊʔ]	to visit
əu – o	/pitəuʔ/	[pɪtəʊʔ]	to hang	/bitoʔ/	[bɪtʊʔ]	neck
əu – u	/ŋajəuʔ/	[ŋɛjəʊʔ]	to tap (on someone's arm)	/tajuʔ/	[tɛjʊʔ]	roof (of boat)
ai – a:	/la:iʔ/	[lɑ:iʔ]	disappointed	/la:ʔ/	[lɑ:ʔ]	loincloth
ai – i:	/pai:iʔ/	[pɑ:iʔ]	bitter	/bi:ʔ/	[bɪ:ʔ]	load (N.)
ai – ai	/pai:iʔ/	[pɑ:iʔ]	bitter	/kapaiʔ/	[kɛpɛiʔ]	near
a:u – a:	/pa:uʔ/	[pɑ:ʊʔ]	to make drunk	/pa:ʔ/	[pɑ:ʔ]	four
a:u – u:	/pa:uʔ/	[pɑ:ʊʔ]	to make drunk	/bu:ʔ/	[bʊ:ʔ]	where
a:u – au	/ma:uʔ/	[mɑ:ʊʔ]	drunk	/gimauʔ/	[gɪmɛʊʔ]	root

Thus, we get the following diphthong phonemes:

4 short diphthongs (/D/), that is /əi/ [ɛɪ], /əu/ [ɛʊ], /ai/ [ɛɪ] and /au/ [ɛʊ]

2 long diphthongs (/D:/), that is /ai/ [ɑ:i] and /a:u/ [ɑ:ʊ]

In this interpretation, I regard the short diphthongs as equivalent to a V segment and the long ones as equivalent to a V: segment.

In the diphthongal interpretation, the ambiguous LTB ultimate rhyme patterns also match the unambiguous ones as shown in Table (ii).

Table (ii). LTB rhyme patterns for the closed ultima with the diphthongal interpretation applied

unambiguous rhyme pattern	VC	/buraʔ/	wasteful
matching ambiguous rhyme patterns	VA	/buraw/	partially sighted
	DC	/kutəiʔ/	to explode
		/kucəuʔ/	to reverse
		/kulaiʔ/	skin
		/murauʔ/	to make
unambiguous rhyme pattern	V:C	/akɑ:n/	knowledge
matching ambiguous rhyme patterns	V:A	/laka:w/	to walk
	D:C	/pai:iʔ/	bitter
		/ma:uʔ/	drunk

In contrast to the VA? analysis, the glides [ɪ] and [ʊ] are not treated uniformly as approximants here, but only as such if they occur in the onset of a syllable or in word final position, as in:

- (3) /paya:w/ [pɛɪɑ:ʊ] ‘sambar deer’
 /biwih/ [biʊih] ‘pig’
 /ata:y/ [ʔɛtɑ:ɪ] ‘liver’

If [ɪ] or [ʊ] occurs in the ultima between a central vowel and a glottal stop (Vʔ), it is considered an off-glide to a base vowel, constituting a diphthong, D(:), with the latter, as in (4).

- (4) /kulaiʔ/ [kʊlɛɪʔ] ‘skin’ (D)
 /paiʔ/ [pa:iʔ] ‘bitter’ (D:)

The length of a short LTB diphthong (D) is perceptually equivalent to the length of an ultimate short vowel nucleus (V). The same is true for long diphthongs (D:) in comparison with long vowels (V:). While this analysis introduces complex nuclei, i.e. diphthongs, it avoids the notion of complex codas, i.e. doubly closed syllables.

If we adopt the diphthongal analysis, we get the generalization about the LTB syllable shown in Figure (i).

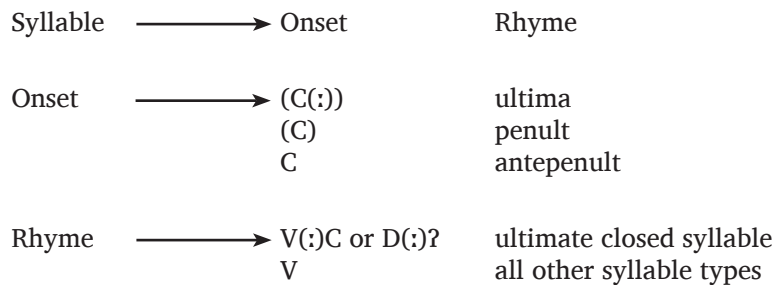


Figure (i). The LTB syllable – diphthongal interpretation

This generalization also correctly predicts the range of ultimate rhyme patterns in LTB (except for the non-occurrence of C:D:C) as shown in Table (iii).

Table (iii): Range of LTB syllable patterns for the closed ultima with the diphthongal interpretation applied

onset		rhyme									
		rhyme with V or D nucleus									
C	VC	/katoh/	always	VA	/buraw/	partially sighted	DC	/kulaiʔ/	skin		
C:	VC	/bitroh/	rock	VA	/ŋaraw/	to disturb	DC	/mapaiʔ/	to stop on a journey		
∅	VC	/bu.əŋ/	beetle	VA	—	—	DC	/ji.əuʔ/	skewer		
		rhyme with V: or D: nucleus									
C	V:C	/kumain/	to eat	V:A	/laka:rw/	to walk	D:C	/pai:ʔ/	bitter		
		/mite:n/	split	V:A	/siku:y/	water-melon	D:C	/sərai:ʔ/	bad smell of cooking oil		
C:	V:C	/mit:e:n/	stand	V:A	/man:ay/	clever	D:C	—	—		
∅	V:C	/gi.u:m/	clouds	V:A	/bu.a:rw/	to migrate	D:C	/au:ʔ/	anus		
		/e:ŋ/	waist	V:A							

References

- Anon. 1991. *Kamus Bahasa Melayu Brunei*. Bandar Seri Begawan: Dewan Bahasa dan Pustaka Brunei.
- Anon. 1995. *Kadazan Dusun-Malay-English dictionary*. Kota Kinabalu: Kadazan Dusun Cultural Association, Sabah.
- Anon. 2005. *Reproduction of The International Phonetic Alphabet*. <http://www.arts.gla.ac.uk/ipa/vowels.html> (accessed 24 September 2007).
- Asmah Haji Omar. 1983. *The Malay peoples of Malaysia and their languages*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Blust, Robert. 1974. *The Proto-North Sarawak vowel deletion hypothesis*. Unpublished Ph.D. dissertation, Honolulu, USA: Department of Linguistics, University of Hawai'i at Manoa.
- Blust, Robert. 1992. The long consonants of Long Terawan. *Bijdragen tot de Taal-, Land- en Volkenkunde* 148: 409–427.
- Blust, Robert. 1995. Notes on Berawan consonant gemination. *Oceanic Linguistics* 34: 123–138.
- Blust Robert. 2000. Low vowel fronting in Northern Sarawak. *Oceanic Linguistics* 39:285–319.
- Clayre, Beatrice. 1994. *Verb affixes in Berawan*. Paper presented at the Seventh International Conference on Austronesian Linguistics. Leiden 22–27 August 1994.
- Clayre, Beatrice with Denny Belawing Wan. 1996. An introduction to Berawan phonology. *Sarawak Museum Journal* 71: 209–240.
- García-Bellido, Paloma and Beatrice Clayre. 1997. Prosodic constraints and representations in the Berawan word. *Oxford Working Papers in Linguistics, Phonology & Phonetics*:17–52. Oxford: Department of Linguistics, Oxford University.
- Wurm, S. A. and B. Wilson.1983. *English finderlist of reconstructions in Austronesian languages*. Pacific Linguistics Series C:No. 33.

Result Complex Predicates in Kimaragang Dusun*

Paul Kroeger
SIL and GIAL, Dallas

This paper discusses a peculiar construction in Kimaragang Dusun in which two verbs occur within the same clause. One use of this construction is to describe the result of an action. The initial verb, which names the result, is the independent form; it may be inflected for tense, aspect, imperative mood, etc., and typically appears in a non-active voice. The second verb, which names the action, appears in an invariant non-past, active voice form. The construction is analyzed as a complex predicate, meaning that the two verbs share a single argument structure. Evidence for monoclausality comes from the pattern of case assignment and the distribution of 2P clitics, while evidence for shared argument structure comes from uniqueness effects and interaction with causative formation.

1. Introduction

This paper discusses a somewhat surprising construction in Kimaragang Dusun, a Philippine-type language spoken in northeastern Borneo. Although the construction contains two verbs, I argue that it is monoclausal and that the two verbs together share a single, complex argument structure. For this reason I refer to the construction as a COMPLEX PREDICATE. Two primary types, or uses, of the complex predicate construction have (thus far) been identified, which I refer to as RESULT and MANNER complex predicates (RCPs and MCPs, respectively). These two types differ in their semantic relations, but to this point I have found no clear evidence of a difference in syntactic structure.

Some examples of the result type, the main focus of this paper, are presented in (1).¹ As these examples illustrate, the first verb in the construction (V_1) is fully inflected for voice, tense-aspect, mood and modality, while the second verb (V_2) is invariant, always appearing in the active voice, non-past tense, and non-imperative mood. Notice that the patient (or Undergoer) argument is marked for nominative case even when it follows V_2 . This is surprising, because (as discussed in the next section) the active voice form of the verb cannot normally select the patient as subject. We would expect instead that the active voice verb would assign accusative case to its

* I would like to thank Jim Johansson and Janama Lantubon for providing many of the examples cited in this paper, and for helpful discussions of the issues it raises. Thanks also to Maria Polinsky for many helpful comments and suggestions.

¹ Note that in the orthography used here, word-final glottal stop is not marked while a word-final *-h* indicates the lack of a final glottal stop, i.e., an open syllable. However, I have chosen (somewhat inconsistently) to mark final glottals with an apostrophe in certain grammatical morphemes where they would be pronounced by speakers of all dialects.

patient; but this marking is not possible here, as illustrated in (1f). It is this unexpected case-marking pattern in particular which makes the construction seem so odd.²

- (1) a. N-a-rasak do karabau monginum(m-poN-inum)³ at weeg.
 past-nvol-dry.up GEN buffalo av-tr1-drink nom water

‘The stream was drunk dry by buffaloes.’

- b. P[in]i-toning-Ø kuh it sapi om karabau Ø-po-ogot.
 RECIP[PAST]-near-OV 1sg.GEN NOM cow and buffalo AV-TR2-tie

‘I tied up the cow and the buffalo near each other.’

- c. Adan-o’ yalo mamasut(m-poN-pasut)!
 faint-OV.IMPER 3sg.NOM AV-TR1-whip

‘Whip him unconscious!’

- d. N-a-dapit-Ø do tulun momokok(m-poN-wokok) at bawang.
 PAST-NVOL-span-OV GEN person AV-TR1-dam NOM river

‘Someone dammed up the river all the way across.’

- e. Tuus-an noh momo’og(m-poN-wo’og) inoh tunturu nuh!
 bare-DV FOC AV-TR1-wash that(NOM) finger 2sg.GEN

‘Wash your fingers totally clean!’

- f. *Tuus-an noh momo’og(m-poN-wo’og) dinoh tunturu nuh!
 bare-DV FOC AV-TR1-wash that(ACC) finger 2sg.GEN

(ungrammatical with the intended meaning; marginally acceptable with a different meaning: ‘Wash it (unspecified) totally clean with your fingers!’)

The oddness of the case-marking pattern is even more evident in the manner complex predicate (MCP); some examples are given in (2). As these examples show, the “pivot” (or grammatical subject) of the MCP construction is typically not a semantic argument of V_1 . Rather, V_1 seems to be predicated of the action as a whole.⁴ If the second verb were not present, all seven of these sentences would be ungrammatical. For example, the root *basag* ‘strong’ in (2c) is only used for animate beings. The verb form *basagon* cannot be used for strengthening a rope (or any other inanimate thing), but describes how the action

² Although the construction seems odd, it is not uncommon. Examples of RCPs are fairly easy to find, and MCPs seem to be even more common. See Holmer (2004) and Chang (in press) for examples of similar constructions in several Formosan languages.

³ The “transitivity” prefixes, glossed “TR1” and “TR2”, are discussed in the last half of section 4.

⁴ In this sense, the assignment of semantic roles in the MCP is somewhat similar to a Raising construction (e.g., *John seems to enjoy his work*). Syntactically, however, I argue that the structure is monoclausal, whereas a Raising construction by definition must be biclausal.

is done ('strongly'). The root *bandan* 'big' in (2d) cannot be predicated of an object (e.g. a chicken), and chopping something cannot make it bigger. Nevertheless, it is V_1 and not V_2 that determines the nominative case marking on the Undergoer subject.

- (2) a. *Induwa-an nopoh momo'og(m-poN-wo'og) inoh wagas toboh.*
 twice-DV only AV-TR1-wash that(NOM) rice PRTCL

'Just wash that (uncooked) rice two times.'

- b. *N-enggotus-an dialo miguguli(m-pi-gu-guli)*
 PAST-hundred.times-DV 3sg AV-RECIP-DUP-return

mongumpug(m-poN-umpug) i niyuw di lapak-on yoh.
 AV-TR1-gather NOM coconut REL split-OV 2sg.GEN

'He went back a hundred times to gather the coconuts that he was splitting.'

- c. *Basag-on noh mongogodong(m-poN-godong) ilo tali ong*
 strong-OV FOC AV-TR1-pull that(NOM) rope if

tarik.tali kou ki.
 tug.of.war 2pl.NOM PRTCL

'Pull hard on the rope when you enter the tug-of-war, okay?'

- d. *Bandan-o = i' momurok(m-poN-purok) inoh manuk toh.*
 big-OV = PRTCL AV-TR1-cut.up that(NOM) chicken PRTCL

'Chop up that chicken into big pieces.'

- e. *Tuyuan-ai Ø-po-wiliw inoh rungou, a-babak dati.*
 careful-DV AV-TR2-lay that(NOM) jar NVOL-shatter likely

'Put that jar down gently or it might break.'

- f. *G[in]ibang-Ø kuh yalo manampar(m-poN-tampar).*
 [PAST]-left-OV 1s.GEN 3sg.NOM AV-TR1-punch

'I hit him with my left hand.'

- g. *K[in]ondiri-Ø dialo mamatai(m-poN-patai) it tasu yoh.*
 [PAST]-self-OV 3sg AV-TR1-kill NOM dog 3sg.GEN

'He killed his dog himself.'

So the "pivot" (i.e. subject) is typically a semantic argument of both verbs in the RCP (a more precise statement is formulated in section 4), but only of the first verb in the MCP. The syntactic properties of the two constructions, however, seem

to be identical, as noted above. The remainder of this paper focuses primarily on the result type. Section 2 provides some basic information about Kimaragang morphology and clause structure. Section 3 presents evidence for the monoclausal nature of this construction, and illustrates some structural differences between the complex predicate construction and a superficially similar biclausal construction. Section 4 discusses the semantic constraints on argument sharing within the RCP construction. Section 5 formulates a lexical rule of RCP formation, and provides supporting evidence for this analysis involving the interpretation of examples in which V_1 is a morphological causative. Section 6 discusses an alternative possible analysis which essentially treats V_2 as a kind of gerund functioning as a causal adjunct.

2. Basic clause structure

In Kimaragang, as in most Philippine-type languages, any argument of the verb can be selected as subject. A definite Undergoer will normally be selected as subject unless some other argument of the clause is extracted. Non-subject arguments carry semantically determined case marking that reflects their thematic role, but subjects always carry nominative case. The case-marking particles of Kimaragang are summarized below:⁵

(3)		NOM	GEN/ACC	DAT/LOC
	Definite	i(t)	di(t)	sid
	Indefinite	o(t)	do(t)	sid
	Unique ref.	a(t)	da(t)	ad

The thematic relationship of the subject argument to the predicate is signaled by a voice- marking affix on the verb. The use of the voice markers in Kimaragang is illustrated in the following examples. In each sentence the subject is italicized.

- (4) a. *Yokuh* ot minonunguw(m[in]-poN-tunguw) dinoh tinorimo
 1sg.EMPH NOM AV[PAST]-TR1-pour that.ACC cooked.rice

dinoh ...
 that

'I was the one who added water to the rice (being cooked)...'

- b. Tungu-on it sada ki-owoh ...
 pour-OV NOM fish PRTCL

'Add water to *the fish*, okay?' (when cooking; to make gravy)

⁵ The genitive and accusative cases are distinguished only in pronominal forms, but for ease of exposition I will gloss non-pronominal NP's as bearing ACC or GEN case depending on which form a pronominal argument would take in that same position. Another possible analysis would be to say that non-pronominal objects take genitive case, while pronominal objects take accusative case. This pattern finds parallels in other Philippine-type languages. For example, in Tagalog definite animate objects take dative case, while indefinite and most inanimate objects take genitive case.

c. N-i-tunguw kuh it weeg di sada sid
 PAST-IV-pour 1sg.GEN NOM water GEN fish DAT

poonumadan([V]-poN-sumad-an) do tasu.
 DUP-TR1-feed-DV GEN dog

‘I poured *the water* from (cleaning) the fish into the dog’s feeding dish.’

d. Tungu-ai poh do tinasak ilot lampu kuh.
 pour-DV.IMPER yet ACC oil that.NOM lamp 1sg.GEN

‘Fill *my lamp* with oil.’

Instrumental Voice (*i-*) is used when the subject is either an instrument, or (as in 4c) a displaced theme. Dative Voice (*-an*) normally indicates that the subject is a goal (as in 4d), a recipient, or a beneficiary; other uses with specific classes of verbs are illustrated in Kroeger and Johansson (2005). Active Voice (*m- ~ -um-*) signals that the subject is the argument that is highest on the thematic hierarchy (agent > experiencer > ...). Objective Voice (*-on*) normally selects the patient as subject. A fifth voice category, not illustrated in (4), is Locative Voice, which is homophonous with the Objective Voice except in the past tense (see ex. 7b). Locative Voice is used primarily when the subject of the sentence is the location or goal of an intransitive verb of motion, posture or position; and with verbs of infestation, infection, etc.

The Kimaragang TAM affixes are summarized in (5). Kimaragang exhibits a simple two-tense system, past vs. non-past. The past tense marker is the infix *-in-*, inserted after the initial consonant of the base form. Before vowel-initial forms, this infix reduces to a prefixed *n-*. Verb forms which lack this tense marker are interpreted as non-past. In addition, three of the voice markers have a distinct atemporal form which is used for imperatives, for the main verb when an auxiliary is present, and for main-line narrative events as illustrated in (6), taken from a folk-tale. Finally, a distinct non-volitive form (glossed “NVOL”) is used to express ability, involuntary actions, result states, and indefinite time reference; see Kroeger (1990), Dell (1983). Note that aspectual uses of reduplication are not included in (5). These include imperfective aspect (i.e., progressive and habitual), repetitive, distributed action, etc.

(5) TAM affixes:

Voice Category	Non-past	Past	Imperative/ atemporal	Non-volitive
Actor (AV)	m- / -um-	m-in- / -in-um-	∅-	(no)ko-
Objective (OV)	-on	-in-_-∅	-o’	(n)o-
Dative (DV)	-an	-in-_-an	-ai	(n)o- -an
Instrument (IV)	i-	n-i-	---	(no)ko-
Locative (LV)	-on	-in-_-on	---	---

- (6) *Ongoy-o'* noh di tasi-asi om *patay-o'* noh om
 fetch-OV.ATMP FOC GEN orphan and kill-OV.ATMP FOC and
- tunuw-ai* noh.
 roast-DV.ATMP FOC

‘The orphan boy fetched (the lizard) and killed it and roasted it.’

In basic verbal clauses the verb always comes first, pronouns almost always precede full NPs, and NP subjects tend to occur in clause-final position. The position of pronominal elements is fairly strictly determined by various constraints which will be discussed immediately below. The relative order of full NPs, on the other hand, is relatively free. There is a general tendency for NPs to precede PPs, and for genitive NPs to precede dative NPs. When the verb is marked for actor voice, the nominative Actor NP may either occur in final position or immediately after the verb. In other voices, the Actor NP normally precedes all other non-pronominal elements of the clause. But non-human Actors, as well as inanimate effectors, may also occur after the subject.

Genitive pronouns for all persons, and first and second person nominative pronouns, are second-position (or “2P”) elements; that is, they must always follow the first constituent in their clause. (Third person nominative pronouns optionally occupy this position.) In a normal verb-initial clause, this means following the verb. For example, (1b), (2f) and (4c) all contain a genitive Actor pronoun immediately following the verb.

When a negative or other adverbial element is fronted to pre-verbal position, 2P clitics will also precede the verb; this is exemplified in (7). In addition to pronouns, a variety of other particles also occur in this position, as seen in (7c). (See also (1e) and (2c), where the focus particle *noh* occupies the 2P position.) Sentence-level conjunctions do not function as a part of the minimal clause, and so do not affect clitic placement.

- (7) a. Suwab-suwab *okuh* manalu(m-poN-salu) do pulut.
 every.day 1sg.NOM AV-TR1-tap GEN rubber
- ‘Every day I tap rubber.’
- b. Sid tana *yah* n-odop-on.
 DAT earth 1pl.ex.GEN PAST-sleep-LV
- ‘We slept on the ground (after the house burned down).’
- c. Amu *okuh* *poh* *dati* ko-guli dot suwab
 NEG 1sg.NOM yet probably NVOL.AV-return LNK tomorrow
- sitih k[um]araja.
 here [AV]work

‘Tomorrow I probably cannot return to work here.’

3. Evidence for monoclausality

The most obvious kind of evidence that the two verbs in the complex predicate construction belong to the same clause is that they cannot be separated by a pause, conjunction, complementizer, linker, or any other marker of clause boundaries. Further evidence comes from the placement of second-position clitics. In a subordinate clause, whether complement or adjunct, clitic pronouns and particles appear immediately after the first element of their minimal clause; this clearly indicates the location of sentence-internal clause boundaries. Example (8a), for example, contains three internal clause boundaries as indicated by the location of the italicized clitic elements. In the complex predicate construction, however, there is no medial position that can host such clitics; note the position of the two clitic arguments in (8b), immediately following the fronted NEG. This is especially striking in the nominative subject pronoun, since non-pronominal subject NPs normally follow V_2 in this construction. Thus the position of the 2P clitics in examples like (8b) provides strong evidence that the RCP is monoclausal.

- (8) a. Ela'an *kuh* = i' dot magaago *yalo* nga'
 know 1sg.GEN = EMPH COMP hurry(AV) 3sg.NOM but

n-antara-Ø *kuh* tu' waro b[in]oros *kuh* sid dialo.
 PAST-intercept-OV 1sg.GEN because EXIST [NMLZ]say 1sg.GEN DAT 3sg

'I knew that he was in a hurry but I held him up because I had something to say to him.'

- b. Amu *kuh* *yalo* n-o-onong-Ø monimbak(m-poN-timbak).
 not 1sg.GEN 3sg.NOM PAST-NVOL-hit-OV AV-TR1-shoot

'I didn't hit him when I shot.'

Additional evidence for the monoclausal status of the complex predicate construction comes from the pattern of voice marking and case assignment. As discussed above, the case marking of an NP argument depends in part on the voice marking of the verb which selects it. In the complex predicate construction, V_2 always appears in the active voice. V_1 typically appears in a non-active voice, and it is this verb which determines the case assignment for both Actor (GEN) and Undergoer (NOM). The nominative case marking of the Undergoer, even when it follows V_2 as in (1a,d,e), (2a-e) etc., is quite surprising, and is one of the main defining features of this construction. It seems to show that, whatever the semantic relations involved, the nominative NP is always a syntactic argument (and therefore a clausemate) of V_1 . This conclusion is supported by the variable position of the nominative NP, which normally follows V_2 but often precedes V_2 when the Actor is a pronoun (see 1b, 14b, etc.). It is further confirmed by the fact

that the nominative NP can be topicalized (9a-b), clefted, or questioned (9c), in which case it appears before V_1 .⁶

- (9) a. It sapi om karabau p[in]i-toning-Ø kuh Ø-po-ogot.
 NOM cow and buffalo RECIPIENT[PAST]-near-OV 1sg.GEN AV-TR2-tie

‘The cow and the buffalo I tied up near each other.’

- b. It togilai yah nga’ n-a-awi-Ø do kara
 NOM maize 1pl.ex.GEN TOP PAST-NVOL-finish-OV GEN monkey

mangkan(m-poN-akan), it parai yah nga’
 AV-TR1-eat NOM rice 1pl.ex.GEN TOP

n-o-tuus-Ø do karabau mongotop(m-poN-kotop).
 PAST-NVOL-bare-OV GEN buffalo AV-TR1-graze

‘Our maize was all eaten by monkeys, and our rice crop was completely devoured by buffaloes.’

- c. Disai do tasu ot n-a-patai-Ø dialo momobog(m-poN-bobog)?
 whose LNK dog NOM PAST-NVOL-kill-OV 3sg AV-TR1-beat

‘Whose dog did he beat to death?’

In a biclausal structure, each verb determines the case marking of its own arguments. In most biclausal structures, moreover, each verb can be marked for tense, and may (in principle) take the full range of voice markers. The biclausal result-reason sentence pattern is illustrated in (10). Notice that both verbs are marked for past tense, each verb is followed by its own arguments, and each verb assigns nominative case to its own subject (omitted under identity in the second clause of 10b).

- (10) a. [N-a-andab i = togilai] tu’ [minonutud(m-[in]-poN-tutud) okuh
 PAST-NVOL-wilt NOM = corn because AV-PAST-TR1-burn 1sg.NOM

di = sakot id = tompil].
 ACC = grass DAT = side

‘The corn wilted (from heat) because I burned some grass right next to it.’

- b. [N-a-babak ilo kasa,] tu’ [noko-dumpau
 PAST-NVOL-shatter that(NOM) bottle because PAST.NVOL.IV-throw.down

⁶ These examples cannot be instances of long-distance extraction, because the voice category of V_2 would not permit it. The topicalization in (9a) is indicated by pre-verbal position. The TOPIC particle *nga’* in 9b is used primarily for contrastive topics.

sid pampang].
 DAT stone

‘That bottle broke because it was thrown against a rock.’

The conjunction *tu* ‘because’ is frequently omitted in such constructions when the subject of the second clause is co-referential with the subject of the first, as in (10b). The result is a biclausal sentence pattern that looks superficially similar to the resultative (RCP) construction; some examples are provided in (11). This biclausal pattern is distinguished from the RCP by a number of diagnostic features including the following: the second verb may be inflected for tense and may occur in non-active voices; either clause may contain an auxiliary verb, as in the second clause of (11c), whereas in the RCP the auxiliary is only possible for V_1 (see ex. 19a–b); an optional pause may occur between the two clauses, as in (11d–e); each verb precedes and assigns case to its own arguments; and the overt subject NP normally appears as an argument of the first clause, and thus follows the first verb, whereas in the RCP the subject NP normally follows V_2 (unless it is a pronominal clitic).

- (11) a. Noko-pilai yalo s[in]ungu-Ø do karabau.
 PAST.NVOL.IV-throw 3sg.NOM [PAST]-horn-OV GEN buffalo

‘He was gored by a buffalo and tossed (into the air).’

- b. N-a-alum-Ø nopoh ilo parai n-ajang-an do sarup
 PAST-NVOL-flatten-OV only that(NOM) rice PAST-affect-DV GEN wind

tologod.
 strong

‘That rice was hit by a strong wind and flattened.’

- c. N-o-wutus-Ø noh it parai minaan da manuk
 PAST-NVOL-pull.out-OV FOC NOM rice AUX.PAST GEN chicken

ka-kakay-o’.
 REDUP.Scratch-OV.ATMP

‘My rice plants were uprooted because the chickens were scratching around them.’

- d. N-o-dompol i tandus kuh, n-i-sungkad dialo sid pampang.
 PAST-NVOL-blunt NOM spear 1sg.GEN PAST-IV-poke 3sg DAT stone

‘My spear was blunted because he poked it against a stone.’

- e. N-o-oyas it walai dialo, n-a-aba-an dit kayu.
 PAST-NVOL-fall.apart NOM house 3sg PAST-NVOL-fall.over-DV GEN tree

‘His house was fallen on by a tree and got smashed to pieces.’

Another important difference between the two patterns is that the conjunction *tu* ‘because’ can never occur between the two verbs of the RCP construction, even though the action named by V_2 is clearly the cause of the result named by V_1 . This fact is illustrated in (12).

- (12) a. N-a-rasak do karabau monginum(m-poN-inum) at weeg.
 PAST-NVOL-dry.up GEN buffalo AV-TR1-drink NOM water

‘The stream was drunk dry by buffaloes.’

- b. *N-a-rasak do karabau tu’ monginum at weeg.
 PAST-NVOL-dry.up GEN buffalo because AV-TR1-drink NOM water

(could only mean: ‘(Something) was dried up by buffaloes because the stream was drinking.’)

- c. N-a-awi-Ø do kara mangakan(m-poN-akan) it togilai
 PAST-NVOL-finish-OV GEN monkey AV-TR1-eat NOM maize

yah.

1pl.ex.GEN

‘Monkeys ate up all of our maize.’ (lit: ‘... finished the maize eating.’)

- d. *N-a-awi-Ø do kara tu’ mangakan it togilai
 PAST-NVOL-finish-OV GEN monkey because AV-TR1-eat NOM maize

yah.

1pl.ex.GEN

(could only mean: ‘Monkeys finished (something) off because our maize was eating.’)

It is generally possible to express the meaning of a result complex predicate like that in (13a) using a biclausal structure, as illustrated in (13b). However, paraphrase in the opposite direction is not always possible. The RCP construction is subject to a number of semantic restrictions which do not apply to the biclausal structure. Some of these will be discussed in the following sections.

- (13) a. N-o-tuus-Ø do karabau mongotop(m-poN-kotop) i parai
 PAST-NVOL-bare-OV GEN buffalo AV-TR1-graze NOM rice

kuh.
1sg.GEN

‘My rice crop was completely devoured by buffaloes.’

- b. N-o-tuus-Ø i parai kuh k[in]otop-Ø do karabaw.
PAST-NVOL-bare-OV NOM rice 1sg.GEN [PAST]-graze-OV GEN buffalo

‘My rice crop was all eaten up, grazed off by buffaloes.’

4. Semantic features of the RCP

In the result complex predicate (RCP), the inflected verb (V_1) which functions as the syntactic head names the result, while the “dependent” verb (V_2) names the activity which brings about this result. In the examples in (1), V_1 names a specific result state; some further examples of this pattern are given in (14). However, in other cases V_2 expresses the extent of the action, as illustrated in (15); see also (12c).

- (14) a. Pi-apat-on nuh mangalapak(m-poN-lapak) inoh niyuw.
RECIP-four-OV 2sg.GEN AV-TR1-split that(NOM) coconut

‘Split that coconut into four parts.’

- b. P[in]i-so-sokot-Ø dialo ah babas monombir(m-poN-tombir).
RECIP[PAST]-REDUP-join-OV 3sg NOM cloth.scrap AV-TR1-sew

‘She sewed the scraps of cloth together (lit: joined the scraps sewing).’

- c. Irot-on noh Ø-po-otub i pati ki.
tight-OV FOC AV-TR2-close.lid NOM box okay?

‘Close the box tight, okay?’

- d. Pa-rasak-ai mangansak(m-poN-ansak) inoh rinapa ki.
CAUS-dry.up-DV.IMPER AV-TR1-cook that(NOM) viand okay?

‘Cook that food until it is dry, all right?’

- (15) a. N-o-rikot kuh momilai(m-poN-pilai) i walai nuh sid
PAST-NVOL-arrive 1s.GEN AV-TR1-throw NOM house your DAT

tupak.
other.side

‘I threw (something) and hit your house all the way across the river.’

b. Amu poh n-o-pupus-an di Odu manangon(m-poN-tangon)
 not yet PAST-NVOL-end-DV GEN Granny AV-TR1-folk.tale

i Majabou.
 NOM Majabou

‘Granny has not finished telling the story of Majabou.’

The action named by V_2 is in some sense presupposed, and this presupposition is preserved when the clause is negated. The negated examples in (16) deny that the result named by V_1 was achieved, without denying that the action named by V_2 was performed. In this respect, the Kimaragang RCP is similar to the English resultative construction, as illustrated in (17).

(16) a. Amu kuh n-o-onong-Ø monimbak(m-poN-timbak) it tambang
 not 1sg.GEN PAST-NVOL-hit-OV AV-TR1-shoot NOM deer

(tu’ naka-alai nogi.)
 because PAST.NVOL.AV-avoid again

‘I didn’t hit the deer when I shot (because my bullet went astray/missed the target).’
 (does not imply ‘I did not shoot.’)

b. Amu kuh n-a-patai-Ø momobog(m-poN-bobog) ih wulanut.
 not 1sg.GEN PAST-NVOL-kill-OV AV-TR1-beat NOM snake

‘I didn’t beat the snake to death (i.e., I beat the snake but it didn’t die).’

c. Amu kuh n-o-togu-Ø yalo monobok(m-poN-tobok)
 not 1sg.GEN PAST-NVOL-pierce-OV 3sg.NOM AV-TR1-stab

tu’ kikobol yalo.
 because invulnerable 3sg.NOM

‘I didn’t/wasn’t able to pierce him when I stabbed him because he has *kebal* protection (i.e., has been made invulnerable by performing certain rituals).’

(17) a. Maestro: You went to a football match??!

Opera star: Don’t worry, I didn’t yell myself *hoarse*.
 (does not imply ‘I didn’t yell.’)

b. I put in eight hours a day, but I’m not going to work myself *to death*.

c. I spanked the child, but I certainly didn’t beat him *black and blue*.

We turn now to the pattern of argument sharing in the RCP. We noted in the introduction that in manner complex predicates (MCPs), the “pivot” (or grammatical subject) of the construction is typically not a semantic argument of V_1 . This is not the case for the RCP. The subject of the RCP must be the Undergoer (i.e., affected argument) of the first verb, and is frequently the Undergoer of the second verb as well.

The two verbs in the RCP express a single complex transitive event which involves a unique Actor and a unique Undergoer. If either A or U is an argument of both verbs, as is frequently the case, each verb must select the same A and/or U. But it is not always necessary for both of these arguments to be shared, because either verb may be inherently intransitive. The Undergoer must always be an argument of V_1 and the Actor must always be an argument of V_2 . This implies that V_1 must be either transitive or unaccusative, while V_2 must be either transitive or unergative. Examples (1a) and (15a) illustrate unaccusative verbs occurring as V_1 ; the examples in (20) illustrate unergative verbs occurring as V_2 .

- (20) a. Amu o-owit-Ø dit tombolog t[um]ulud it wulanut.
not NVOL-lift-OV GEN bird [AV]fly NOM snake

‘The bird was not able to fly off carrying the snake.’

- b. N-a-dansaran-an dati m-ogom ah takod da tanak
PAST-NVOL-sit.on-DV likely AV-sit NOM foot GEN child

sagai m[in]og-iad noh.
therefore AV[PAST]-pog-weep already

‘Someone probably sat down squarely on the child’s foot, that is why he started crying.’

It is possible for V_2 to introduce a new argument, provided that argument is not its Undergoer. In example (21), V_2 introduces an instrument (the spear) which is not an argument of V_1 .

- (21) N-o-onong-Ø dialo monokon(m-poN-tokon) do tandus i
PAST-NVOL-hit-OV 3sg AV-TR1-hurl ACC spear NOM

koridai.
barking.deer

‘He hit the deer with a spear.’ (lit: ‘He hurling a spear hit the deer.’)

The pair of sentences in (22) illustrates the uniqueness constraint for Actors. Example (22a) is a biclausal sentence like those discussed in the previous section. However, there is no way to paraphrase this example using the result complex predicate construction. The attempted paraphrase in (22b) is ungrammatical because the two verbs have distinct Actors. The root *odop* ‘sleep’ is an unergative verb, as seen in the use of *noko-* rather than *no-* for the non-volitive form; its only argument (‘child’) is an

which the rice plants themselves are cut down. The semitransitive form, however, does not grammatically specify an Undergoer, and so the correct reading is allowed.¹⁰

(24) a. Amu okuh ka-waya dikau suwab tu'
not 1sg.NOM AV.NVOL-accompany 2sg.ACC tomorrow because

g[um]amas / #mangagamas(m-poN-gamas) okuh do parai.
[AV]-cut.grass AV-TR1-cut.grass 1sg.NOM ACC rice

'I can't go with you tomorrow because I have to clear away the grass around my rice plants (#cut down my rice plants).'

b. Amu nuh n-a-awi-Ø g[um]amas / #mangagamas i
not 2sg.GEN PAST-NVOL-finish-OV [AV]-cut.grass AV-TR1-cut.grass NOM

parai.
rice

'You didn't finish cutting grass around the rice plants (#cutting down the rice plants).'

The examples in (24) illustrate an aspect of Kimaragang morphology that provides an excellent way to test the Single Undergoer constraint. When the Undergoer of a transitive verb is not selected as subject, the verb bears one of two "transitivity" prefixes, glossed here as TR1 and TR2. These prefixes provide partial information about the semantic role of the Undergoer: TR1 is used when the Undergoer is a patient, goal or recipient, while TR2 is used when the Undergoer is an instrument or displaced theme (see Kroeger 1996, Kroeger and Johansson 2005 for details).

We can illustrate this pattern with the verb 'give'. The act of giving involves three participants: an Actor (the giver), a theme (the gift), and a recipient. When the TR1 prefix (*poN-*) is used, the recipient is selected as Undergoer. The action is viewed as primarily affecting, or being directed towards, the recipient, and must involve a change of ownership. When the TR2 prefix (*po-*) is used, the Undergoer is the theme. In this case there need not be any change of ownership, but there must be a physical transfer of possession. This semantic contrast is illustrated in (25). The noun *tana* is ambiguous between the meanings 'land' and 'dirt'. Thus example (25a) could mean either 'I will give you some land' or 'I will give you some dirt'; but the former meaning is more likely, since the *poN-* form implies change of ownership and dirt is seldom given as a gift. However, the *po-* form in (25b) implies a physical transfer of possession. Since a piece of land cannot be physically moved (at least, not by human agency), example (25b) can only mean 'I will give you some dirt'.

(25) a. Mana'ak(m-poN-ta'ak) okuh dikau do tana.
AV-TR1-give 1sg.NOM 2sg.DAT GEN earth

'I will give you some land/dirt.'

¹⁰ Kroeger (1996) argues that in a technical sense, semitransitive forms have no Undergoer.

- b. Ø-pa-ta'ak okuh dikau do tana.
 AV-TR2-give 1sg.NOM 2sg.DAT GEN earth

'I will hand you some dirt (*land).'

Since the two “transitivity” prefixes occur only in verb forms whose Undergoer is distinct from the subject, a verb that contains one of these prefixes can never occur as V_1 of the RCP. This is because subject of the RCP must be the Undergoer of V_1 .

The contrast between TR_1 and TR_2 provides a valuable probe for testing the structure of the resultative construction. Since the second verb in the RCP must appear in active voice, it will (if it is transitive) bear one of these two prefixes. If the transitivity prefix on V_2 selects an argument other than the Undergoer of V_1 , the construction is ungrammatical. To illustrate, the RCP examples in (26b–c) are both intended to be equivalent to the biclausal sentence in (26a). In all three sentences, the Undergoer of the first verb (*naawi*) is the theme ('money'). However, in (26b) the use of TR_1 indicates that the Undergoer of the second verb (*manaak*) is the recipient. This violation of the single Undergoer constraint renders the sentence ungrammatical. Sentence (26c) is identical except for the use of TR_2 , and is fully grammatical.¹¹

- (26) a. N-a-awi it siin kuh n-i-taak kuh sid tongo
 PAST-NVOL-finish NOM money my PAST-IV-give 1sg.GEN DAT PL

tanak.
 child

'I used up all my money, I gave it to my children.'

- b. *N-a-awi kuh manaak sid tongo tanak it
 PAST-NVOL-finish 1sg.GEN AV- TR_1 -give DAT PL child NOM

siin kuh.
 money my

- c. N-a-awi kuh pataak sid tongo tanak it siin kuh.
 PAST-NVOL-finish 1sg.GEN AV- TR_2 -give DAT PL child NOM money my

'I used up all my money giving it to my children.'

A similar contrast can be shown with the root *pilay* 'throw'. The examples in (27) are based on (15a). The use of TR_1 in (27a) indicates that the Undergoer of V_2 is the goal or target ('throw at, pelt'). Since the nominative argument ('mangoes') must be interpreted as the Undergoer of V_1 , the use of this prefix forces a bizarre interpretation under which the house is being thrown at the mangoes. The intended interpretation is

¹¹ The verb form *naawi* in (26) is ambiguous between an unaccusative sense 'all gone; used up' and a transitive sense 'use up, finish off (OV)'. The gloss reflects the unaccusative (intransitive) sense.

only possible when V_2 bears the TR2 prefix, as in (27b), indicating that the Undergoer is the displaced theme.

- (27) a. *N-a-awi kuh momilai(m-poN-pilai) dit walai nuh
 PAST-NVOL-finish 1sg.GEN AV-TR1-throw ACC house your
- it tongo mangga.
 NOM PL mango

(can only mean: ‘I finished off the mangoes by throwing your house at them.’)

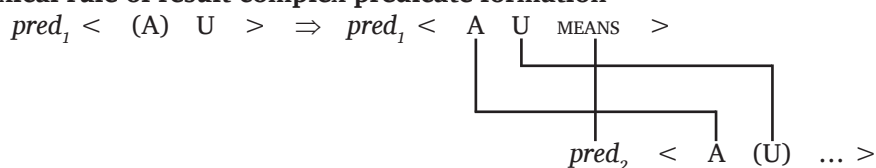
- b. N-a-awi kuh Ø-po-pilai sid walai nuh it tongo
 PAST-NVOL-finish 1sg.GEN AV-TR2-throw DAT house your NOM PL
- mangga.
 mango

‘I threw all the mangoes at your house.’

5. A lexicalist analysis

I have suggested that the two verbs in the resultative construction form a complex predicate, meaning that the two verbs together share a single, complex argument structure. For the sake of concreteness, let us assume that the second verb, expressing the manner or means by which the result is achieved, is incorporated into the argument structure of the first verb by a lexical process something like the following (A = ‘Actor’, U = ‘Undergoer’):

(28) lexical rule of result complex predicate formation



The informal representation in (28) captures the following intuitions: V_1 may be either transitive or unaccusative when used by itself, but the complex predicate as a whole is always transitive, and both the Actor and Undergoer are governed and case marked by V_1 . V_2 may be transitive or unergative, and may even introduce additional arguments such as instruments (see ex. 21); but its Actor and Undergoer (if any) may not be distinct from the Actor and Undergoer of V_1 . Shared arguments are identified by the process of argument structure merger (Alsina 1992), and receive a single syntactic expression.

One reason for treating this as a lexical process rather than a syntactic combination of elements comes from interactions with other derivational processes, in particular with causativization. When the first verb in a complex predicate contains the reciprocal prefix *pi-*, as in (1b) and (14a–b), the interpretation is predictable under a syntactic

analysis. However, when V_1 contains the causative prefix *po-*, the interpretation is much more difficult to account for under a syntactic analysis. Let us begin with the reciprocal pattern.

When a reciprocal verb is marked for Objective Voice, it takes on a causative interpretation as illustrated in (29b–c). (Contrast the normal, non-causative sense of the AV reciprocal form in 29a.) In the OV form the reciprocal action or relationship is predicated of the Undergoer-causee, which must be a group of two or more individuals, and is understood to be caused or brought about by the Actor-causer. When such a verb occurs as the V_1 in a RCP, this Actor-causer is also interpreted as the Actor of V_2 , as illustrated in (30); see also (1b), (14a–b). This identification of Actor arguments seems to be required by the uniqueness constraints discussed above, and is easily accounted for under either a lexical or a syntactic analysis.

- (29) a. Migogol(m-pi-gogol) ilo tasu, minogiad(m[in]-pog-iad)
 AV-RECIP-wrestle that(NOM) dog AV[PAST]-pog-weep

ilo tanak tu' n-ajang-an.
 that(NOM) child because PAST-affect-DV

‘Those dogs were fighting, and that child cried because he got mixed up in it.’

- b. Okon-ko pi-gogol-on inoh tasu, o-pilat-an dati.
 don't RECIP-wrestle-OV that(NOM) dog NVOL-wound-DV likely

‘Don't make those dogs fight each other, they will get wounded.’

- c. Pi-toluod-on noh yo'alo songobpinee, okon.ko pi-odu-on.
 RECIP-kindness-OV FOC 3pl.NOM siblings don't RECIP-quarrel-OV

‘Cause those siblings to treat each other kindly, don't make them quarrel with each other.’

- (30) a. P[in]i-uruk-Ø dialo Ø-pa-akan at sompusasawo.
 [PAST]-RECIP-eat.together-OV 3sg AV-CAUS-eat NOM married.couple

‘He caused/allowed the married couple to eat together.’

- b. Amu obbulih kawoh dot pi-go-gowul-on mangansak(m-poN-ansak)
 NEG allowed PRTCL COMP RECIP-REDUP-mix-OV AV-TR1-cook

o tumbus tu' ko-bunsala kah.
 NOM veggie because AV.NVOL-prone.to.error QUOT

‘You shouldn't cook different kinds of vegetables together, they say it can make you prone to slips of the tongue.’

- c. Pi-abpay-on nopoh Ø-po-wiliw ilo torigi toh!
 RECIP-CROSS-OV only AV-TR2-lie.flat that(NOM) house.post PRTCL

‘Lay those posts crossways to one another.’

When V_1 contains the causative prefix *po-*, we would again expect the causer to be interpreted as the Actor of V_2 . However, as illustrated in (31), it is the causee rather than the causer that receives this interpretation.

- (31) a. N-i-pa-awi kuh di Jaiwan mangakan(m-poN-akan) it
 PAST-IV-CAUS-finish 1sg.GEN ACC Jaiwan AV-TR1-eat NOM

rinapa.
 viand

‘I caused/allowed Jaiwan to eat up all the cooked food.’

- b. Subai ipeerot(i-po-irot) noh di tanak Ø-po-otub i tuung.
 PRTCL IV-CAUS-tight FOC ACC child AV-TR2-cover NOM box

‘You’d better make the child close the lid of the box tightly.’

- c. I-po-birud nopoh dialo manganu(m-poN-anu) at tobu toh.¹²
 IV-CAUS-twist only 3sg AV-TR1-take NOM sugar.cane PRTCL

‘Just have him twist the sugar cane off (pick by twisting).’

At first glance, these sentences seem to violate the Single Actor constraint: the causer is clearly the Actor of V_1 , but the Actor of V_2 is the causee. Based on the evidence presented in section 4 this interpretation should not be possible for this construction. If these sentences are formed by a purely syntactic combination of lexical items, this paradox is quite difficult to deal with. However, if complex predicate formation is a lexical process, it may “feed” other lexical/ derivational processes.

If the rule of result complex predicate formation can apply before the rule of causative formation, there would be no violation here. A basic, non-causative RCP would serve as input to the morphological causative rule. This input form would have a single, unambiguous Actor (e.g. ‘the child’ in 31b), and this Actor would become the causee of the derived causative. A simplified sketch of the derivation for (31b) is presented in (32).

¹² Ex. 31c may be an example of a manner complex predicate (MCP) rather than a RCP.

- (32) a. **Base form:** *tighten* < A , U >
 b. **complex pred.** *tighten* < A , U , MEANS >
 c. **causative** CAUSE < A , EVENT >
 |
 tighten < A , U , MEANS >

6. A possible alternative analysis

In the preceding section we referred somewhat vaguely to a possible “syntactic” analysis of the construction under discussion. One specific analysis that might be suggested is that V_2 is a kind of reduced relative clause modifying the Actor of V_1 . There are a number of reasons why this does not seem like a viable proposal. Clitic pronouns never serve as the head of a relative clause in Kimaragang, yet they are quite common as Actors in the RCP. Stressed pronouns and proper names could only function as the head of a non-restrictive relative, which should be set off by pauses; but such pauses are impossible in the RCP. Relative clauses in Kimaragang normally contain a linking particle (optional, but generally present), but these linkers are not used in the RCP. The modifying clause in a Kimaragang relative construction immediately follows the head noun, but V_2 is frequently not adjacent to the Actor NP. For these reasons we reject this proposal without further discussion.

A more plausible analysis has been suggested to me by Maria Polinsky (p.c.): could V_2 be analyzed as a kind of nominalized adjunct (like a gerund) that is predicated of the Actor of V_1 and gets a causal interpretation? This analysis is somewhat similar to the closest English equivalent to the literal form of RCP examples like (1a), repeated below: ‘The stream was dried up by buffaloes drinking.’

- (1) a. N-a-rasak do karabau monginum(m-poN-inum) at weeg.
 PAST-NVOL-dry.up GEN buffalo AV-TR1-drink NOM water

‘The stream was drunk dry by buffaloes.’

This proposal seems plausible because there are some gerundive uses of “infinitival” verb forms (equivalent to a simple non-past form which is unmarked for aspect, mood or modality). This usage seems to be more common with Actor Voice forms, as illustrated in (33a–b), but is also possible with non-active voices as in (33c–d). (Notice that the implicit arbitrary or generic referent in (33c) is the grammatical subject, while in (33d) it is the agent.)

- (33) a. Ara’at ot [Ø-po-pi-oduw do tulun sompusasawo].
 bad NOM AV-CAUS-RECIP-quarrel ACC person married.couple

‘To cause a married couple to quarrel is evil.’

- b. [Monimba'al(m-poN-simba'al) do tulun] nga ara'at = i'.
 AV-TR1-slander ACC person also bad = PRTCL

'Slandering people is evil too.'

- c. [Tinduk-on do wulanut] nga ka-patay = i'.
 peck-OV GEN snake also AV.NVOL-kill = PRTCL

'Getting bitten by a snake can kill you too.'

- d. [Pi-igol-on ot tasu] nga amu = i' obbulih tu
 RECIP-dance-OV NOM dog also NEG = PRTCL allowed because

ko-bunsut.

AV.NVOL-curse

'Making dogs dance with each other is not allowed either, because you could be struck by the *bunsut* curse (be swallowed into the ground).'

A gerundive analysis would explain why V_2 cannot be inflected for tense/aspect/mood, and would account for the monoclausal word-order pattern as well. The fact that the gerund is predicated of the Actor would explain the uniqueness constraint for Actors, and might be made to explain why V_2 always appears in the Actor Voice form. The uniqueness of the Undergoer is not so easy to account for under this analysis; some special syntactic mechanism, perhaps a special kind of control, would have to be assumed. But the great advantage of this proposal is that it makes an exotic (i.e., unusual) construction seem much more familiar, that is, describable in more conventional terms.

One argument against the gerundive analysis was given in the previous section: if the nominalized adjunct V_2 must be predicated of the Actor, there is no obvious way to account for the meaning of the causative examples in (31). We might imagine that the nominalized adjunct could be predicated of either causer or causee, depending on semantic and pragmatic plausibility, since both are in a sense Actors. But this does not seem to be the case. When V_1 is a reciprocal form with causative meaning, as in (30), V_2 can only be predicated of the causer. This restriction is confirmed by the unacceptability (34b), where the form of V_2 would be appropriate to denote the action of the causee. These facts are what we would expect under either the gerundive analysis or the lexicalist analysis outlined in the preceding section.

- (34) a. P[in]i-uruk-Ø dialo Ø-pa-akan at sompusasawo.
 [PAST]-RECIP-eat.together-OV 3sg AV-CAUS-eat NOM married.couple

'He allowed the married couple to eat together.' [= 30a]

- b. *P[in]i-uruk-Ø dialo mangakan(m-poN-akan) at
 [PAST]-RECIP-eat.together-OV 3sg AV-TR1-eat NOM
- sompusasawo.
 married.couple

However, just the opposite is true when V_1 is a causative. Here the form of V_2 must be appropriate to denote the action of the causee, as illustrated in (35). This contrast strongly suggests that the predication relation is grammatically determined, since either form would be semantically plausible in both cases. The difference in interpretation between the reciprocal and the causative, when they appear in the V_1 position, is easy to explain using ordered lexical rules; but it presents a significant challenge to the nominalized adjunct analysis.

- (35) a. N-i-pa-awi kuh di Jaiwan mangakan(m-poN-akan) it
 PAST-IV-CAUS-finish 1sg.GEN ACC Jaiwan AV-TR1-eat NOM
- rinapa.
 viand

‘I caused/allowed Jaiwan to eat up all the cooked food.’ [= 31a]

- b. *N-i-pa-awi kuh di Jaiwan Ø-pa-akan it rinapa.¹³
 PAST-IV-CAUS-finish 1sg.GEN ACC Jaiwan AV-CAUS-eat NOM viand

A second problem for the nominalized adjunct analysis concerns the lexical properties of V_1 . We have noted that V_1 may be either transitive or unaccusative when it is used by itself. In the result complex predicate, however, V_1 is always syntactically transitive as indicated by the pattern of case assignment for the Actor and Undergoer. Along with this shift in valence, there may be minor but unpredictable changes in semantic content, a characteristic feature of lexical processes.

Some examples of these changes are presented in (36). The verb form that appears as V_1 in (36a), *norikot*, can be used alone as an unaccusative verb meaning ‘has arrived’; but in (36a) it functions as a transitive verb meaning ‘to reach (e.g. by throwing)’. If the second verb were not present, *norikot* could not be used in this way. Similarly, the verb form *narasak* in (36b) is an unaccusative verb meaning ‘dried up’. It takes on a transitive sense in the RCP; but to be used in this sense without the second verb, e.g. if someone dries up a stream or pond to catch the fish, an extra suffix would be needed: *narasakan* or *rinasakan*.

- (36) a. N-o-rikot kuh *(momilai[m-poN-pilai]) i walai nuh.
 PAST-NVOL-arrive 1sg.GEN AV-TR1-throw NOM house your

‘I threw (something) all the way to your house.’ (cf. 15a)

¹³ Jim Johansson (p.c.) informs me that *paakan* would be possible in (35b) with an indefinite plural causee, e.g. ‘people’; the resulting sentence would imply distributing a large quantity of food to many people until it was all gone. This interpretation may involve a non-causative sense of *paakan*, but more investigation is required.

b. N-a-rasak do karabau *(monginum[m-poN-inum]) a
 PAST-NVOL-dry.up GEN buffalo AV-TR1-drink NOM

weeg.
 stream

‘The stream was drunk dry by buffaloes.’ (cf. 1a)

A somewhat similar pattern also occurs in English. Unergative verbs occurring in resultative constructions often take on a slightly different sense, as illustrated in (37). Some authors, following Simpson (1983), have interpreted this to mean that these verbs must undergo a change in argument structure in order to appear in the resultative.¹⁴

- (37) I shouted myself *(hoarse).
 I cried myself *(to sleep).
 I worked my fingers *(to the bone).
 She drank him *(under the table).

The same considerations apply even more so to Manner examples like those in (2), in which the subject is rarely if ever a possible semantic argument of V_1 . I will not claim that it is impossible for the gerundive analysis to account for these examples, but I do not see how to make it work. Given the information currently available, the complex predicate analysis seems more plausible.

7. Conclusion

The analysis presented here should be considered tentative, since research into this construction is still in its preliminary stages. The evidence against a clause boundary between the two verbs seems to me fairly solid. The uniqueness effects and obligatory argument sharing seem most naturally described in terms of a common argument structure. The ordering paradox that arises when V_1 is a causative form is easy to resolve if complex predicate formation is a lexical process, but very awkward if the two verbs stand in a purely syntactic relationship to each other. Comparative evidence from the Formosan languages could help to resolve some of these questions. It would also be helpful to know whether this construction occurs in any languages within the Philippines.

References

- Alsina, Alex. 1992. “The argument structure of causatives.” *Linguistic Inquiry* 23:517–555.
- Carrier, Jill, and Janet Randall. 1992. “The argument structure and syntactic structure of resultatives.” *Linguistic Inquiry* 23:173–234.

¹⁴ Note however that this conclusion is rejected by Levin and Rappaport Hovav (1995) and Carrier and Randall (1992).

- Chang, Henry Yungli. (in press). "The guest playing host: adverbial modifiers as matrix verbs in Kavalan." To appear in *Clause structure and adjuncts in Austronesian languages*, ed. by Hans-Martin Gaertner and Paul Law. Berlin: Mouton de Gruyter.
- Dell, Francois. 1983. "An aspectual distinction in Tagalog." *Oceanic Linguistics* 22–23:175–206.
- Holmer, Arthur. 2004. "Intraposition and Formosan adverbial heads." Handout from AFLA 11, Berlin.
- Kroeger, Paul R. 1990. "Stative aspect and unaccusativity in Kimaragang Dusun." *Oceanic Linguistics* 29.2:110–131.
- Kroeger, Paul R. 1996. "The morphology of affectedness in Kimaragang Dusun." In H. Steinhauer, ed., *Papers in Austronesian Linguistics No. 3*. Pacific Linguistics A–84. Canberra: Research School of Pacific and Asian Studies, Australian National University.
- Kroeger, Paul R. and Jim Johansson. 2005. "Object alternations in Kimaragang." In J. Heinz & D. Ntelitheos, eds., *Proceedings of the Twelfth Annual Conference of the Austronesian Formal Linguistics Association (AFLA)*. *UCLA Working Papers in Linguistics*, no. 12.
- Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity: at the syntax-lexical semantics interface*. Linguistic Inquiry monographs 26. Cambridge, Mass.: MIT Press.
- Simpson, Jane H. 1983. "Resultatives." In Rappaport, Levin and Zaenen, eds., *Papers in LFG*. Indiana University Linguistic Club.

Transitivity, Valence and Voice in Mandar*

Jason Kwok Loong Lee
Department of Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200

The categorisation of voice systems in non-Oceanic Austronesian languages has long been an issue. One categorisation distinguishes symmetrical voice from asymmetrical voice languages (e.g. Himmelmann 2005). However, a necessary precursor to identifying voice variation is identifying transitivity and valence of clauses. A necessary precursor to identifying transitivity and valence is the identification of core arguments. Mandar's voice system is claimed to be asymmetrical, like its close relation Makassar (Jukes 1998; Jukes 2005). Although Mandar's voice system is not symmetrical, there are some clause types that appear to be neither transitive nor intransitive, but "semi-transitive". The difficulty in identifying the transitivity status of these clauses is due largely to the difficulty in identifying whether an argument is core or non-core. In this paper I will describe transitivity, valence and voice in Mandar in its typological context.

1. Introduction

In this paper I discuss how transitivity and voice are encoded in Mandar. I pay particular attention to what others have called the "semi-transitive" (Friberg 1991; Jukes 2005) construction in Mandar and closely related languages. After examining transitivity more closely, I propose calling them "extended intransitives" instead. I also show how the encoding of voice is closely tied to the encoding of transitivity.

What is "transitivity"? Hopper and Thompson (1980) see transitivity as a cline involving "a number of components, only one of which is the presence of an object of the verb" (1980:251). Their components are basically semantic (including discourse related properties) and concern the degree to which one thing affects another, but the presence of an object of the verb is a syntactic encoding feature. I prefer to refer to the semantic aspects of transitivity as "valence" and to reserve the term "transitivity" for their manifestation in morphosyntax.

* An earlier version of this paper was presented at the ALS (Australian Linguistics Society) conference in Melbourne, Australia in September 2005. I thank John Bowden, Malcolm Ross, I Wayan Arka from the ANU, Bill Foley, and the rest of the audience at the ALS conference for their questions and comments. All remaining errors are mine. This paper presents research made possible by the Indonesian Institute of Sciences (LIPI) and my core sponsor the *Balai Bahasa Ujung Pandang, Departemen Pendidikan dan Kebudayaan*. The greatest thanks go to the Mandar speakers for teaching me their language.

Secondly, what is “voice”? Before we define voice, we need to recognise some basic semantic-syntactic roles or macroroles. In the realm of semantics, we can refer to “actor” and “undergoer” or “agent” and “patient”. As we step into the realm of morphosyntax, the semantic macroroles are translated into what are commonly called grammatical relations, relations between arguments and predicates. It is convenient to be abstract and follow Comrie (1978) in referring to “A” and “P” respectively as the more agent-like and the more patient-like argument in a multi-argument clause. And we can refer to “S” as the only argument of a single argument clause. (Dixon and others use “O” instead of “P” to refer to the same thing, as in Dixon 1994). Consequently, I take “voice” to refer to alternations in morphosyntax that affect the mapping between grammatical relations and semantic macroroles.

2. Core arguments

Before statements on the syntactic transitivity of a structure can be made, one must either mention or make assumptions about the core status of the argument or arguments of a structure. Mandarin is a head-marking language where pronominal clitics occur on clause heads and may thus be considered the core arguments. Examples (1) and (2) can be considered canonical intransitive clauses and examples (3) and (4) canonical transitive clauses. These examples characterise two basic clause types, represented schematically in (5) and (6) (a summary of the schematic representation of all the clause types discussed in this paper is presented in Table 1). It is also clear from these examples that the pronominal clitics pattern in an ergative-absolutive manner. The S of the intransitive clauses pattern with the P of the transitive clauses. Table 2 lists the pronominal clitics.

- (1) *Umande aq.*¹
 -um-ande = aq
 -A_{ITR}²-eat = 1s.ABS

‘I eat.’

- (2) *Matindo aq.*
 ma-tindo = aq
 U_{ITR}-sleep = 1s.ABS

‘I sleep.’

¹ Most of the examples in this paper are elicited or recorded in language-learning sessions. A few are taken from Muthalib (1977). In this orthography, *q* represents the glottal stop.

² Abbreviations used in this paper are: 1 – first person; 2 – second person; 3 – third person; A – actor, agent, most agent-like argument of a multi-argument clause; ABS – absolutive; APP – applicative; AUX – auxiliary; AV – actor voice; BEN – benefactive; DIST – distal deictic; ERG – ergative; ITR – intransitive; NEG – negative; NP – noun phrase; P – patient, most patient-like argument of a multi-argument clause; PERF – perfect; PERS – personal name; POSS – possessive; PROX – proximal deictic; s – singular; S – only argument of a single-argument clause; U – undergoer; V – verb.

Table 1. Summary of clause types (types marked with an asterisk (*) indicate clause types that I would expect but that I have not yet found in my data)

	Intransitive (examples)	Transitive (examples)
basic (a)	ITRV = ABS ^S (NP ^S) (1), (2)	ERG ^A = TrV = ABS ^P (NP ^P) (3), (4)
(b)	AUX = ABS ^S (NP ^S) (9) ITRV	AUX = ABS ^P ERG ^A = TrV (NP ^P) (13)
(c)	NP ^S (AUX) ITRV (10)	NP ^P ERG ^A = TrV (14)
extended (a)		*ERG ^A = TrV = ABS ^{P1} (none) NP ^{P2}
(b)		*AUX = ABS ^{P1} ERG ^A = TrV NP ^{P2} (none)
(c)		NP ^{P1} ERG ^A = TrV NP ^{P2} (24)–(27)
maC-extended (a)	ITRV = ABS ^A NP ^P (7)	
(b)	AUX = ABS ^A (NP ^A) (20) ITRV NP ^P	
(c)	NP ^A (AUX) ITRV (18), (21) NP ^P	NP ^A TrV = ABS ^{P1(recipient)} (29), (30) NP ^{P2(theme)}
Ditransitive?		
basic (a)		
(b)		
(c)		NP ^{P1} ERG ^A = TrV (32) = ABS ^P NP ^{P2}

- (3) *Mu pipal aq.*
mu = pipal = aq
 2.ERG = slap = 1s.ABS

‘You slap me.’

- (4) *U ande i lokamu.*
u = ande = i loka-mu
 1s.ERG = eat = 3.ABS banana-2.POSS

‘I eat your bananas.’

- (5) ITRV = ABS^S 3 (NP^S) [“intransitive”, examples (1) and (2)]
- (6) ERG^A = TRV = ABS^P (NP^P) [“transitive”, examples (3) and (4)]

In example (4), the undergoer is represented by both the third-person pronominal enclitic =*i*, and a free noun phrase *loka-mu* ‘your bananas’. Here, the enclitic can still be considered the core argument while the free noun phrase is an optional co-referential NP in apposition to the enclitic. However, leaving out the NP still implies a definite or specific P, identifiable from context.

Table 2. Pronominal clitics and some possessive suffixes

	Ergative	Absolutive	Possessive
1 st person	u =	= aq	-u
2 nd person	mu =	= o	-mu
3 rd person	na =	= i	-na

When the undergoer is indefinite, another common construction tends to be used. Jukes describes parallel constructions in Makassar as “semi-transitive” (2005:664). I, however, am calling such constructions “extended intransitives” (after Dixon and Aikhenvald (2000:3–4), but closer in definition and usage to Ross and Teng (in press:16–19)). Example (7) illustrates the extended intransitive construction, represented schematically in (8). Here, only the actor is represented by a pronominal clitic. The undergoer is not represented by a clitic, but by a full noun. The noun here is obligatory.

- (7) *Maqande aq loka.*
maC-ande = aq loka
 AV-eat = 1S.ABS banana
 ‘I eat bananas.’

- (8) ITRV = ABS^A NP^P [“extended intransitive”, example (7)]

Before examining the extended intransitive more closely, I will first go back and examine the basic intransitive and transitive clause structures.

2.1. Intransitive

The schema for the basic intransitive clause given in (5) has a few common variations, as exemplified by (9) and (10) and schematised in (11) and (12).

³ The superscripted abbreviations identify the argument’s macrorole, e.g. =ABS^P is an absolutive enclitic which is a P argument, etc.

- (9) *Meloq aq umande.*
meloq = *aq* -*um-ande*
 want = 1_{S.ABS} -A_{ITR}-eat

‘I want to eat.’

- (10) *I Ali membuni.*
i = *Ali meC-buni*
 PERS = Ali A_{ITR}-hide

‘Ali is hiding.’

- (11) A_{UX} = A_{BS}^S (NP^S) I_{TRV}

- (12) NP^S (A_{UX}) I_{TRV}

We can also note that almost all intransitive verbs are derived forms. That is, it is difficult to find bare-root intransitive verbs in clauses. Table 3 lists affixes that are found in the examples.

Table 3. Some verbal affixes

Affix	Function	Comments
- <i>um</i> -	actor intransitive	archaic; unproductive; but on very common verbs
<i>ma</i> -	undergoer intransitive	commonly found on stative and ‘adjectival’ roots
<i>ma</i> -	actor intransitive	rare could be an allomorph of <i>meC</i> -
<i>me</i> -	actor intransitive	could be an allomorph of <i>meC</i> -
<i>meC</i> -	actor intransitive	more common (productive?) than - <i>um</i> -
<i>maC</i> -	actor voice	
- <i>ang</i>	benefactive applicative	benefactive; distributive; collective action
- <i>i</i>	locative applicative	

It is difficult to predict which verbs can take which actor intransitive affixes. The choice of -*um*- over *meC*- for the root *ande* ‘eat’ is not predictable. The valence of the root (semantic transitivity or lexical transitivity) helps only a little in predicting whether a verb will have both intransitive and transitive forms. It is necessary to work the other way round and deduce the root valence of a verb by testing and searching whether they occur as intransitive and/or transitive forms.

2.2. Transitive

The basic transitive clause as schematised in (5) has similar variations to those of the basic intransitive clause. These are exemplified in (13) and (14) and schematised in (15) and (16). If we look at the summary of clause types (Table 1), we see that the (b) variations involve an auxiliary and the absolutive enclitic attaching to that instead of to the main verb. The (c) variations involve a fronted NP.⁴

- (13) *Andiang i u ita duriammu.*
andiang =i u= ita duriang-mu
 NEG = 3.ABS 1s.ERG = see durian-2.POSS

‘I did not see your durian.’

- (14) *Lokamu u ande.*
loka-mu u= ande
 banana-2.POSS 1s.ERG = eat

‘It is your bananas that I eat. / I eat your bananas.’

- (15) $AUX = ABS^P \text{ ERG}^A = \text{TrV} (\text{NP}^P)$

- (16) $\text{NP}^P \text{ ERG}^A = \text{TrV}$

The above examples show bare root verbs. However, derived forms can also occur in basic transitive clauses with the causative *pa-* and applicatives *-ang* (benefactive/distributive applicative) and *-i* (locative applicative). Example (17) is a simple transitive clause with a derived verb stem.

- (17) *U itai i.*
u= ita-i =i
 1s.ERG = see-APP = 3.ABS

‘I search for it.’

2.3. Core arguments

In the previous two sections on the basic intransitive and transitive clauses, I have made implicit assumptions about the identity of core arguments. To be able to identify the transitivity of a clause, we need to first be able to identify the core arguments.

Ross (2002:28) lists three conditions for an argument being core. They are,

⁴ An examination of structures involving a fronted NP is interesting in its own right, but is beyond the scope of this paper, where I only briefly touch on it. It probably has much to do with discourse features but that is not discussed here.

- (a) The argument has morphosyntactic relationship to the verb. This relationship may be marked by coding on the verb (e.g. agreement affixes), by coding on the arguments (e.g. case marking), or by position in the clause. At the same time, the argument is not oblique: an argument is oblique if an argument with the same structure may also occur as a peripheral argument (one not required by verbal valency), as in *I was working on the floor*.
- (b) The argument is required by the valency of the verb (or, ‘subcategorised for by the verb’). This is a necessary, but not a sufficient condition, as verbal valency may also require an oblique argument, as in *I gave the apple to the man* or *I put the apple on the floor*.
- (c) The argument has reference-related functions. If the argument is not the pivot, then it will have fewer reference-related functions. This again is a necessary, but not a sufficient condition, as in some languages an oblique argument may also have reference-related functions. (Ross 2002:28).

He states that only the first condition is sufficient. The latter two conditions are “necessary” but not “sufficient”. I have identified core arguments in Mandar by two forms of morphosyntactic coding. The first form of coding is the marking on the verb by pronominal clitics. That is, the pronominal clitics are core arguments. The second form of coding is position in the clause as a fronted NP. This fronted NP is also a core argument.

Arka (2005) has demonstrated the importance of using language-specific core-indices in determining the core status of arguments. He has 12 tests for core properties for Balinese and 11 tests for Indonesian. Unfortunately, I have not conducted sufficient elicitation to thoroughly determine such an extensive lists of tests that are applicable to Mandar in determining core status. This is an important further step I need to take in my research.

2.4. Extended intransitive

We now return to the extended intransitive construction.

I mentioned in regard to example (7) that the NP^P is usually indefinite or non-specific. However, a definite P is sometimes possible, as in example (18).

- (18) *Yau pura maqatang asunna.*
yau pura maC-atang asu-na
 1s already AV-hit dog-3.poss

‘I already hit his dog.’

Examples (17) and (18) look like antipassive-voice clauses. The A argument occurs in the absolutive form and position. Indefinite bananas are less salient than definite bananas. What is salient is the action of eating. Even in example (18), the left-detached position of ‘I already’ makes that salient, not whose dog nor what was done. However, in canonical antipassives, the P argument should be optional or marked by an oblique. This is not so in Mandar.

Examples (7) and (18) also look like a canonical intransitive clauses in terms of their having only a single core argument. However, the obligatory NP^P raises challenges. Example (19) presents a comparison to (18) where the same verb root is used in an intransitive (19) and extended intransitive (18) clause.

- (19) *Tappana pura meatang, raqmusang mi maindong.*
tappa-na pura me-atang raqmus-ang =m=i ma-indong
 straight-3. POSS already A_{TR}-hit hurriedly-BEN =PERF=3. ABS A_{TR}-run

‘Right after hitting, they hurriedly ran.’

Anna Margetts (1999) got around a similar descriptive issue in Saliba by advocating a distinction between an “inner” and “outer” core; and, a separation of transitivity onto three levels. The three levels of transitivity that she proposed are “valence” for the domain of the verb root, “word-level transitivity” for the domain of the inflected verb, and “clause-level transitivity” for the domain of the clause. For Margetts, word-level transitivity mapped onto the inner core and was defined by the number of pronominal affixes on the Saliba verb. Clause-level transitivity mapped onto the whole clause and was defined by the overall number of syntactic arguments in the clause.

Attempting to apply Margetts’ distinctions to Mandar has provided some answers, but also raised other questions. If word-level transitivity applies to the inner core, example (7) could be seen as being intransitive at the word level (only one pronominal clitic) but transitive at the clause level (two referents, one marked by a clitic, the other by a noun). However, because Mandar has clitics rather than verb affixes, talking about the “word” level is a misnomer. Mandar clitics attach to phrases rather than words, so if the verb is modified by something like a negative, a different pattern of cliticisation occurs, as in example (20).

- (20) *Andiang aq maqande duriang.*
andiang =aq maC-ande duriang
 NEG =1s.ABS AV-eat durian

‘I am not eating durian.’

Another approach would be to consider ‘bananas’ in example (7) to be an incorporated noun in the verb. That is, that ‘eat bananas’ is an intransitive verb, and so only requires a single S argument.

We can return to defining the transitivity of a clause by the number of core arguments in the clause: these core arguments are identified as being the pronominal enclitics and/or a fronted NP. On this basis, the NP^P does not qualify as a core argument. Others might make a case that the obligatory status of the NP^P makes it a candidate for core status, but I do not.

There is an apparent mismatch between form and meaning. In form the clause is intransitive. In meaning it is transitive and its verb is “lexically transitive” (divalent?) in the sense that there is a corresponding form in which ‘bananas’ is the subject.

The extended intransitive has parallel variations to the variations listed for the intransitive and transitive construction. Example (20) is an extended intransitive with

an auxiliary. Examples (18) and (21) are extended intransitive clauses with a fronted NP instead of an absolutive enclitic. These are represented schematically in (22) and (23).

- (21) *Yau maqande diqo lokamu o.*
yau maC-ande diqo loka-mu = o
 1s AV-eat DIST banana-2.POSS = DIST
 ‘I ate those bananas of yours.’

- (22) $AUX = ABS^A (NP^A) ITRV NP^P$

- (23) $NP^A (AUX) ITRV NP^P$

2.5. Extended transitive

Having seen that there is an extended intransitive construction in Mandar, it is not surprising to find that there is a corresponding extended transitive construction. These occur particularly with semantically trivalent verbs. The verb ‘give’ is an example of a semantically trivalent verb. Examples (24) to (27) parallel the transitive example (14). In (24) and (25) the recipient is fronted. In (26) and (27) the theme is fronted. Both recipients and themes can be considered undergoers. These examples show that recipients and themes are treated in the same manner. In all the examples (24) to (27), the ergative proclitic marks the agent. Note too that the benefactive applicative *-ang* is used for all these examples. The schematic representation of these examples (24) to (27) is given in (28).

- (24) *Innai mu bengang gulanggu?*
innai mu = be-ang gulang-u
 who 2.ERG = give-BEN rope-1s.POSS
 ‘Who did you give my rope to?’

- (25) *Anaqu u bengang gulammu.*
anaq-u u = be-ang gulang-mu
 child-1s.POSS 1s.ERG give-BEN rope-2.POSS
 ‘I gave your rope to my child.’

- (26) *Apa mu bengang anaqmu?*
apa mu = be-ang anaq-mu
 what 2.ERG = give-BEN child-2.POSS
 ‘What did you give your child?’

- (27) *Gulammu u bengang anaqu.*
gulang-mu u = be-ang anaq-u
 rope-2.POSS 1s.ERG = give-BEN child-1s.POSS

‘I gave your rope to my child.’

- (28) NP^{P1} ERG^A = TRV NP^{P2}

When the actor is fronted, no ergative proclitic occurs and the verb is prefixed with *maC-*. The absolutive enclitic marks the recipient. Instead of the benefactive applicative *-ang*, the locative applicative *-i* is used. Examples (29) and (30) show a fronted actor NP and are schematically represented in (31).

- (29) *Innai mambei o diqo gulang?*
innai maC-be-i = o diqo gulang
 who AV-give-APP = 2. ABS DIST rope

‘Who gave you that rope?’

- (30) *Kamaqu mambei aq diqe gulang.*
kamaq-qu maC-be-i = aq diqe gulang
 father-1s.POSS AV-give-APP = 1s.ABS PROX rope

‘My father gave me this rope.’

- (31) NP^A TRV = ABS^{P1(recipient)} NP^{P2(theme)}

2.6. Ditransitive?

I have only one example of what looks like a ditransitive construction, example (32). This is the only example I have of a fronted NP co-occurring with what looks like a co-referential absolutive enclitic. It is not helpful in showing the recipient marked by the enclitic because both the recipient and theme are in the third person. I’m not too sure about this example. It could be an error as it is structurally and semantically very similar to example (27).

- (32) *Gulammu u bengan i anaqu.*
gulang-mu u = be-ang = i anaq-qu
 rope-2.POSS 1s.ERG = give-BEN = 3. ABS child-1s.POSS

‘I gave your rope to my child.’

3. Voice

At the start of this paper, I stated that voice refers to alternations in morphosyntax that affect the mapping between grammatical relations and semantic macroroles.

From the above discussion, we can see that the *maC*- extended constructions are in a paradigmatic voice relationship with parallel constructions.

The *maC*- extended intransitive construction is in a paradigmatic voice relationship with the transitive construction. The extended intransitive is similar to a canonical antipassive. However, the antipassive is commonly defined with the demotion of the P argument from a core to an oblique. Here, the P argument is demoted from core to a non-core obligatory extended argument. Because of the differences in transitivity, the voice alternation may be described as asymmetrical.⁵

With the semantically trivalent verb 'give', there appear to be two kinds of voice alternation. The first is a simple syntactic alternation between the fronting of either the recipient or the theme. This could be described as a syntactic voice alternation between a recipient voice and theme voice. Here, the alternation is purely syntactic, not morphological.

The morphosyntactic variation between the *maC*- extended transitives and basic extended transitives also marks a voice alternation. If the verb is prefixed with *maC*-, the fronted NP is the actor. Otherwise, the fronted NP is an undergoer (either a theme or recipient).

4. Conclusion

In this paper I have defined and restricted the terms transitivity and valency to morphosyntax and semantics respectively. By doing so I have shown that in Mandar clauses are syntactically either intransitive or transitive. The evidence for ditransitive clauses is not enough for me to be sure that this can be considered a third class of clauses in terms of transitivity. I have shown that the *maC*- extended constructions are in a voice relationship with related transitive constructions, but these voice alternations are asymmetrical rather than symmetrical.

References

- Arka, I Wayan. 2005. *The core-oblique distinction and core index in some Austronesian languages of Indonesia*. Paper presented as a keynote address at the International Association of Linguistic Typology conference (ALT VI) in Padang, Indonesia.
- Comrie, Bernard. 1978. Ergativity. In Winfred P. Lehmann, ed., *Syntactic typology*, 329–394. Austin: University of Texas Press.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge, New York, Melbourne: Cambridge University Press.
- Dixon, R.M.W. and Aikhenvald, Alexandra Y. 2000. Introduction. In R.M.W. Dixon and Alexandra Y. Aikhenvald, eds., *Changing valency: case studies in transitivity*, 1–29. Cambridge, New York, Melbourne: Cambridge University Press.

⁵ Of course, those who argue that what I have called the 'extended argument' is indeed a core argument can then also say that Mandar is symmetrical.

- Friberg, Barbara. 1991. Ergativity, focus and verb morphology in several South Sulawesi languages. In Ray Harlow, ed., *VICAL 2: Western Austronesian and contact languages, parts I and II: Papers from the fifth International Conference on Austronesian Linguistics, Auckland, New Zealand*, 103–30, Auckland: Linguistic Society of New Zealand.
- Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: typological characteristics. In Alexander Adelaar and Nikolaus P. Himmelmann, eds., *The Austronesian languages of Asia and Madagascar*, 110–181, London, New York: Routledge.
- Hopper, Paul J. and Thomson, Sandra A. 1980. Transitivity in grammar and discourse. *Language*. 56(2). 251–299.
- Jukes, Anthony. 1998. *The phonology and verbal morphology of Makassar (Basa Mangkasara’): An Austronesian language of South Sulawesi*. Unpublished MA thesis. University of Melbourne.
- Jukes, Anthony. 2005. Makassar. In Alexander Adelaar and Nikolaus P. Himmelmann, eds., *The Austronesian languages of Asia and Madagascar*. 649–682. London; New York: Routledge.
- Margetts, Anna. 1999. *Valence and transitivity in Saliba: an Oceanic language of Papua New Guinea*. Nijmegen: MPI.
- Muthalib, Abdul. 1977. *Kamus bahasa Mandar – Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.
- Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In Wouk, Fay and Ross, Malcolm, eds., *The history and typology of western Austronesian voice systems*. 17–62. Canberra: Pacific Linguistics.
- Ross, Malcolm and Teng, Stacy Fang-ching. In press. Formosan languages and linguistic typology. In *Language and linguistics*.

Is there a VP in Pendau?*

Phil Quick
SIL International

*In this paper I will address whether or not Pendau (a Tomini-Tolitoli language in Central Sulawesi, Indonesia) has in fact a VP constituent. Himmelmann (2005:142–143) suggests that in another Tomini-Tolitoli language, the Totoli language, there is a VP constituent similar to a number of other western Austronesian languages. Pendau has a symmetrical voice system that contrasts active voice and inverse voice (see Quick 1997, 2002, 2003). Subjects and second objects may never be placed between the verb and the grammatical object. Both voices can be described as having the same pragmatic flexible word order: 1) S[VO]_{VP} and 2) [VO]_{VP}S. The main part of the paper will present data and discuss the evidence for analyzing the Pendau data as having a VP. Although core arguments can never be placed between the verb and the object, exceptions to this will include data with fronted oblique arguments and the optional ‘floating’ adverb *moje* ‘also, too’. A brief description on the floating adverb will be given, but as this parallels other exceptions found in languages with VPs, this will not be a problem to the VP analysis. The research demonstrates that the data for the fronted obliques inside the VP has only been found in elicited material. I briefly discuss why this data does not negate the viability of a VP constituent in Pendau. Finally, there is a brief discussion on viewing the VP as part of a topic-comment construction. This will include a brief discussion on the discourse possibilities for fronting the VP and future research.*

1. Introduction

In this paper I will address whether or not Pendau (a Tomini-Tolitoli language in Central Sulawesi, Indonesia) has in fact a VP constituent. Himmelmann (2005:142–143) suggests that in another Tomini-Tolitoli language, the Totoli language, there is a VP constituent similar to a number of other western Austronesian languages. Some Sulawesi languages that have been discussed as having a VP or at least an obvious VO include *Tukang Besi* (Donohue 1995, 1999), *Kaili-Ledo* (Evans 2003), and *Bambam* (Campbell 1989). Other non-Sulawesi Indonesian languages with a VP are represented by *Balinese* (Arka 2003, Manning 1996) and *Toba Batak* (Manning 1996).

* Abbreviations: 1PL, first plural, 1SG first singular, 2PL second plural, 3SG third singular, 3PL third plural, A actor/agent, AB absolute, CN common noun, COM comitative, COMP completive aspect, DIR directional, DY dynamic verb class, EXC exclusive, GE genitive case, INC inclusive, IR irrealis, INSTR instrument case, IV inverse voice, LOC locative case, NP noun phrase, OBJ object, P undergoer/patient, PN proper noun, PP prepositional phrase, PT primary transitive verb class, RE realis, SF augmented stem former, ST stative verb class, SUBJ subject, TZ transitivizer, V verb, VP verb phrase

Pendau has a symmetrical voice system that contrasts active voice and inverse voice (see Quick 1997, 2002, 2003, Himmelmann 2002, 2005, and Ross 2002a, 2002b). Transitive word order constructions have a ‘rigid’ position following the verb. The simplest and most elegant solution is to consider that this constituent is the grammatical object. Pendau also has variable word order in which the ‘flex’ word order position appears before the verb or after the ‘rigid’ argument, and can be called the grammatical subject. Subjects and second objects may never be placed between the verb and the grammatical object. In summary then both voices can be described as having a variable word order with these two possibilities in which I will assume tentatively that there is a VP constituent: 1) S[VO]_{vp} and 2) [VO]_{vp}S. The main part of the paper will present data and discuss the evidence for analyzing the Pendau data as having a VP.

Although core arguments can never be placed between the verb and the object, some elicited examples show that oblique arguments can be placed in between the verb and object when an oblique is fronted for prominence purposes. The only other known exceptions are an optional ‘floating’ adverb *moje* ‘also, too’ and the floating quantifier *jojoo* ‘all’ (see Quick 2003 for examples of this). A brief description will be given of the floating adverb *moje*, but as this parallels other exceptions found in languages with VPs, this will not be a problem to the VP analysis (see Kroeger 2004 for example).

The remaining syntactic discussion will show that fronted obliques do not negate the viability of a VP constituent in Pendau. Following this section there will be a brief discussion on viewing the flexible word order as VP fronting and the probable correlation with topic-comment structure.

2. Basic Grammar Notes

This section presents basic grammar notes that are necessary for this paper. Examples (1)–(4) give typical transitive and intransitive constructions in their default word order. Examples (1)–(2) contrasts the symmetrical transitive voices. Example (3) gives an intransitive example of the dynamic verb class which is a mixed verb class of transitive and intransitive verbs. Example (4) gives an example of the stative verb class. There are seven canonical verb classes in Pendau, all of which can be identified either by the particular stem former associated with its verb or in the case of statives with no stem former (Quick 1999, 2003, 2005). In this paper most of the examples will use verbs from the primary transitive verb class. These can be inflected in either active voice¹ or in inverse voice.² In order to expedite the discussion I will sometimes refer to active voice verb constructions as *nong*-verbs, and refer to the inverse voice verb as the *ni*-verb (using the realis form for both voices).

¹ Active voice is not marked in the interlinear representations as it is a composite result of the fusion of several formatives. For informal purposes the active voice may be referred to as the *nong*-formative.

² This is a pragmatic inverse voice, *not* a semantic inverse voice system (see Givón 1994, 2001). The analysis of inverse voice coincides with the ‘focus system’ school of thought, and does not in principle clash with this view.

- (1) **Active Voice**
Siama'u nonuju siina'u.
si = ama = 'u N-pong-tuju si = ina = 'u
 PN/AB = father = 1SG/GE RE-SF/PT-send PN/AB = mother = 1SG/GE
Pivot = A non-pivot = P
 'MY FATHER sent my mother.'
- (2) **Inverse Voice**
Siama'u nituju niina'u.
si = ama = 'u ni-tuju ni = ina = 'u
 PN/AB = father = 1SG/GE IV/RE-send PN/GE = mother = 1SG/GE
Pivot = P non-pivot = A
 'My mother sent MY FATHER.'
- (3) **Dynamic Intransitive Construction**
SiYusup neriing.
si = Yusup N-pe-riing
 PN/AB = Joseph RE-SF/DY-bathe
Pivot = S_A
 'Joseph bathed.'
- (4) **Stative Intransitive**
SiYusup nanabu.
si = Yusup no-nabu
 PN/AB = Joseph ST/RE-fall
Pivot = S_p
 'Joseph fell (down).'

Table 1 gives the basic case paradigm. Note that in example (1) that both NPs mark their core arguments with the same case. I call this the 'absolute case', as it turns out that this is the same 'case' marking set used for all other positions of grammatical constructions except for the actor/agent of the inverse voice and in genitive constructions, in which case it is called the 'genitive case', and except for the instrument case. See Quick 2003 for a full discussion of the voice system including the reasons why the Pendau language does not have an ergative system.

Table 1. Pronouns and noun phrase markers in Pendau

		Absolute ³	Genitive ⁴	Instrument	IV Pronominal Prefixes
SG.	1	a'u	= 'u		'u-, no'u- (irrealis, realis)
	2	oo	= mu		mu-
	3	io	= nyo		--
PL.	1 INC	ito	= to		--
	1 EXC	ami	mami		--
	2	emu	miu		--
	3	jimo	nijimo		--
Proper Nouns		si =	ni =	--	
Common Nouns		∅	nu =	nu =	

Examples (5)–(7) show the marked word order places the ‘pivot’ after the verb and another core argument. Example (5) illustrates the *nong-* verb construction and examples (6)–(7) illustrate the *ni-* verb construction.

- (5) *Nongkomung asu jimo ono mbengimo ri'uo.*
N-pong-'omung asu jimo ono mbengi = mo ri = 'uo
 RE-SF/PT-take dog 3PL/AB when night = COMP LOC = yonder

‘They took dogs over there when it was already night.’

- (6) *Tarus nilolo niinanyo unga uo.*
tarus ni-lolo ni = ina = nyo unga 'uo
 continue IV/RE-search PN/GE = mother = 3SG/GE child yonder

‘Her mother continued looking for her daughter (lit. child).’

- (7) *Paey rasaur miu ami.*
paey ro-saur miu 'ami
 and.then IV/IR-defeat 2PL/GE 1PL.EXC/AB

‘And then you all will defeat us.’

Pendau transitive clause types can be summarized as below:

³ The absolute case set should not be confused with the ergative ‘absolute’ case set, it is only a coincidence that the terms appear to be similar.

⁴ The genitive pronoun set also includes the fronted pronouns ‘u- and mu- for 1st and 2nd person respectively, effectively becoming verbal prefixes. The genitive pronoun set is a mixed set, some are enclitics, and some are free words (distinguishable by phonological criteria).

Pendau has two transitive verb forms distinguished by *nong-* and *ni-* prefixes.⁵ These verb constructions both have A and P arguments.

<i>nong-</i>	AVP or VPA word order Absolute is used in A and P positions Genitive rarely occurs in the P position
<i>ni-</i>	PVA or VAP word order ⁶ Absolute is used in the P position Genitive is used in the A position

The *ni-* verb construction indicates that A is in the non-pivot position, and the P is in the pivot position. The *nong-* verb construction indicates that A is in the pivot position, and the P is in the non-pivot position. The choice between *ni-* and *nong-* verb constructions seems to be dependent on the degree of topic continuity. The *ni-* verb construction seems to be the favoured verb construction when the A argument has a low referential distance (nearly a 3:1 ratio; for discussion and evidence of voice selection criteria see Quick 2002, 2003, in press).

3. Identification of the VP

Figure 1 compares the etic word orders for basic transitive clauses and their associated transitive verb affixes in Pendau. Each verb type has a *rigid* argument position that is postverbal,⁷ and each verb type has a *flex* argument position that is in either (a) a pre-verbal position or (b) in a post-verbal position which must follow the *rigid* argument position. The *flex* positions are marked in Figure 1 by circles around the arguments which have more than one word order position. However, what is relevant is that this pattern suggests that both the *nong-* verb clause and the *ni-* verb clause have one single underlying word order (the emic word order). The *flex* position is identified as that of the *pivot* since preverbally this is the same position the *pivot* occurs in relative clauses, and the *rigid* position as that of the *non-pivot*. The emic word order variation is a pragmatic discourse function that is discussed in §7.

⁵ There are other AV and IV prefixes, however these are the two most commonly encountered and represent the full range of possibilities. Also note that *nong-* is a short hand of the fused form of *N-* and it's underlying stem former *pong-*. In this section I also only refer to the realis form of affixes for convenience, and the reader should note that these all also may appear in the irrealis.

⁶ I am not including the inverse constructions in which the pronominal prefix functions also as inverse voice. These can be annotated as: PA-V and A-VP. The A represents the pronominal prefix in an inverse voice construction.

⁷ Floating adverbs and serial verbs may occur between the verb and the rigid argument in either voice. Serial verbs are by definition part of the verb event and therefore do not invalidate this analysis. Although the floating adverb *moje* 'also, too' and the floating quantifier *jojoo* 'all' may also occur between the verb and the rigid argument they are a special exception and do not invalidate this analysis.

1.	A	nong-V	P	
2.		nong-V	P	A
3.	P	ni-V	A	
4.		ni-V	A	P

Figure 1. A and P argument positions in Pendau transitive clauses

The four etic transitive word orders AVP/VPA and PVA/VAP can be conflated into two emic word orders if we assume that the similarity of pivot and non-pivot positions (or flex and rigid positions) captures an emic word order pattern (Figure 2). The best candidates for these emic word patterns are the grammatical relations subject and object. This would mean that there are two basic transitive patterns we can initially posit as SVO and VOS. These two word orders in fact correlate with the single argument positions of intransitive clauses which occurs as SV and VS word order positions.

The subject can be defined as the syntactic clause's pivot. The initial evidence is provided by the conflation of the etic word order into the emic word order (Quick 2003). This is based on identifying the pivot as a flex position versus the rigid post-verbal non-pivot position. This conclusion is reached via the fact that the verb prefix assigns a semantic role to the NP which has variable or a flexible word order position together with the fact that the etic word order difference between active and inverse voice clauses can be captured or conflated as one emic word order (see Quick 2003 for complete discussion). This evidence indicates that if there is a VP in Pendau, then the VP must be the verb and its dependent argument in the rigid position. This also points to the flex position as being the pivot/subject. This notion of subject is supported by what Manning (1996) calls 'grammatical subject' in contrast to the 'a-structure subject.' Figure 2 gives the tree diagrams for the two word orders.

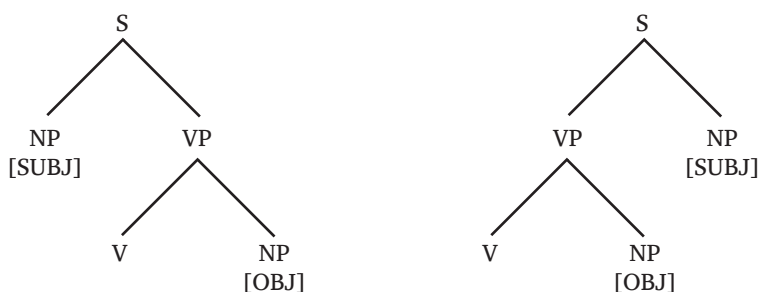


Figure 2. The VP and SVO/VOS word orders

The basic SVO word order is based on several facts. This order is the only order which occurs in relative clauses, and it is also the more frequent word order. Quick (2003, in press) shows that for one story SV/SVO occurs 64% of the time and the VS/VOS order occurs 36% (for both transitives and intransitives).

Additional evidence for the VP is found with conjoined clauses as in (8).

- (8) *Unga miu tonialap nutoo ape nipiarā*
unga miu to = ni-alap nu = too ape ni-piara
 child 2PL/GE RM = IV/RE-get CN/GE = people or IV/RE-care

nutoo.

nu = too

CN/GE = people

‘Someone got your child or someone took care of (your child).’

4. The VP and Double Objects

In this section we will examine double object constructions and the evidence they provide for a VP in Pendau. Examples (9)–(10) illustrate double object constructions in both the *nong*- verb construction and the *ni*- verb construction. The examples show the default word order.

- (9) *A’u mongolia’ io vea.*
a’u M-pong-oli-a’ io vea
 1SG/AB IR-SF-buy-TZ 3SG/AB raw-rice
A = Pivot **P** **3rd argument**
Agent **Recipient** **Theme**
SUBJECT **OBJECT** **2nd OBJECT**

‘I will buy him/her rice.’

[EN97-003.59]

- (10) *Io niolia’o’u vea.*
io ni-oli-a’=’u vea
 3SG/AB IV/RE-buy-TZ = 1SG/GE raw-rice
P = Pivot **A** **3rd argument**
Recipient **Agent** **Theme**
SUBJECT **OBJECT** **2nd OBJECT**

‘I bought him/her rice.’

[EN97-003.59]

The evidence that double object constructions gives for a VP is the fact that the second object can never occur between the grammatical object and the verb, although it can occur in nearly any other position. When the grammatical subject occurs post-verbally the second object may also occur between the grammatical object and the subject. The second object can also be fronted preverbally, but it cannot occur between the grammatical subject and the verb. The more common positions are diagrammed in (11), where O2 refers to the second object and the comma indicates an intonation pause. These orders may occur in either active voice or inverse voice constructions. The addition of prepositional phrases is not included, but it increases the number of possible variations without changing anything of significance in relation to the discussion on VPs. Regardless of the pivot in a ditransitive clause the A and P arguments must always maintain their relative linear position

- (11) S V O O2 *default order*
 V O S O2 *subject occurs between two objects, O2 in default position*
 O2 V O S *second object is fronted preverbally*
 V O O2 S *second object is fronted postverbally*

The examples in (12)–(14) shows that the second object *vea* ‘raw-rice’ can vary in all postverbal positions except between the verb and the grammatical object which is marked as ‘ill-formed.’ Example (15) demonstrates that the meaning changes when an attempt is made to swap the order of the pronouns.

- (12) *Nongolia’* *io* *a’u* *vea.*
N-pong-oli-a’ *io* *a’u* *vea*
 RE-SF/PT-buy-TZ 3SG/AB 1SG/AB raw-rice

‘I bought him rice.’

- (13) *Nongolia’* *io* *vea* *a’u.*
N-pong-oli-a’ *io* *vea* *a’u*
 RE-SF/PT-buy-TZ 3SG/AB raw-rice 1SG/AB

‘I bought him rice’

- (14) **Nongolia’* *vea* *io* *a’u.*
N-pong-oli-a’ *vea* *io* *a’u*
 RE-SF/PT-buy-TZ raw-rice 3SG/AB 1SG/AB

*‘I bought him rice.’

- (15) *Nongolia’* *a’u* *io* *vea.*
N-pong-oli-a’ *a’u* *io* *vea*
 RE-SF/PT-buy-TZ 1SG/AB 3SG/AB raw-rice

‘He bought me rice.’

Examples (16)–(18) illustrate similar word order possibilities for the inverse voice construction.

- (16) *Nisambalea’omo* *niCeku* *jimo* *manu’* *niYusup.*
ni-sambale-a’=mo *ni=C.* *jimo* *manu’* *ni=Y.*
 IV/RE-butcher-TZ=COMP PN/GE=C. 3PL/AB chicken PN/GE=Y.

‘Ceku butchered Joseph’s chicken for them.’

- (17) *Rusa* *uo* *nisoputa’o’u* *jimo* *riMalawa.*
rusa *’uo* *ni-soput-a’=’u* *jimo* *ri=Malawa*
 deer yonder IV/RE-shoot-TZ=1SG/GE 3PL/AB LOC=Malawa

‘I shot that deer for them at Malawa.’

- (18) *Niatora'onyo* *teule* *ma'o* *junjungonyo* *unga* *uo*.
ni-ator-a' = *nyo* *teule* *ma'o* *junjung = nyo* *unga* *'uo*
 IV/RE-deliver-TZ = 3SG/GE return go house = 3SG/GE child yonder

'He took (lit. delivered) that child home to his house.'

5. The VP and the floating Adverb *moje* 'also, too'

Adverbs such as *moje* 'again, also' are floating adverbs. The term 'floating' is not used in a technical sense, but in the looser sense of word order variability.⁸ Examples (19)–(25) illustrate the possible positions that floating adverbs may occur in (in these examples the adverb *moje* 'again, also' is used). Floating adverbs must normally occur in a post-subject constituent position (when it precedes the subject it is distinctively topicalization and requires an intonation pause, see Quick 2003). When the floating adverb occurs in these positions there is no semantic change, and the scope is over the predication itself and not the entire proposition. The adverb *moje* is in bold font and floats or moves in the positions of the otherwise same clauses from right to left. Examples (19)–(21) illustrate the three positions that an adverb may occur in with an inverse voice transitive clause.⁹

- (19) *Ami* *rimoo* ***moje*** *nidua'* *nubali*.
'ami *ri = moo* ***moje*** *ni-dua'* *nu = bali*
 1PL.EXC/AB LOC = this also IV/RE-arrive CN/GE = enemy

'The enemy again arrived here beside us.'

- (20) *Ami* *rimoo* *nidua'* ***moje*** *nubali*.
'ami *ri = moo* *ni-dua'* ***moje*** *nu = bali*
 1PL.EXC/AB LOC = this IV/RE-arrive also CN/GE = enemy

'The enemy again arrived here beside us.'

- (21) *Ami* *rimoo* *nidua'* *nubali* ***moje***.
'ami *ri = moo* *ni-dua'* *nu = bali* ***moje***
 1PL.EXC/AB LOC = this IV/RE-arrive CN/GE = enemy also

'The enemy again arrived here beside us.'

Examples (22)–(25) illustrate the four positions that adverbs in ditransitive active voice instrument clauses may occur in.

⁸ See Donohue 1995:173-175 for an example of floating adverbs in *Tukang Besi* and 'launching' that occurs according to criteria "that is relevant to a non-nominative-argument." At this point in time there has been no observable reason that distinguishes the 'floating' in Pendau, but like *Tukang Besi* it is possible to nominate a canonical position which in Pendau is preverbal (but postverbal in *Tukang Besi*).

⁹ Non-floating adverbs most commonly occur in the preverbal position, see Quick 2003.

- (22) *SiYusup* **moje** *monyambale* *japing* *uo* *nupiso*.
si = Yusup **moje** *M-pong-sambale* *japing* *'uo* *nu = piso*
 PN/AB = Joseph also IR-SF/PT-butcher cow yonder INSTR = machete

'Joseph also butchered the cow with the machete.'

- (23) *SiYusup* *monyambale* **moje** *japing* *uo* *nupiso*.
si = Yusup *M-pong-sambale* **moje** *japing* *'uo* *nu = piso*
 PN/AB = Joseph IR-SF/PT-butcher also cow yonder INSTR = machete

'Joseph also butchered the cow with the machete.'

- (24) *SiYusup* *monyambale* *japing* *uo* **moje** *nupiso*.
si = Yusup *M-pong-sambale* *japing* *'uo* **moje** *nu = piso*
 PN/AB = Joseph IR-SF/PT-butcher cow yonder also INSTR = machete

'Joseph also butchered the cow with the machete.'

- (25) *SiYusup* *monyambale* *japing* *uo* *nupiso* **moje**.
si = Yusup *M-pong-sambale* *japing* *'uo* *nu = piso* **moje**
 PN/AB = Joseph IR-SF/PT-butcher cow yonder INSTR = machete also

'Joseph also butchered the cow with the machete.'

Some adverbs such as *sura* 'only' are restricted to where they can float to as examples (26)–(30) illustrate. In these examples *sura* 'only' can occur only in two positions, either before the verb or before the instrument phrase (examples (28) and (29) contrast the instrument phrase with the prepositional phrase with the instrument).

- (26) *SiYusup* **sura** *monyambale* *japing* *uo* *nupiso*.
si = Yusup **sura** *M-pong-sambale* *japing* *'uo* *nu = piso*
 PN/AB = Joseph only IR-SF/PT-butcher cow yonder INSTR = machete

'Joseph will only butcher the cow with a machete.'

- (27) **SiYusup monyambale sura japing uo nupiso*.

- (28) *SiYusup* *monyambale* *japing* *uo* **sura** *nupiso*.
si = Yusup *M-pong-sambale* *japing* *'uo* **sura** *nu = piso*
 PN/AB = Joseph IR-SF/PT-butcher cow yonder only INSTR = machete

'Joseph will only butcher the cow with a machete.'

- (29) *SiYusup monyambale japing uo sura sono piso.*
si = Yusup M-pong-sambale japing 'uo sura sono piso
 PN/AB = Joseph IR-SF/PT-butcher cow yonder only with machete

‘Joseph will only butcher the cow with a machete’

- (30) **SiYusup monyambale japing uo nupiso sura.*

Since other languages such as English also have adverbs similar to *moje* ‘also, too’ in the VP this will not be considered to be evidence against a VP (see Kroeger 2004). On the other hand, the restrictions on other adverbs from occurring in this same position, such as *sura* ‘only’, indicates that there is a constituent boundary.

6. The VP and Obliques

Prepositional phrases normally occur in clause final position as in (31)–(33). When the VP is in the initial position the normal position for a prepositional phrase is also in final position as in (34).

- (31) *A'u mongkomung bau rijunjung.*
a'u M-pong-'omung bau ri = junjung
 1SG/AB IR-SF/PT-carry fish LOC = house

‘I will carry the fish to my house.’

- (32) *Rusa uo nisoputa'o'u jimo riMalawa.*
rusa 'uo ni-soput-a' = 'u jimo ri = Malawa
 deer yonder IV/RE-shoot-TZ = 1SG/GE 3PL/AB LOC = Malawa

‘I shot that deer for them at Malawa.’

- (33) *Tavala ni'omuni nikai ribuut.*
tavala ni-'omung-i ni = kai ri = buut
 spear IV/RE-carry-DIR PN/GE = grandfather LOC = mountain

‘The grandfather held the spear on the mountain.’

- (34) *Paey ratabola'oto io uo rirano.*
paey ro-tabol-a' = oto io 'uo ri = rano
 and.then IV/IR-throw-TZ = 1PL.INC/GE 2SG/AB yonder LOC = lake

‘And then we’ll throw him into the lake.’ (monkeys talking about the turtle)

Prepositional phrases can be fronted to highlight or put some prominence on the prepositional phrase, as shown schematically in (35). There are two basic fronting positions: 1) postverbal, and 2) preverbal.

- | | | | | |
|------|-----------|-----------|-----------|---------------------|
| (35) | V | NP | PP | normal |
| | V | PP | NP | postverbal fronting |
| | PP | V | NP | preverbal fronting |

Examples (36)–(37) show the prepositional phrase fronted in two different preverbal positions.

- | | | | | | |
|------|--------------------|-------------|------------|-------------------------|--------------|
| (36) | <i>RiMalawa</i> | <i>rusa</i> | <i>uo</i> | <i>nisoputa'o'u</i> | <i>jimo.</i> |
| | <i>ri = Malawa</i> | <i>rusa</i> | <i>'uo</i> | <i>ni-soput-a' = 'u</i> | <i>jimo</i> |
| | LOC = Malawa | deer | yonder | IV/RE-shoot-TZ = 1SG/GE | 3PL/AB |

'I shot that deer for them at Malawa.'

- | | | | | | |
|------|-------------|------------|--------------------|-------------------------|--------------|
| (37) | <i>Rusa</i> | <i>uo</i> | <i>riMalawa</i> | <i>nisoputa'o'u</i> | <i>jimo.</i> |
| | <i>rusa</i> | <i>'uo</i> | <i>ri = Malawa</i> | <i>ni-soput-a' = 'u</i> | <i>jimo</i> |
| | deer | yonder | LOC = Malawa | IV/RE-shoot-TZ = 1SG/GE | 3PL/AB |

'I shot that deer for them at Malawa.'

Example (38) illustrates fronting of the prepositional phrase to the immediately postverbal position.

- | | | | | | | |
|------|------------|------------|-----------------------|---------------------|-------------|------------|
| (38) | <i>Ila</i> | <i>uo</i> | <i>nitalaunyo</i> | <i>rijunjung</i> | <i>unga</i> | <i>uo.</i> |
| | <i>ila</i> | <i>'uo</i> | <i>ni-talau = nyo</i> | <i>ri = junjung</i> | <i>unga</i> | <i>'uo</i> |
| | from | yonder | IV/RE-leave = 3SG/GE | LOC = house | child | yonder |

'From there he left that child at the house.'

Example (39) shows a complex sentence with three clauses. Each of the clauses is transitive and has *abato* 'grub' as its subject. In the third and final clause of the sentence the prepositional phrase is fronted before the P argument to highlight the location in which the grub is deliberately put, i.e. *engenyoyo* 'his nose'.

- | | | | | | | |
|------|------------------------|--------------|----------------------|--------------|--------------------------|--------------|
| (39) | <i>Bai</i> | <i>uo</i> | <i>ni'itonyo</i> | <i>abato</i> | <i>uo,</i> | <i>tarus</i> |
| | <i>bai</i> | <i>'uo</i> | <i>ni-'ito = nyo</i> | <i>abato</i> | <i>'uo</i> | <i>tarus</i> |
| | like | yonder | IV/RE-see = 3SG/GE | grub | yonder | continue |
| | <i>nialaponyo</i> | <i>abato</i> | <i>uo,</i> | <i>paey</i> | <i>nintamaa'onyo</i> | |
| | <i>ni-alap = nyo</i> | <i>abato</i> | <i>'uo</i> | <i>paey</i> | <i>ni-ntama-a' = nyo</i> | |
| | IV/RE-get = 3SG/GE | grub | yonder | and.then | IV/RE-enter-TZ = 3SG/GE | |
| | <i>riengenyoyo</i> | <i>abato</i> | <i>uo.</i> | | | |
| | <i>ri = enge = nyo</i> | <i>abato</i> | <i>'uo</i> | | | |
| | LOC = nose = 3SG/GE | grub | yonder | | | |

'After that he saw that grub, and then he took the grub, and then he put into his nose the grub.'

The previous examples are all based on natural texts. After an extensive review of my data, I was only able to find a few elicited examples in which obliques and instrument NPs¹⁰ could appear in the VP, i.e. between the V and the O. Examples (40)–(41) illustrate accepted elicited examples of the locative oblique phrase occurring in the VP of an active voice and inverse voice construction respectively. Example (42) illustrates the comitative *sono* ‘with’ inside the VP of another elicited example.

- (40) *A’u monyoput riMalava rusa uo.*
a’u M-pong-soput ri = Malava rusa ’uo
 1SG IR-SF/PT-shoot LOC = Malava deer yonder

‘I will shoot the deer in Malava.’

- (41) *Japing uo nisambale riMalava niYusup*
japing ’uo ni-sambale ri = Malava ni = Yusup
 cow yonder IV/RE-butcher LOC = Malava PN/GE = Joseph

nupiso.
nu = piso
 INSTR = machete

‘Joseph used a machete to butcher that cow in Malava.’

- (42) *Japing uo nisambale sono piso niYusup*
japing ’uo ni-sambale sono piso ni = Yusup
 cow yonder IV/RE-butcher COM machete PN/GE = Joseph

riMalava.
ri = Malava
 LOC = Malava

‘Joseph butchered that cow with a machete in Malava.’

Examples (43)–(44), also elicited, show that it is possible to place the instrument noun phrase between the A argument and the verb.

- (43) *SiYusup monyambale nupiso japing uo.*
si = Yusup M-pong-sambale nu = piso japing ’uo
 PN/AB = Joseph IR-SF/PT-butcher INSTR cow yonder

‘JOSEPH will slaughter the cow with a knife.’

¹⁰ In Quick 2003 I analyze instrument NPs as second objects, although the evidence is ambiguous whether they should be treated as obliques or as core arguments.

- (44) ...*paey unga nirembasi nuuram niYusup.*
paey unga ni-rembas-i nu = uram ni = Yusup
 and.then child IV/RE-hit-DIR INSTR = medicine PN/GE = Joseph

‘...and then Joseph hit the CHILD with the medicine.’

The evidence for fronted obliques to occur within the VP is weak as the only examples so far have come from elicited clauses. Further research is necessary, but at this stage it looks like obliques are not found in natural texts within the VP because it would break up this VP constituent. On the other hand, if further research shows that obliques are found inside the VP in natural texts, this in itself is not enough evidence to counter the other evidence for a VP. Lexical Functional Grammar (LFG) for example allows discontinuous VPs (Bresnan 2001:126):

The noncompositionality of LFG thus implies that VPs can be discontinuous phrases whose heads may appear external to the rest of the phrase.

7. Word Order Choice, Topic-Comment and the VP

In this section I want to consider briefly what the basis for selecting between the two pragmatic word orders may be. At this point it seems likely that when the VP occurs before the S that we can consider this to be VP fronting rather than saying the S moves. In the previous section I examined briefly oblique fronting which places obliques into a position of pragmatic prominence (see Quick 2003 for second object fronting). There is reason to believe that the VP is part of a topic-comment framework which may explain this variable word order. Van Valin and LaPolla suggest (1997:218),

It is reasonable to suppose, then, that the universal basis for the language-specific phrasal category VP is focus structure.

In Quick (in press) I checked for topic continuity of the S in both positions, and found that in one text there was no significant difference that was based on the topic continuity (i.e. for referential continuity). The only distinction found demonstrated that the full noun phrase is favored in about a 3:1 ratio when the subject is in the post-verbal position. For future research I suggest that based on those statistics and the evidence presented in this paper for a VP in Pendau that VP fronting occurs for one or more of these reasons:

- Thematic continuity and/or action continuity (see Dooley and Levinsohn 2001, Levinsohn 2003)
- Pivotal storyline or primary event (see Longacre 1989, Dooley and Levinsohn 2001)
- Heavy NPs preferred in a postverbal S position

Since fronting of other constituents, such as the obliques and second objects can occur in the same construction as a fronted VP, the two types of discourse functions cannot obviously be for the same purposes when they appear in a clause at the same time.

8. Conclusion

In conclusion the data shows that there is a VP in Pendau. This is obtained from basic word order and other constituents that either can't occur within the VP or seem to be unnatural inside the VP. This major finding points to the probability that the variation in word order is due to fronting of the VP, which is itself a pragmatic discourse function. Future research addressing the VP as part of a topic-comment construction should provide productive results which can provide practical solutions such as helping to translate material naturally.

Identifying that the AV and IV constructions each have a constituent which behaves the same way is striking evidence that this constituent is a VP. What is especially striking is the fact that the NP argument inside this constituent can be either an actor or a non-actor argument. The following list summarizes the restrictions which are identical for both voice constructions:¹¹

- The O cannot move, it is a 'rigid' argument and must occur immediately following the V. It must be [VO]_{VP}
- The S (or 'flex' argument) may not appear between the V and O, although it may appear in almost any other word order position
- The O2 (second argument) may not intervene between the V and the O
- Obliques do not naturally occur between the V and the O
- Nuclear directional serial verbs follow the V1 and occur immediately before the O
- Core directional serial verbs do not occur inside the VP and must follow the O somewhere outside the VP

Exceptions to the rigid word order in the VP also show the same behaviour and pattern for both voices. These exceptions were discussed and not considered to bear weight against the evidence for a VP.

References

- Arka, I Wayan. 2003. *Balinese morphosyntax: a lexical-functional approach*. Canberra: Pacific Linguistics.
- Bresnan, Joan. 2001 *Lexical-functional syntax*. Oxford: Blackwell.
- Campbell, Philip J. 1989. Some aspects of Pitu Ulunna Salu grammar: A typological approach. MA thesis. The University of Texas at Arlington.
- Donohue, Mark. 1995. The Tukang Besi language of Southeast Sulawesi, Indonesia. PhD dissertation. Canberra: Australian National University.

¹¹ Serial verbs were not discussed here. For a general discussion on serial verbs in Pendau see Quick 2003.

- Donohue, Mark. 1999. *A grammar of Tukang Besi*. Berlin: Mouton De Gruyter.
- Dooley, Robert A. and Stephen H. Levinsohn. 2001. *Analyzing discourse: a manual of basic concepts*. Dallas: Summer Institute of Linguistics
- Evans, Donna. 2003. Tata bahasa Kaili-Ledo yang singkat. In *Kamus Kaili-Ledo Indonesia Inggris*, Donna Evans, pp. 485–543. Palu, Indonesia: Pemerintah Daerah Propinsi Sulawesi Tengah, Dinas Kebudayaan dan Pariwisata.
- Givón, Talmy, ed. 1994. *Voice and Inversion*. Amsterdam: John Benjamins.
- Givón, Talmy. 2001. *Syntax an introduction (vol II)*. Amsterdam: John Benjamins.
- Himmelman, Nikolaus. 2002. Voice in Western Austronesian: An update. In Fay Wouk and Malcolm Ross, eds., *The History and Typology of Western Austronesian Voice Systems* (PL 518). Canberra: Pacific Linguistics.
- Himmelman, Nikolaus. 2005. The Austronesian languages of Asia and Madagascar: typological characteristics. In *The Austronesian Languages of Asia and Madagascar*, Alexander Adelaar and Nikolaus P. Himmelman, eds., 110–181. London: Routledge.
- Kroeger, Paul. 2004. *Analyzing syntax a lexical-functional approach*. Cambridge: Cambridge University Press.
- Levinsohn, Stephen H. 2003. Analysis of narrative texts. Manuscript, SIL International.
- Longacre, Robert. 1989. Two hypotheses regarding text generation and analysis. *Discourse Processes* 12:413–460.
- Manning, Christopher D. 1996. *Ergativity argument structure and grammatical relations*. Stanford, California: CSLI Publications.
- Quick, Philip A. 1997. Active and inverse voice selection criteria in Pendau, a Western Austronesian language. In *Proceedings of the Seventh International Conference on Austronesian Linguistics*, Cecilia Odé and Wim Stokhof, ed. Amsterdam: Editions Rodopi B.V.
- Quick, Philip A. 1999. A panoramic view of 17 pV(C)- formatives in Pendau distinguished by the paradigmatic landscape. In *Studies in Sulawesi Linguistics*, NUSA 46, Part VI, Wyn Laidig, ed, 95–117.

- Quick, Philip A. 2002. A sketch of the primary transitive verbs in Pendau. In Fay Wouk and Malcolm Ross, eds., *The History and Typology of Western Austronesian Voice Systems* (PL 518). Canberra: Pacific Linguistics.
- Quick, Philip A. 2003. A grammar of the Pendau language. PhD dissertation. Canberra: Australian National University.
- Quick, Philip A. 2005a. Augmented stem forming morphology in Pendau. Paper presented at KOLITA 3 (Konferensi Linguistik Tahunan Atma Jaya: Tingkat Nasional), in unpublished proceedings, 111–116, Jakarta, Pusat Kajian Bahasa dan Budaya, Unika Atma Jaya.
- Quick, Philip A. 2005b. Topic continuity, voice, and word order in Pendau. In *The many faces of Austronesian voice: Some new empirical studies*, Wayan Arka and Malcolm Ross, eds., (PL 571) 209–228. Canberra: Pacific Linguistics.
- Quick, Philip A. In press. Applicatives and double objects in Pendau. In *Studies in Sulawesi Linguistics*, NUSA 49, Part VII, Wyn Laidig, ed.
- Ross, Malcolm. 2002a. The history and transitivity of western Austronesian voice and voice-marking. In Fay Wouk and Malcolm Ross, eds., *The History and Typology of Western Austronesian Voice Systems* (PL 518). Canberra: Pacific Linguistics.
- Ross, Malcolm 2002b. Final words: research themes in the history and typology of western Austronesian languages. In Fay Wouk and Malcolm Ross, eds., *The History and Typology of Western Austronesian Voice Systems* (PL 518). Canberra: Pacific Linguistics.
- Van Valin, R.D. and R.J. La Polla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.

Rhythm in Bidayuh*

Calvin R. Rensch
SIL International

The approximately 25 dialects of Bidayuh (Land Dayak) spoken in Sarawak, Malaysia, can be grouped into four clusters of generally intelligible dialects: Eastern, Central, Highland and Western. Although there is considerable variation among the clusters, the phonological patterns are very similar

In both the ultima and penult there are differences of vowel duration. The ultima vowels can occur as normal or as lengthened. In most dialects there are two types of penult vowel, full and reduced. A full penult vowel is sustained and has consistent quality. A reduced vowel frequently is brief, has lax articulation and lengthens a following consonant. In certain environments it is deleted entirely.

*The lengthened vowels in the ultima have at least two historical sources: 1) loss of a Proto Bidayuh medial *h; 2) Proto Bidayuh lengthened vowels that developed from loss of an Austronesian medial laryngeal.*

The full vs. reduced vowel contrast seems to be related to differences in accentual pattern in some other AN languages and apparently developed from early accentual differences.

1. Introduction

Bidayuh, a Land Dayak language, is a network of dialects spoken in the inland sections of the Kuching and Samarahan Divisions in southwestern Sarawak. The approximately 25 dialects can be grouped into four clusters of generally intelligible dialects: Eastern Bidayuh (Bukar-Sadung), Central (Biatah-Penrissen-Lower Padawan), Highland (Tringgus-Sembaan-Upper Padawan) and Western (Singai-Jagoi).¹

In addition to Bidayuh the Land Dayak family of languages includes both the rather closely-related Bakati' group and a number of other languages spoken farther south in West Kalimantan.

* The author wishes to acknowledge the leaders of the Bidayuh Language Development Project for their facilitating and encouraging the collection of the Bidayuh language data that forms the basis for this analysis as well as dozens of Bidayuh friends who have generously devoted hours to answering many questions about their variety of Bidayuh.

¹ The dialect groups correspond roughly to political districts (Eastern to Serian District, Central to Kuching District and Western to Bau District), but the Highland dialects are spoken in highland areas of both Kuching and Bau Districts.

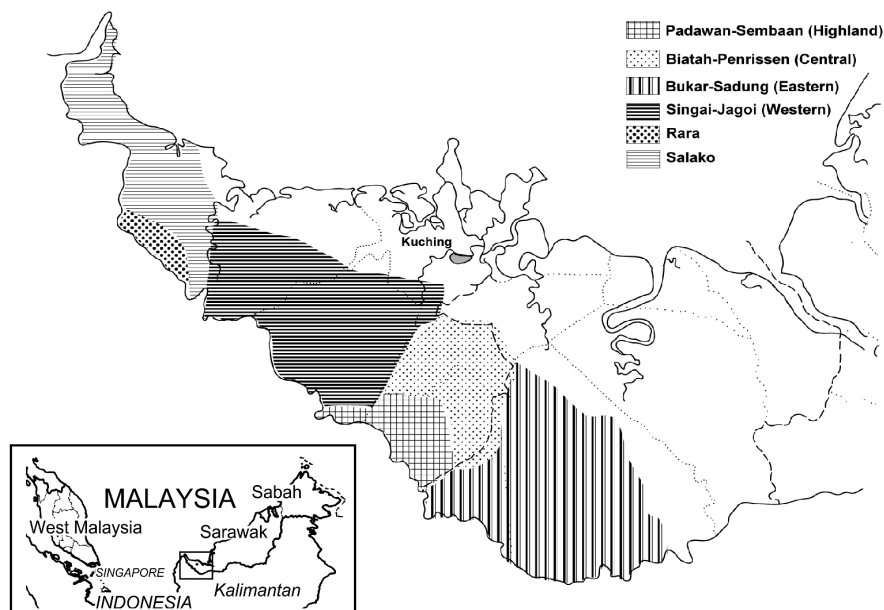


Figure 1. Location of Bidayuh dialect clusters with Rara and Salako (inset shows location of Bidayuh regions within Malaysia and Borneo)

2. Rhythm differences

Bidayuh stems are typically disyllabic. Stress falls on the ultima of the Bidayuh word, at least in isolation.

In both the ultima and the penult there are differences of vowel duration that create contrasting rhythm patterns. The longer (L) and shorter (S) vowels in both ultima and penult occur in all four possible combinations: S-S, L-S, S-L, L-L. Consider the following examples from the Tringgus Bireng (Highland) and Tebakang (Eastern) dialects:

Tringgus Bireng	<i>sūkun</i> ² ‘elbow’	<i>pūkut</i> ‘punch (someone)’
	<i>būko:ʔ</i> ‘parang, bush knife’	<i>sūka:n</i> ‘hide’
Tebakang	<i>kābəs</i> ‘dead’	<i>kābət</i> ‘tie (impv.)’
	<i>mābə:r</i> ‘fly (verb)’	<i>tābi:r</i> ‘wall’

2.1 Vowels of the ultima

In the ultima there are six vowel qualities: high *i*, *u*; central *ə*; mid *e*, *o*; low *a*. In nearly half the dialects the mid vowels are pronounced as centralizing diphthongs, [iə/iə] and [ua/uə]; elsewhere they are pronounced as [e/ɛ] and [o/ɔ]. The central

² The consonants of Bidayuh are voiceless stops *p*, *t*, *c* (contrastive in Eastern dialects only), *k*; voiced stops *b*, *d*, *j*, *g*; nasals *m*, *n*, *ɲ*; liquids *r*, *l* (contrastive in Eastern dialects only); semivowels *w*, *y*; fricative *s*; laryngeals *ʔ*, *h*.

vowel is pronounced as [i] except in the Jagoi, Penrissen and Lower Padawan dialects, where it has a more retracted pronunciation, [ɤ].

Each of the six vowels in the ultima can occur as normal (V) or as lengthened (V:)³ although the mid and low vowels are lengthened more commonly than the others. In some dialects it is difficult to be sure whether the contrast occurs with all the vowels. Curiously, it appears that the Tringgus Raya (Highland) dialect has no length contrast in final syllables even though the other Highland dialects do.⁴

Bunan (Eastern) *tuʔa:n*, Biatah (Central) *tuʔa:n*, Sembaan (Highland)
tuʔa:n, [Serambu (Western) *tuʔan*] ‘primary forest’.

Tebakang (Eastern) *sibo:ŋ*, Biatah (Central) *sibo:ŋ*, Tringgus Bireng
(Highland) *sibo:ŋ*, [Singai (Western) *siboŋ*] ‘sheath for a bush knife’.

2.2 Vowels of the penult

In the penult there are four vowel qualities: high i, u;⁵ central ə;⁶ low a. The mid vowels e and o do not occur contrastively in the penult.

It is worth noting in passing that the antepenult in Bidayuh has just one vowel i,⁷ which does not seem to exhibit either the lengthening found in the ultima or the full/reduced difference found in the penult.

In Eastern, Central and Highland (but not Western) dialects there are two types of penult vowels, full (V̄) and reduced (V̆). Note the following contrasts in selected varieties:

Biatah (Central)	<i>bāʔuh</i> ‘new’, <i>băʔuh</i> ‘eagle’
	<i>dāya:ʔ</i> ‘person’, <i>dăya:ʔ</i> ‘blood’
Bukar (Eastern)	<i>āsuh</i> ‘smoke’, <i>āsih</i> ‘who’
Anah Rais (Highland)	<i>bīsaʔ</i> ‘can, be able’, <i>bīsaʔ</i> ‘wet’
Biya (Highland)	<i>tūroh</i> ‘egg’, <i>tūrəŋ</i> ‘bone’
Tringgus Bireng (Highland)	<i>sūka:n</i> ‘hide’, <i>sūkuh</i> ‘elbow’

³ In Western Bidayuh vowels may be rearticulated, i.e., there are sequences of identical (or diverse) vowels with the onset of the second vowel marked by an articulatory pulse coinciding with the onset of stress. For example, Gumbang (Western) [ta.'as] ‘ironwood’, [ti.'iʔ] ‘itchy’. These are sequences of vowels that form the nucleus of successive syllables. Western Bidayuh does not have lengthened vowels in the usual sense.

⁴ For a discussion of long consonants as well as long vowels in the ultima of Long Terawan Berawan, a Lower Baram language of northern Borneo, see Jürgen Burkardt, “Long Terawan Berawan phonology: Questions on diphthongs and syllabicity” in this same volume.

⁵ As penult vowels i and u can vary in some dialects to e and o, respectively, especially when the vowel of the ultima is e or o.

⁶ The mid vowel ə (in non-Western dialects) is more restricted than the other vowels in its occurrence in that it may occur in the penult only if the vowel of the ultima is also a mid vowel, e.g., Bukar (Eastern), Biatah (Central) *təʔəp*, Tringgus Raya (Highland) *təʔəp* ‘deep’; Bukar (Eastern), Biatah (Central) *bəʔəŋ*, Tringgus Bireng *bəʔəŋ* (Highland) ‘round’.

⁷ The other Land Dayak languages also have just one vowel that occurs in the antepenult, but in the Bakati’ languages and some others farther south it is a low vowel, a, rather than i.

The following examples show that the same type of penult vowel is often found in the equivalent forms in various dialect areas:

- Tebakang (Eastern) *bāuh*,⁸ Biatah (Central) *bāʔuh*, Tringgus Raya (Highland) *bāʔuh*, [Serambu (Western) *baʔuh*] ‘new’
 Tebakang (Eastern) *bǎʔuh*, Biatah (Central) *bǎʔuh*, Tringgus Raya (Highland) *bǎʔuh*, [Singai (Western) *bəʔuh*] ‘eagle’
 Tebakang (Eastern) *mūkut*, Benuk (Central) *mūkut*, Sembaan (Highland) *mūkut*, [Singai (Western) *bukut*] ‘punch (vb.)’
 Tebakang (Eastern) *būko:ʔ*, Benuk (Central) *būko:ʔ*, Tringgus Bireng (Highland) *būko:ʔ*, [Singai (Western) *bukoʔ*] ‘bush knife’

A full penult vowel has normal duration and is sustainable, with consistent vowel quality. Although there is no single phonetic feature that is present in every instance of a reduced penult vowel, the following features most commonly characterize this type of vowel:

- quick articulation that is not sustainable
- lengthening of a following consonant; cf. Tringgus Bireng /*sūkuh*/ [suk:uh] ‘elbow’
- devoicing before a voiceless stop; cf. Biya /*prǎkis*/ [prahkis] ‘boil, infection’, /*ǎtək*/ [ʔatik] ‘brain’
- deletion, in at least some utterances, especially between a stop and a liquid or between *s* and a stop or nasal; cf. /*tūraŋ*/ > Benuk /*traŋ*/ [tra^hŋ] ‘bone’; /*sīpo:t*/ > Sembaan /*spo:t*/ [spo:t] ‘blowpipe’, /*sīno:d*/ > Benuk /*sno:d*/ [snu:^hd] ‘comb’
- lax vowel quality; e.g., *i* > *I*, *u* > *u*, *a* > *ə*; cf. Bukar /*sīkuh*/ [sɪk:uh] ‘elbow’, Bukar /*būda:ʔ*/ [bʊd:a:ʔ] ‘white’, Biatah /*dārəd*/ [dərɪd] ‘mountain’
- lowering of high vowels; e.g., *i* > *e*, *u* > *o*, cf. Tebakang /*jūleh*/ [jɛl:ɛh] ‘tongue’, Tebakang /*tūlo:ʔ*/ [tol:o:ʔ] ‘egg’
- exchange of vowel quality, especially among high vowels; e.g., *i* > *u*, *u* > *i*, *u* > *ə*, cf. /*sīkuh*/ > Anah Rais /*sīkuh*/, /*sūkuh*/ ‘elbow’, /*sūdeʔ*/ > Tringgus Raya /*sīdeʔ*/ ‘younger sibling’, /*mūrib*/ > Tringgus Raya /*mārib*/ ‘to fly’

The predominant canonical pattern in Bidayuh stems is CVNCVC.⁹ Each of the consonant positions is optional. That is, the penult may have no initial consonant, and there may be no medial consonant. If the penult is closed by N, it must be followed by an initial consonant in the ultima. The ultima may have no closing consonant.

The contrast between full and reduced penult vowels is neutralized in the patterns CVVC and CVhVC.¹⁰

⁸ A medial ʔ is regularly deleted in Tebakang following a full vowel (but not a reduced vowel) in sequences of diverse vowels.

⁹ C = consonant, V = vowel, N = nasal of the same place of articulation as the following consonant.

¹⁰ Although a similar neutralization might be expected in CVʔVC sequences, the contrast between full and reduced vowels does in fact operate in that environment. For example, Biatah (Central), Tringgus Raya (Highland) and Bukar (Eastern) *bāʔuh* ‘new’, Biatah (Central), Tringgus Raya (Highland) and Bukar (Eastern) *bǎʔuh* ‘eagle’.

When the pattern is CVVC, i.e., when there is no following consonant, the penult vowel is full, as in the following examples:

Anah Rais and Biya (Highland) *nīap* ‘count’
 Biya (Highland), Tebakang (Eastern) *ūi* ‘rattan’
 Biatah (Central), Tringgus Raya (Highland), Sadong (Eastern) *māin* ‘play’
 Tebakang (Eastern) *kāi?* ‘negative’

When the pattern is CVhVC, i.e., when the medial consonant is h, the penult vowel is reduced, as in the following examples:

Tringgus Bireng (Highland) *dīhen*, Tebakang (Eastern) *dīhan* ‘durian’
 Tringgus Raya (Highland) *bāhit*, Tebakang (Eastern) *ku māhit* ‘left (hand)’
 Tringgus Raya (Highland), Bukar (Eastern) *tīhi?* ‘itchy’
 Tringgus Raya (Highland), Bukar (Eastern), Tebakang (Eastern) *māhi* ‘eight’

Pattern A: complete set of full and reduced vowels	high	ī, ĩ	ū, ũ
all Highland dialects	central	ē, ě	
Tebakang (Eastern)	low	ā, ǎ	
Pattern B: contrast in all but central vowel	high	ī, ĩ	ū, ũ
Benuk (Central)	central	ə	
	low	ā, ǎ	
Pattern C: contrast in only the low vowel	high	i	u
Biatah (Central)	central	ə	
some Penrissen (Central)	low	ā, ǎ	
most Eastern dialects			
Pattern D: no contrast	high	i	u
all Western dialects	central	ə	
Lower Padawan (Central)	low	a	
Bunan (Eastern)			

Not all dialects exhibit the same system of full and reduced vowels in the penult. In some groups the contrast occurs with only certain vowel qualities, and in one group the contrast does not operate at all.

3. Sources of Bidayuh rhythmic differences

One may wonder how these rhythmic differences in the Bidayuh ultima and penult have developed. Apparently the differences in the ultima and the penult have different sources.

3.1 Sources of lengthened vowels in the ultima

The lengthened vowels in the ultima seem to have had at least two sources: (a) vowels brought together by the loss of a Proto Bidayuh medial *h and (b) retention of a Proto Bidayuh lengthened vowel, which developed when a medial laryngeal at an earlier stage dropped out.

3.1.1 Loss of Proto Bidayuh medial *h

Some lengthened vowels in Central Bidayuh and the rearticulated vowels in Western Bidayuh, which are not lengthened vowels, have developed where a medial *h of Proto Bidayuh has dropped out.

- Proto Bidayuh **tāhas* > Bukar (Eastern) *tāhas*, Tringgus Raya (Highland) *tāhas*, Biatah (Central) *ta:s*, Serambu (Western) *taas* ‘ironwood’
 Proto Bidayuh **muhun* > Bukar (Eastern) *muhun*, Sembaan (Highland) *muhun*, Biatah (Central) *mu:n*, Serambu (Western) *muun* ‘descend’
 Proto Bidayuh **tihī?* > Bukar (Eastern) *tihī?*, Tringgus Raya (Highland) *tīhī?*, Biatah *ti:?*, Serambu (Western) *tii?* ‘itchy’
 Proto Bidayuh **bāhə(h)* > Sangking (Eastern) *bāhə*, Tringgus Raya (Highland) *bāhə*, Biatah (Central) *əbə:*, Serambu (Western) *bāəh* ‘charcoal’

3.1.2 Loss of Austronesian medial laryngeal

Nearly all non-Western dialects have lengthened vowels that have developed from Proto Bidayuh lengthened vowels, which in turn have developed where an Austronesian (or PMP) medial laryngeal has dropped out.

- PAN **buSek* > Proto Bidayuh *(V)*bo:k* > Sembaan (Highland) *bo:k*, Biatah (Central) *ubo:k*, [Serambu (Western) *bok*] ‘hair (of the head)’
 PMP **bə[n]tiʔis* > Proto Bidayuh **bīte:s* > Tebakang (Eastern) *bīte:s*, Anah Rais (Highland) *bāte:s*, Bistaang (Central) *bāte(:)s*, [Serambu (Western) *bites*] ‘calf (of leg)’

Another Proto Bidayuh form with a lengthened vowel may reflect an earlier form with a medial laryngeal: Proto Bidayuh **ma:n*, **maʔan* > Sembaan (Highland) *ma:n*, Biatah *ma:n*, [Singai *man*]; Bukar (Eastern) *maʔan* ‘eat’. The consonant m of the Proto Hesperonesian reconstruction **amaq* fits well with these forms, but the Proto Austronesian reconstruction **kaʔən*, also glossed ‘eat’, with a medial laryngeal may offer a source for the reconstructed long vowel of Proto Bidayuh. (Note the long vowel in the Maloh form in the following section.)

The proposal that the lengthened vowels of Proto Bidayuh developed where an Austronesian medial laryngeal dropped out is supported by evidence from some other language groupings in Borneo.

Note, for example, the following examples from Maloh/Embaloh (Tamanic) in which long vowels have developed through the loss of a medial *h, *q or *y:

- PAN **buSek* > Maloh *bu:k* ‘hair (of the head)’
 PMP **daqan* > Maloh *da:n* ‘branch’
 PMP **taqu* > Maloh *kato:* ‘right (hand)’
 PMP **kaʔen* > Maloh *aŋ|ka:n* ‘eat’
 PHN **sayap* > Maloh *sa:p* ‘wing’ (Adelaar 1994)

In Salako a medial **l* is often lost, e.g., PMP **bulan* > Salako *buatn* ‘moon’, PMP **kulit* > Salako *kuit* ‘skin’. When the same vowel occurs before and after the liquid, a rearticulated vowel results, which is similar to those of Western Bidayuh and in some contexts is pronounced as a long vowel.

- PAN **Zalan* > Salako *bajaatn* ‘walk’
 PMP **laləj* > Salako *aat* ‘(house)fly’
 PAN **tələn* > Salako *taatn* ‘swallow’
 PAN *dəm* > Salako *maam*, cf. Malay *malam* ‘night’

3.2 Source of full and reduced vowels in the penult

Some scholars, e.g., Zorc (1978, 1983) and Ross (1994:62), suggest that PAN had differences in placement of the accent. The contrast between full and reduced vowels in the Bidayuh penult seems to be related to differences in accentual pattern in some other Austronesian (especially Philippine) languages and apparently has developed from accentual differences in Proto Austronesian.

Specifically, a Bidayuh form with a full vowel has developed from an earlier form with final stress while a Bidayuh form with a reduced vowel has developed from an earlier form with penultimate stress.

Note the following examples of Bidayuh full and reduced vowels that have developed from earlier forms with final and pre-final stress patterns, respectively.¹¹

Forms with PAN final stress

- PAN **batú* > Proto Bidayuh **bātuh* ‘stone’
 PMP **manúk* > Proto Bidayuh **mānuk* ‘bird’
 PAN **limáH* > Proto Bidayuh **riməh* ‘five’
 PAN **bi + bíR* > Proto Bidayuh **bībih* ‘lip’
 PMP **luZáq* > Proto Bidayuh **ŋ-ir-ūja?* ‘spit’
 PHN **buká?* ‘open’ > Proto Bidayuh **būka?* ‘wide’
 PAN **telúR* > Proto Bidayuh **tūloh* ‘egg’
 PAN **belí* > Proto Bidayuh **mīrih*, **bīrih*, ‘buy’

¹¹ A form from Proto Austronesian (PAN) is provided when available. However, a form from Proto Malayo-Polynesian (PMP) is provided when that is the earliest horizon from which a reconstructed source is available. For a fuller set of examples of this contrast see Rensch et al. (2006), pp. 318–322.

Forms with PAN penultimate stress

PAN **káyuH* > Proto Bidayuh **kǎyuh* ‘tree, wood’

PAN **láyit* > Proto Bidayuh **rǎñit* ‘sky’

PAN **d₂áya* ‘inland’ > PLD **dǎyǎh* ‘(up)land’

PAN **búlan* > Proto Bidayuh **bǔran* ‘moon’

PAN **búluq* > Proto Bidayuh **bǔru?* ‘bamboo’

PMP **hútek*, **qútek* > PBB **ǎntək* ‘brain’

PMP **ípen* > Proto Bidayuh **jǐpən* ‘teeth’

PAN **sí + kuH* > Proto Bidayuh **sǐkuh*, etc. ‘elbow’

4. Conclusion

Most varieties of Bidayuh exhibit rhythmic features not shared by many of their Austronesian neighbors. In the ultima there are lengthened vowels that contrast with the considerably more common normal-length vowels. In the penult the full and reduced vowels are distinguished from each other not only through differences in sustainability but through tense vs. lax articulation and even the potential for complete deletion.

These differences in phonetic realization of the two types of contrast are paralleled by differences in historic sources of the contrasts. The lengthened vowels in the ultima derive primarily from lost medial consonants whereas the full-reduced vowel contrast in the penult relates to accentual differences of an earlier period.

References

- Adelaar, K. Alexander. 1994. The classification of the Tamanic languages. In *Language Contact and Change in the Austronesian World*, ed. by Thomas E. Dutton and Darrell T. Tryon, 1–41. Berlin: Mouton de Gruyter.
- Rensch, Calvin R., Carolyn M. Rensch, Jonas Noeb and Robert Sulis Ridu. 2006. *The Bidayuh Language: Yesterday, today and tomorrow*. Kuching: Dayak Bidayuh National Association.
- Ross, Malcolm D. 1994. Some current issues in Austronesian linguistics. In *Comparative Austronesian Dictionary*, ed. Darrell T. Tryon, Vol. 1, 45–120. Berlin: Mouton de Gruyter.
- Zorc, R. David. 1978. Proto-Philippine word accent: innovation or Proto-Hesperonesian retention? In *Papers from the Second International Conference on Austronesian Linguistics*, ed. by S. A. Wurm and Lois Carrington, fascicle 1, 67–113. *Pacific Linguistics* C-61. Canberra: Australian National University.
- Zorc, R. David. 1983. Proto Austronesian accent revisited. *Philippine Journal of Linguistics* 14 (1) 1–24.

Between Actor and Undergoer: The -om- predicates in Kankanaey

Janet L. Allen
SIL International

This study looks at the range of predicates that are formed with the infix -om- in Kankanaey, and attempts to define the conditioning factors that govern its use. In most cases, a particular verbal affix is restricted to cross-referencing specific semantic roles that group together as either Actors or Undergoers, giving rise to such traditional labels as “actor-focus verb” or “object-focus verb.” This study finds that -om- defies this restriction by cross-referencing a special group of Actors and a special group of Undergoers. In Kankanaey, verbal affixes contribute to the agency implicature of the verb’s Actor. The affix -om- is used to license the Actor status of arguments with low or reduced agency or effectiveness. With stative roots, -om- cross-references Undergoer arguments that participate in independent change-of-state events. When a participant is both Actor and Undergoer of an event, or when an Actor is lower in inherent agentivity than the Undergoer, -om- cross-references the Actor. This study concludes that with -om- affixed verbs in Kankanaey, the status of their cross-referenced argument is compromised or modified in some way, occupying a middle ground between typical Actors and typical Undergoers.

1. Introduction

Kankanaey is spoken by 150,000 people in communities in northern Luzon, mostly in Benguet Province. It is among the many Philippine languages that count -om- (-um- in many orthographies) among their verbal affixes. In Kankanaey, verbal affixes are restricted to cross-referencing an argument that fills specific semantic roles. These roles normally group together as either Actors or Undergoers. Examples of Actors include AGENTS, EFFECTORS, and MOVERS, while Undergoers include PATIENTS, THEMES, and GOALS. The focus of this study, the infix -om- (-inom- with completive aspect), does not conform to this restriction, but forms predicates that cross-reference Actors in some contexts, and Undergoers in others. This study looks at the range of verbal predicates that are formed with this affix in Kankanaey, and attempts to define the conditioning factors that govern its use.

The theoretical background is compatible with Role and Reference Grammar as presented in *Syntax* (1997) by Robert D. VanValin, Jr. and Randy J. LaPolla (hereafter: VVLP, 1997). Data for this study was excerpted from texts gathered in Benguet between 1975 and 1996.

2. Theoretical Background

Kankanaey verbal predicates are formed by combining a lexical root with one of several predicating affixes. A lexical root denotes a state of affairs and the possible participants in it. The lexical roots that combine with predicating affixes may be categorized as denoting either dynamic situations (actions), or states, including attributes. In a clause, the predicating affixes cross-reference one argument of the verb, that is, one participant in the state of affairs indicated by the root.¹ Subsections 2.1–2.3 below provide more detail on the different types of clauses.

In this study, the Logical Structure of each verb will be presented in accordance with the lexical representation of the Aktionsart classification system. This indicates the temporal properties and the participants (arguments) in the state of affairs. These arguments may be grouped by the macro roles Actor and Undergoer, as defined by their place in the logical structure of the verb.

In this representational system, as shown in Table 1, boldface type with a prime indicates semantic constants, capital letters indicate modifications to the predicate, and normal type in parentheses indicates the argument variables. Table 1 shows six types of predicates with their logical structures (LS). Note that Activities are represented by the presence of **do'** and that a change of state may be expressed as an Achievement or Accomplishment depending on the time variable, whether instantaneous (INGR) or requiring time (BECOME). Accomplishments express changes with an end point while Process Accomplishments (PROC) are open ended. Combinations of predicates include Active Accomplishments and Causatives of every kind. The table is taken from VVLP (1997:109), with PROC added.

Table 1. Lexical representations for Aktionsart classes

<i>Verb class</i>	<i>Logical Structure</i>
State	predicate' (x) or (x,y)
Activity	do' (x, [predicate' (x) or (x,y)])
Achievement	INGR predicate' (x) or (x,y) or INGR do' (x, [predicate' (x) or (x,y)])
Accomplishment	BECOME predicate' (x) or (x,y) BECOME do' (x, [predicate' (x) or (x,y)]) PROC predicate' (x) or (x,y)
Active accomplishment	do' (x, [predicate' ₁ (x,(y))]) & BECOME predicate' ₂ (z, x) or (y)
Causative	α CAUSE β where α, β are LSs of any type

Of further relevance to this study is the Actor-Undergoer Hierarchy (as taken from VVLP, 1997:146), shown below in figure 1. The hierarchy shows that if there is a **do'** in the representation of a predicate, the first argument is unmarked as the Actor

¹ Another term for this cross-referencing that has been used in many Philippine studies is *focus*.

macrorole (at the left of figure 1), while the single argument of a state predicate (the right of figure 1) is unmarked as an Undergoer. Predicates with two arguments will typically have an Actor as first argument and Undergoer as the second. This study examines predicates formed in Kankanaey by *-om-*, and shows that in these situations, the macrorole is the marked, unexpected choice.

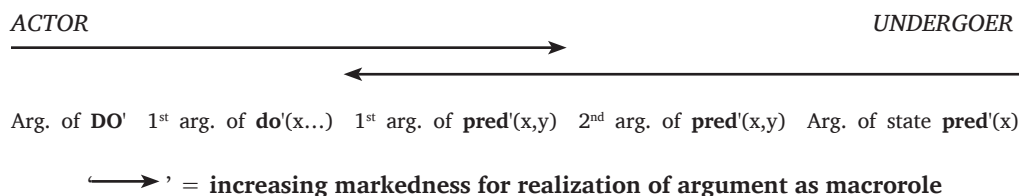


Figure 1. The Actor-Undergoer Hierarchy

2.1. State predicates

Kankanaey State predicates are formed by affixing stative roots. These roots denote inherent or unvarying situations or may denote the result of some causative force without specified intentionality. The typical argument of a State predicate is an inactive, unintentional and totally affected Undergoer.

2.2. Activity predicates

Kankanaey Activity predicates are formed by affixing action roots. Most action roots have lexicalized agency, which typically involves animacy, intention, independent action, and control on the part of the Actor. Exceptions to this general rule include roots that denote physical motion or position. Agency can also include causation, which implies success or effectiveness. The Kankanaey array of verbal affixes give choices for expressing varying degrees or aspects of agency. Some may contribute to the agency implicature, others may block it. Activity predicates may have only an Actor, or both an Actor and one or more Undergoer arguments.

2.3. Achievement and Accomplishment predicates

While Kankanaey State predicates denote a non-dynamic ongoing situation, changes of state are a different kind of predicate. As pointed out in VVLP (1997:93), Achievements and Accomplishments are not static. They are “happenings,” but they are not dynamic, in that they do not involve any action on the part of the participant. There is effectiveness, but no intention. These change-of-state predicates have Undergoer arguments.

3. Thesis: The role of *-om-* in Kankanaey predicates

This section looks at representative examples of verbs formed with *-om-* and notes that in each case, the agency of the cross-referenced argument is modified from the default value in some way. These anomalies vary according to the type of roots

and type of verbs that are formed; they do not at first seem to form a homogenous group. A careful look at the logical structures of the examples and a consideration of the common thread of atypical features of agency in each of them sheds light on the role of the affix *-om-*.

The following discussion looks first at verbs formed with *-om-* from action roots, then at verbs formed from stative roots. The examples include the Logical Structure of the verb. For ease in understanding the examples, please note that the cross-referenced argument follows the verbal predicate and is an absolutive-case pronoun, or a noun preceded by the absolutive nominal marker *din*.²

3.1. Action roots with *-om-*

With action roots, the affix *-om-* forms intransitive Activity predicates (see Table 1 above). The cross-referenced argument of these Activity predicates is the first argument of the **do'** predicate, which is an EFFECTOR with no inherent agency implied (VVLP, 1997:118). The following examples reveal the various agency modifications covered by *-om-*.

3.1.1. Unintentional Activity predicates

With roots that denote an atelic action (that is, one with no end-point) or movement, *-om-* forms Activity verbs with inherent or unspecified unintentionality. In example (1), the mass-noun argument (blood) is not intentional, and is moved by uncontrolled natural forces.

- (1) Omaloyas din dada na.
 Om- aloyas din dada na
 -om- flow NM blood 3sE

'His blood flows down.
do' (blood, [**flow'** (blood)])

In (2), the baby is crying uncontrollably or unintentionally. When an older person cries, the root is more likely to be affixed with *man-* (Actor-referencing), or with *i-* or *-an*, which cross-reference other entities related to the crying. The use of *-om-* blocks any agentivity implicature for this verb.

- (2) Omogaoga din moyang.
 Om- oga- oga din moyang
 -om- INTENS- cry NM baby

'The baby is bawling.'
do' (baby, [**cry'** (baby)])

² *din* (NM) is further analyzed as *di-n* (NM-def)

3.1.2. Dual-role Activity predicates

In example (3), *-om-* is used with a transitive root. (The Actor is chosen for cross-reference in this antipassive voice construction because of pragmatic, topical implications in the discourse.) *Onod*, ‘to follow someone/thing’ is an atelic movement verb with implied intentionality, and the MOVER is also a THEME that is the entity moved by the action. This dual role complicates or reduces the agentivity of the participant, and *-om-* is the Actor-referencing affix chosen.

- (3) Omonodak en agik.
 Om- onod -ak en agik
-om- follow -1Sa OPNM cousin.my

‘I follow/am following my cousin.’
do' (I, [follow' (I, cousin)])

3.1.3. Inchoative Activity predicates

The root *tayaw* ‘to fly’ most often occurs with *man-*, the unmarked Actor-referencing affix, to indicate the atelic action of birds overhead (*mantayaw*). This root lexicalizes some degree of intention or control (birds cannot accidentally fly!). When *-om-* is used instead of *man-*, as in (4), it cannot override the agentivity but rather specifies punctuality, i.e., the moment of inception of the the bird’s activity. This example does not immediately support the “reduced-agency” hypothesis for *-om-*-referenced Actors, and we will return to it shortly.

- (4) “Wit dokit” kanana yan pag tomayaw.
 Wit dokit kana -na yan pag -om- tayaw
 wit dokit say -3sE and then -om- fly (3sA)

“Wit-dokit” he said and then flew away.’
 INGR **do'** (3s, [fly' (3s)])

3.1.4. Antipassive Activity predicates with partial effectiveness

With 2-argument roots, the norm in Kankanaey is to use an Undergoer-referencing affix such as *-en* to form an Active Accomplishment verb as in (5) with the root *gisgis*, ‘to split something’. When the assertion answers the question “What is he doing?”, Actor-referencing *man-* is the affix used to form an Activity verb as in example (6). Note that the PATIENT, bamboo, is not referential. When the PATIENT is referential but only partially affected, as in example (7), the Undergoer argument can not be cross-referenced on the verb as a full-fledged argument of INGR **split'** because not all of it is affected. The less-than-effective Actor is cross-referenced by *-om-* in an antipassive construction, while the undergoer, still referential, is marked as oblique and definite.

- (5) Gisgisem din anes ay doy.
Gisgis -em *din* *anes* *ay* *doy*
 split -en.2sE NM bamboo LK that

'Split that bamboo.'

do' (you, [**split'** (you, bamboo)]) &INGR **split'** (bamboo)

- (6) Mangisgisgis si anes.
 Man- gis- gisgis si anes
 man- PROG- split (3sA) ONM bamboo

'He is splitting bamboo.'

do'(3s, [**split'** (3s, bamboo)])

- (7) Gomisgis ka sin anes ay doy.
 -om- gisgis ka si -n anes ay doy
 -om- split 2sA ONM -def bamboo LK that

'Split some of that bamboo.'

do' (you, [**split'** (you, bamboo)])

3.1.5. Antipassive Activity predicates with agentivity hierarchy inversion

As noted above, 2-argument roots typically take an Undergoer-referencing affix in active voice. This is the case in (8), where the Undergoer (GOAL here) of *ayag* 'to call someone' is cross-referenced by the affix *-an*.

- (8) Ay ayagam sisya?
 Ay ayag -am sisya
 Q call -an.2sE 3sA

'Are you calling him/her?'

do' (you, [**call'** (you, him/her)])

In (9), *-om-* is used with this root to cross-reference the Actor while the Undergoer is surprisingly implicit but not syntactically expressed. This is the choice of affixation with many verbs when the Undergoer is first person, or when the Undergoer is human and the Actor an animal, as in (10), or even an inanimate entity, as will be seen later in (11).

- (9) Ay omayag ka?
 Ay -om- ayag ka
 Q -om- call 2sA

'Are you calling me?'

do' (you, [**call'** (you, (me))])

- (10) Komat din aso!
 -om- kat din aso
 -om- bite NM dog
- ‘(Careful!) The dog bites!’
 do' (dog, [bite' (dog, (people))])

Silverstein (1976) has proposed the ‘Inherent Lexical Content Hierarchy’ (ILCH), which has the following order: 1st person > 2nd person > 3rd person > Proper Name > Human > Animate > Inanimate. When *-om-* is used to form antipassives with no expressed oblique second argument (as in the above two examples), it is only for situations in which the implied Undergoer is higher on the ILCH than the Actor. In other words, the implied Undergoer has more inherent agentivity than the Actor. This provides additional evidence that Actors of *-om-* predicates have reduced agentivity.

3.2. Stative Roots with *-om-*

With stative roots, the affix *-om-* forms several different types of predicates. The cross-referenced arguments of these predicates are atypical in their status as Undergoers. Causative Accomplishments with *-om-* cross-reference unintentional causers of situations. Position changes with *-om-* could be interpreted as Causative Accomplishments in which the Undergoer causes the change, or as Active Accomplishments in which the Undergoer is the Actor. Achievements with their inchoative INGR modification, signalling the sudden beginning of a state, come very close to being dynamic events involving only an Undergoer. Accomplishments and Processes cross-reference Undergoers who are independently participating in non-static situations.

3.2.1. Causative Accomplishment Predicates

Typical causatives in Kankanaey are derived with the prefix *pa-* in combination with other affixes. However, some stative and nominal roots form causative verbs with *-om-*. In (11) the inherent nature of wine is seen as causing drunkenness to unspecified arguments. In (12) something about the night or walking abroad at that time is seen as causing the presence of ghosts. These examples are interesting, because although the root is stative and intransitive, the causative introduces a second participant into the logical structure, the EFFECTOR that causes the change of state. This is the participant that *-om-* cross-references. In (11) the CAUSER is inanimate and the PATIENT (the one getting drunk) must be animate. In (12) the CAUSER is a state of affairs, and the THEME (ghost) is perceived as animate. In both examples, the CAUSER is less animate than the affected entity. Thus *-om-* is used to signal an inversion of the agentivity hierarchy with causative predicates.

- (11) Bometeng din alak.
 -om- beteng din alak
 -om- drunk NM wine
- ‘Wine is intoxicating.’
 [do' (wine, Ø)] CAUSE [BECOME drunk' (Ø)]

- (12) Adi ka mandan sin labi, tan bomanig.
 Adi ka man- dan sin labi, tan -om- banig
 neg 2sA man- walk ONM night because -om- ghost

‘Don’t walk at night, because it will make ghosts come out.’
 [.....] CAUSE [BECOME **be.present**] (ghost)]

3.2.2. Causative Accomplishment or Active Accomplishment Predicates

In (13), the state root *tokdo* ‘seated’ when affixed with *-om-* means ‘to sit down/up.’ This could be interpreted as a Causative Accomplishment, in which one does something to assume a seated position. Alternatively, this predicate could be seen as an Active Accomplishment, as represented in the example. The single argument is both MOVER (Actor) and the THEME (Undergoer) of the resultant state. In example (3) above with *onod*, ‘to follow someone or thing’ there was no question of assigning Undergoer status to the MOVER/THEME, due to the presence of a second argument of the root. With an intransitive stative root like *tokdo*, there is a tension between assigning Undergoer or Actor to the single argument. In these cases, the Actor-referencing option is handled with *-om-*. (The Undergoer-referencing option would use a different affix, *i-*.)

- (13) Tomoktokdoak.
 -om- tok -tokdo -ak
 -om- PROG -sit -1sA

‘I am (in the act of) sitting down.’
do’ (I, [sit’ (I)]) & BECOME **seated**’ (I)

3.2.3. Achievement and Accomplishment Predicates

Roots that express a static condition or a resultant state after some action typically form verbs with *ma-* (completive aspect *na-*). The actual moment (INGR) or process (BECOME or PROC) of changing a state, however, is not static but active. When these modifications appear in the logical structure, *-om-* is used with many roots.

In (14) and (15), the State and Achievement (instantaneous change-of-state) uses of *b(e)tak* ‘to burst’ are compared. Example (14) shows the static situation of a flat tire. When *-om-* is used, in (15), the predicate describes an active event as the participant independently begins to be in the state denoted by the lexical root.

- (14) Nabtak din goma na.
 Na- btak din goma na
 na- burst NM innertube 3sE

‘Its innertube is burst/flat.’
burst’ (innertube)

- (15) Bomtak din goma na.
 -om- btak din goma na
 -om- burst NM innertube 3sE

‘Its innertube will burst/pop.’
 INGR **burst**’ (innertube)

(16) is an example of an Accomplishment verb formed with *-om-*. The progressive aspect supports the ongoing non-static interpretation. The THEME (bus) that moves toward the “arrival” state is participating in a active event. A more subtle example is with Process verbs, open-ended (atelic) changes of state. (17) shows a Process verb formed by *-om-* with a color-state root, and a participant with no agentivity yet participating in an event that is changing over time.

- (16) Domatdateng din bas.
 -om- dat dateng Din bas
 -om- PROG- arrive NM bus

‘The bus is approaching.’
 BECOME **be.at.ref.point**’ (bus)

- (17) Ngometit din lokto mo ibilag mo.
 -om- *ngetit* *din* *lokto* *mo* *i-* *bilag* *mo*
 -om- black NM yam if i(U-ref)- be.in.sun 2sE

‘The yam will darken if you put it out in the sun.’
 PROC **black**’ (yam)

3.2.4. Return to inchoative Activity predicates

Example (4) may now be compared with other examples in which *-om-* is used when BECOME or INGR is found in the logical structure. Although the argument maintains its inherent agentivity with *tayaw* ‘to fly’, the marked construction with *-om-* indicates the presence of INGR in the Logical Structure, denoting the beginning of the activity. One could argue that more intentionality might be implied by the inceptive aspect, another example of a modification of agentivity, and further study with more inchoative activity verbs (few examples have been noted) would be of interest. On the other hand, perhaps the use of *-om-* with INGR was extended from states to activities at some point in the history of the development of this language.

4. Conclusion

This study has examined various lexical and grammatical contexts in which the agency of a cross-referenced argument in Kankanaey diverges from the norm or default value when the predicate is formed with *-om-*. The examples include situations in which the Actor argument has no intentionality, or when it does not have full effectiveness. Actors that are both EFFECTOR and THEME of a predicate and Actors that

have lower inherent agentivity than the Undergoer also are cross-referenced by *-om-*. In situations where states are changed, the static meaning of the root is changed to an active event with no other participant than an Undergoer. This study has shown that in Kankanaey, these various complications and modifications of meaning are handled by the verbal affix *-om-*, which legitimizes the anomalous cross-referenced arguments of the predicates it forms.

Abbreviations

A	Absolutive (cross-referenced) case
def	definite
E	Ergative case
PROG	Progressive (reduplication)
INGR	Ingressive (punctual)
INTENS	Intensive aspect (reduplication)
LK	Linker
LS	Logical Structure
NM	Nominal Marker of cross-referenced NP
O(P)NM	Oblique (Personal) Nominal Marker
PROC	Process
Q	Question
s	singular
1,2,3	personal pronouns

References

- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Dixon, R.M.W., (ed.) *Grammatical categories in Australian languages*. Canberra: Australian Institute of Aboriginal Studies.
- Van Valin, Robert D., Jr., and Randy J. LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.

A Study of Participant Reference in Central Bontok*

Byung-Hoon Lee
SIL International

This paper is an analysis of narrative discourse in Central Bontok. In Central Bontok narrative discourse, new participants may be introduced in background information, in new information in the discourse body, in a participant's own speech, or in another participant's speech. New participants are usually referred to by proper names or NPs. However, if there is little ambiguity, new participants can be referred to by pronouns. Major and minor participants are treated differently in the discourse plot structure. Usually, major participants have a speaking role, while minor participants do not, as has been noted by Walrod (1979). Overspecifications of participant reference (full names or proper names plus description) have special functions. In an account containing a protagonist and an antagonist, the relative order in which these are referred to is significant. Besides these characteristics, the present study demonstrates the use of pronouns and NPs when the participants are reintroduced or have been mentioned in a series of sentences in the immediately preceding discourse.

1. Introduction

Central Bontok is a member of the Central Cordilleran subgroup of the Northern Philippine languages. It is most closely related to Northern Kankanaey, Kankanaey, Balangao, Ifugao, Kalinga, and Isinai (McFarland 1980:62). The language is spoken in the municipality of Bontoc and several nearby Bontoc villages in Mountain Province. The Central Bontok language group consists of approximately 35,000 speakers.

The present study describes how and where major and minor participants are introduced and reintroduced in Central Bontok narrative texts.

2. Data (Corpus)

Nine narratives were used to study Bontok narrative discourse. Three personal narratives were collected — one oral (P3) and two written (P1, P2). In addition, one written traditional narrative (F1) was collected, three traditional narratives (F2, F3, F4) were selected from published books for children, and two (F5, F6) were taken

* This paper is a development of *A Discourse Analysis on Central Bontok Narrative* written by a group of Bontok Bible translators at a discourse workshop led by Mike Walrod in January 2004. I am deeply indebted to Mike Walrod for his help and suggestions. I am grateful to Steve Quakenbush and Doug Trick for helpful comments and to Allan Johnson, Tammy Ruch, and Ginny Larson for editing help.

from an unpublished dissertation. See Tables 1 and 2 for a summary of the narratives. English translations of four narratives have been included as appendices.

Table 1. Personal narratives¹

Text	Title		Author
P1(31)	Nan inippengko id Mainit	'My experience at Mainit'	Mr. Eduardo Yango
P2(82)	Nan nensolowak id Betwagan	'My teaching at Betwagan'	Mrs. Josefa Maskay
P3(61)	Nan aliglowan nan iGolo ya iYangnen	'Peace pact between the Golo and Yangnen tribes'	Mr. Nomi Suo

(): number of sentences in the text

Table 2. Traditional Narratives

Text	Title		Author
F1(18)	Nan kosa ya nan otot	'The cat and the mouse'	Mrs. Josefa Maskay
F2(88)	Nan og-okhod cha Ap-apatto ken Changchang-tayan	'Ap-apatto and Changchangtayan'	Mr. Anasor Wayyas
F3(54)	Nan am-ama ay nangasawa isnan talaw	'The man who married the star'	Mr. Anasor Wayyas
F4(15)	Nan chamon si tilin	'The beginning of the rice birds'	Mr. Apo Anchemang
F5(6)	No apay nga ad-i maligo nan Lanao	'Why Lanao is never flooded'	Mrs. Chopochoopen Fakayan
F6(12)	Nan fafai ay iFontok ya nan falo	'The Bontok woman and the widower'	Mrs. Angelita Fagyan

(): number of sentences in the text

¹ To protect the privacy of people when necessary, pseudonyms are used in personal narrative texts.

3. Findings and discussion

3.1. Participants in the plot structure

3.1.1. Places where participants are first introduced in the plot structure

Table 3 below shows where participants are first mentioned in each narrative. (Note that an unspecified group of people or children is considered as one participant.) There are six stages of plot structure in Bontok narratives; setting, inciting incident, developing tension, climax, denouement, and conclusion. It can be seen that denouement and climax are optional.

Table 3. First Mention of Participants in Each Narrative

Text	Setting	Inciting incident	Developing tension	Climax	Denouement	Conclusion	Total
P1	3	5	8	2	2	0	20
P2	3	3	8	2	3	1	20
P3	7	2	6	4	2	1	22
F1	2	1	1	0	—	—	4
F2	3	0	0	0	---	0	3
F3	3	2	1	0	0	1	7
F4	2	1	0	0	—	—	3
F5	3	1	—	0	—	2	6
F6	2	1	1	0	—	0	4
Total	28	16	25	8	7	5	89

As can be seen from the table above, more participants are introduced in the setting than in the other parts; fewer participants are introduced in the conclusion, denouement, and climax. This pattern is more distinctive in traditional narratives than in personal narratives. In comparing the stages of developing tension with the inciting incident, far more participants are introduced in the former.

3.1.2. Major and minor participants

Major participants are whom the story is about and occur throughout the story, or at least throughout a large part of the story. Minor participants, on the other hand, are not essential to the plot and appear momentarily (Pebley 1997:22). For the study of the contrast between major and minor participants, Text P3 “Peace pact between the Golo and Yangnen tribes” was chosen, because it is thought by several Bontok persons to be the most well developed of all the narratives. Table 4 below contains a list of all the participants in text P3 along with the following information; which stages in the plot

structure they appear in, the form of their initial reference, how they are introduced, whether they had a speech role, and whether or not they are considered a major character. Reading the texts with Bontoc persons to seek their impressions on major and minor participants, occurrences, major speech roles, and place of participants' role was very instructive. Below the table is a discussion of the data.

Table 4. How a Participant Is First Introduced In Text P3

Location	Participant	Form of initial reference	How introduced initially	Speech role	Major
sdCDC (13)	I (Nomi Suo)	pronoun/proper noun	self introduction	+	+
sdC (18)	Yangnen people	NP	background information	-	-
sdC (17)	Golo people	NP	background information	+	-
sidDC (21)	We (Exclusive)	pronoun	background information	+	-
sd (4)	Iroda	proper noun	background information	+	+
sd (2)	Kano's younger sister	NP	background information	-	-
sdC (22)	Kano	proper noun	background information	+	+
idC (7)	Mayang	proper noun	with his direct speech	+	+
idC (12)	we (Inclusive)	Pronoun	in someone's speech	-	-
id (4)	Khayaman	proper noun	with his direct speech	+	+
idC (2)	Sechida	proper noun	new information	-	-
dc (4)	Gihong	proper noun	new information	-	-
dc (6)	officials	NP	new information	-	-
dc (5)	Lida	proper noun	with his direct speech	+	+
dcD (27)	Chong	Proper noun	in someone's speech	+	+

d (1)	Chong's children	NP	in someone's speech	-	-
c (6)	they	pronoun	with their direct speech	+	-
c (1)	PC	NP	new information	-	-
c (1)	Khoiron	proper noun	new information	-	-
c (3)	soldiers	NP	new information	-	-
c (8)	Golo woman	NP	new information	-	+
DC (2)	Tongan	proper noun	with his direct speech	+	-
D (2)	Golo old men	NP	with their direct speech	+	-
C (2)	Chakan people	NP	new information	-	-

s: setting, i: inciting incident, d: developing tension, c: climax, D: denouement C: conclusion.
(): total occurrences in the narrative.

The author, Nomi Suo, as a person seen at each stage except during the inciting incident, is more than just one of the major participants. He actually leads the whole story, which may be described as his official adventure. The two opponents, Kano, a brother of the murdered woman, and Chong, the killer, are found most frequently throughout the story and are major participants. However, though first person plural pronouns (exclusive/inclusive) are found throughout the story, they are not prominent. In several places in the text, the identity of the antecedents of these first person pronouns is not clear. In terms of distinguishing major and minor participants, the residents of Golo and Yangnen are controversial — several Bontoc readers identify them as major participants, but it is questionable because the people of Golo and Yangnen in the text move as a crowd.

In the “Peace pact” text, many participants are introduced by means of proper names. Many of them are high-ranking government officials. Some of them make direct speeches in the text, others do not. According to Walrod (1979:48), in Ga'dang,² minor participants are those who do not have a speaking role. This appears to be the case in this “Peace pact” text. Though it is not always straightforward to determine whether a particular participant is a major or a minor one, a person who speaks a lot (major speech) is much more prominent than one who speaks very little or not at all. Governor Iroda and Governor Khayaman who have speaking roles are more prominent than Bishop Sechida or Captin Gihong. On the other hand, even if someone

² Both Ga'dang and Central Bontok belong to the Cordilleran language group.

has a speech role, if it is nominal, s/he is not so prominent and thus is classified as a minor participant. Tongan is an example of such a participant.

New participants are introduced in four different ways. Some are introduced by background information in the setting or by new information given by the narrator as the story proceeds. Others are introduced in another participant's speech, or with his/her own speech. A participant who is introduced with his own speech tends to be more prominent than those introduced by other means. Most high officials who are introduced with their own speech are prominent. A new participant may be referred to by means of a proper noun, a full NP, or a pronoun. Roughly speaking, a participant who is introduced by means of a proper noun is more prominent than one whose first reference is by means of a full NP or pronoun. (However, the use of first person singular pronoun is an exception; such a referent tends to have high prominence.)

A participant who plays an important role around the time of the climax is more prominent than those whose role is primarily at other stages in the discourse. A Golo woman referred to by means of a full NP, and who did not speak, and who was seen for only a short time, happens to play a key role in the climax as a major participant. An old Golo man has a major speech role, but it occurs only at the end of the denouement; he is locally prominent and classified as a minor participant.

3.1.3. *The order of participants reintroduced*

The word order of some collocational words like 'butter and cheese' or 'the cat and the mouse' is not important, but some other collocational words have a typical order. Regt (1998) points out that the order of participants in one constituent can deviate from what is usual. Some biblical examples include 'mother and father' in Leviticus 19:3 and 'concubines and wives' in 2 Samuel 5:13. The relative order of such participants may mean something significant.

In order to investigate the relative order of participants, text P3 was used, which mentions the Golo tribe and the Yangnen tribe several times. Table 5 below shows the results.

Table 5. Contrasting Internal Order between the Two Tribes in Text P3

Ref. no	Plot structure	Circumstance	Connotation	First	Second
2	Setting	Enmity	Negative	Yangnen	Golo
5	Inciting incident	Confrontation	Negative	Yangnen	Golo
6		Denial of crime	Negative	Yangnen	Golo
7		Negotiation	Positive	Golo	Yangnen
8	Developing tension	Geographic distance	Neutral	Golo	Yangnen
16/17		Position of sitting	Neutral	Golo	Yangnen
29		Celebration after making a peace pact	Positive	Yangnen	Golo

In narrating the Yangnen and Golo conflict, the author, who is from Bontoc, appears to be neutral at first glance. But on clear inspection of how the two parallel participants are introduced, it appears that he is more sympathetic to the Golo. When he refers to two people under negative circumstances, he mentions the Yangnen first and the Golo next. But in hopeful, positive, or neutral circumstances he mentions the Golo first and Yangnen second. The only exception is in ref. no. 29, where the Bontoc are mentioned preceding the two tribes.

3.2. The use of full noun phrase, pronoun and zero anaphor

This section will look at how the participants in the texts are encoded, whether using full NPs, pronouns or zero anaphor.

3.2.1. General rules/ default encoding

3.2.1.1. Full NP at first mention, and next pronoun and zero anaphor

When a participant is first introduced, a full NP (or proper name) is usually used. If it is mentioned in the same sentence (or immediate sentence), a pronoun is used. If it is referred to again in the subsequent clause, a zero anaphor may occur. Examples (1) and (2) illustrate this.

- (1) Enliwid **nan kosa ya nan otot** issan chey.
 IMPFT.AG-friend TRM cat and TRM mouse OBL D3

‘Before **the cat and the mouse** were friends.’

Nenfinachang **cha** issan amin ay chonocha.
 PFT.AG-REC-help 3PL OBL all LK work-GEN.3PL

‘**They** helped each other in all their work.’ (F4)

- (2) Isaa tako **siya** tay foweg tako.
 IMPFT.TH-take.home 1PL.IN 3SG because go.together 1PL.IN

‘We took **him** home because we were going together.’

Saet kanan Mayor Tongan en, “Siya man, khawis is
 then say Mayor Tongan COMP yes please good OBL

isaa **tako** \emptyset et.”
 IMPFT.TH-take.home 1PL.IN \emptyset ??

‘Mayor Tongan said, “Yes, it is better if **we take (him)** home.”’(F4)

Table 6 below shows a summary of all the participant references, and how they are encoded, in text F1. A discussion follows.

Table 6. Full NP, Pronoun, and Zero Anaphor Comparison in Text F1

Ref.no.	Full NP		Pronoun		Zero		Plot structure
	Cat	Mouse	Cat	Mouse	Cat	Mouse	
1	+	+					Setting
2			they, their				
3a	+	+	my, we(ex)	you			Inciting incident
3b			I, their				
4		+					
5	+	+	she	her			Developing tension
6	+	+	you(pl), your(pl)				
7	+		her				
8					+		
9			they				
10	+	+	her				
11			she				
12		+	she, her		+		Climax
13			she	them	+		
14			she				
15			she				
16	+		her	your			
17		+					
18	+			they, them			
total	16		25		3		

+: check mark

Text F1 “The cat and the mouse” shows that the most common form of reference is by means of a pronoun, the next common form of reference is a full NP, and the least common form of reference is a zero anaphor. Though a full NP occurs much more often than is the case in other texts (because of the contrast of two major participants all through the text), a pronoun is still preferred to a full NP when there is no ambiguity of identity. In this text, a zero anaphor is found only at developing tension and climax in the plot structure.

3.2.1.2. *The place of zero anaphors in the plot structure*

According to Walrod (1979:25) maximum deletion marks the climactic part of a story, distinguishing it from the setting and development parts by deleting nonverbal elements of clauses. “Routine participant reference may be disturbed” (Longacre 1983:25) in the peak. As such, in the climax less direct referent marking is used (that is, not full NPs). Zero anaphors occur in the buildup of tension in the plot where the action moves quickly and is encoded by shorter clauses with zero anaphors. Some nouns and pronouns are used to distinguish which participant did what to which other participant, but the obvious references are zero anaphors (Pebly 1997:44)

Table 7 below indicates that a zero anaphor is more often found in the climax and developing tension than in any other places in the plot structure. According to the table this phenomenon is distinctive in relatively long traditional narratives but not in personal narratives.

Table 7. The Place and Numbers of Occurrence of a Zero Anaphor in Each Text

Text	Set-ting	Inciting incident	Developing tension	Climax	Denoue-ment	Concl-usion	Total
P1(31)					1	1	2
P2(82)	1	1	1				4
P3(61)	1			1	1		3
F1(18)			1	2			3
F2(88)			1	6		1	7
F3(54)	1	2	4	7		3	17
F4(15)		1					1
F5(6)							0
F6(12)	1						1
Total	4	4	7	16	2	5	38

(): length of text

Note: Zero anaphors in relative clauses were not counted.

3.2.1.3. *A zero anaphor to describe a vague circumstance*

When a vague circumstance is described, a zero anaphor is used. In the following examples, no overt referent is found.

- (3) Masapor ay ikatpe nan likhat.
 must LK IMPFT.TH- endure TRM hardship

‘(You) must endure the hardship.’ (P2)

- (4) Chalanen nan siyam ay kilometro ay omey
 walk-IMPFT.TH TRM nine LK kilometer LK IMPFT.TH-go
- ischi mo marpo ka id Betwagan
 L3 if IMPFT-come.from 2SG OBL Betwagan
- ‘(You) have to walk nine kilometers to go there if you come from Betwagan.’(P2)

3.2.1.4. A zero anaphor in a relative clause

If the head of a relative clause is co-referential with the participant of the relative clause, a zero anaphor occurs in the relative clause. Example (5) illustrates this.

- (5) Wad-ay esang ay fafai ay \emptyset kaki-imen.
 EXT one LK woman LK \emptyset give.birth
- ‘There was a woman who \emptyset had just given birth.’ (F6)

3.2.2. Special rules or marked encoding

3.2.2.1. Highlighted reference

Although a participant may have been mentioned in a previous clause, if it is highlighted, a full NP is used in the subsequent clause instead of a pronoun even in the same paragraph.

In example (6) below, there is nothing in the speech content of the first sentence that might cause ambiguity of referent in the next sentence (and would thus require a noun phrase to disambiguate the referents). That is, the use of a full NP in the second sentence is conditioned by something other than disambiguation.

- (6) Saet kanan **Ot-otot** en, “.....”
 and ay mouse COMP
- ‘And **the mouse** said, “.....”’
- Ngem si **Ot-otot** kananan, “.....”
 but NM mouse say-3SG-COMP
- ‘But **the mouse** thought, “.....”’ (F2)

As can be seen in the example above, left dislocation occurs in Bontok often for highlighting a referent.

- (7) Kanan cha en, “Yake ka kak-imen ya cha ka
 say 3PL COMP then 2SG give.birth and CONT 2SG

kayet mamoknag?”
still IMPT.AG-work

‘They asked, “Why are you working when you have just given birth?”’

Ikhegkhenek angkhay nan nay **anochicha**
IMPFT.TH.CONT-be.quiet only TRM D1 sister-3PL.GEN

‘**Their sister** just kept quiet.’ (F6)

In the second clause of example (7), a third person singular pronoun would normally be adequate because the referent is the addressee of the previous clause. However, a full NP (‘their youngest’) is used in order to suggest a close relationship between the participants.

3.2.2.2. Use of full name

A full name may be used in order to express intensification. In text P3, *Chong* (Ref. 022, 023) was initially introduced by means of his first name (as were other participants), but as the story climaxes, he is referred to as *Chong Misu* (Ref. 024, 032).

3.2.2.3. Overspecification

Usually in developing tension or in the climax, as the story peaks, a referent may be overspecified. This serves to heighten tension in the narration. See the following example.

- (8) Ensafatcha pay ay enlamano ay
IMPFT.AG-meet-3PL till Lk IMPFT.AG&TH-shake.hands LK
- chey ya osto ay sisya ay egay chomapos nan
D3 and when LK still Lk NEG IMPFT.AG-reach TRM
- limacha ay ensafat, sanat **si** **Kano** ay
hand-GEN.3PL Lk IMPFT.AG-meet then.he ABS Kano Lk
- etad** **san** **fafai** **ay** **natey**, oksotna et isnan
brothers late woman Lk PFT.dead pull.out-3SG PAUS OBL
- fadfadyokna ya fadkong isnan poon nan
DIM-small.sword-GEN.3SG and stab OBL lower.part GEN
- fakhang Chong is ka likod.
neck Chong OBL DEF back

‘They met and prepared to shake hands, but before their hands reached each other, **Kano, the brother of the dead woman** suddenly appeared and took out his small sword and struck at the nape of Chong’s neck.’ (P3)

In the context, the long specification preceded by the proper name ‘Kano’ is not strictly necessary because the audience already knows that he is the brother of the dead woman (by previous mention), but the author repeats it as overspecification.

3.2.2.4. *First mentioned pronouns*

The plural pronouns *kami* (first person exclusive) and *cha* (third person) are often used to introduce an unspecified referent. In the personal narratives, *kami* (including the narrator) is found at the early part of the narration. Eventually, the identity of *kami* is gradually made known from subsequent discourse. See, as an example, the following translation from P1.

When I was with Keith Benn, my job was to show the video to the fifteen villages that surround Bontoc and the eight villages that surround the municipality of Sadanga.

Now the time came for me to go to Mainit village. My baggage consisted of a generator, a beta and a player, two gallons of gasoline, ten books, and my backpack full of clothes. We had gone six kilometers from Bontoc when it began to rain.

When we were between Bontoc and Mainit, the tires of the truck got stuck because the road had become muddy and slippery. The children were crying because the rain was very strong. What I did was this: I borrowed the bolo of one of those who was riding, and I went to get a very long young pine tree, and then we raised up the tire. We did that for one hour till we got it out of the mud. All the men joined in pushing so that the truck could be lifted from the mud onto solid ground. (Translation of P1)

In the above text, ‘we’ is not clearly known as first introduced, but eventually it is found to be passengers including the author himself.

3.2.3. *Anonymous referents*

Some participants in a text can be introduced without immediately receiving a name. Regt (1998) gives the examples of ‘a man’ (Gen. 32:25–31) or ‘youngest’ (1Sam. 16:11–13).

In text P2 “My teaching at Betwagan”, participants are mentioned in relation to the authors, and thus both major and minor participants are not referred to by proper names (except for one minor participant). It may be that proper names are not mentioned because the narration is a personal journal.³ This phenomenon is also seen in the text P1 “My experience at Mainit”. However, in text P3 “Peace pact”, a narration as a public and official journal, names of most major participants are clearly mentioned including the author’s own name.

In text F6 “The Bontoc woman and the widower”, major participants as well as minor ones are anonymous when the narration is about something derogated in Bontoc social value. Major participants are introduced vaguely and their names are not revealed until the end (e.g., *esang ay fafai* ‘one woman’, *chowa ay etadna ay lallaki* ‘two elder male siblings’, *esang ay am-ama ay falo* ‘one widower’). A similar phenomenon is observed in text F4 “The beginning of the rice birds”.

³ When the author of P2 was asked why most of the participants are mentioned anonymously, she said that it was for the respect of people older than her.

3.3. Speech margins

“The term *speech margin* is used to refer to the words that introduce actual speech” (Levinsohn 1992:128). In Central Bontok, a speech margin usually occurs before the speech. However, when the conversation is highlighted, it occurs after the speech.⁴ The margin may be split by the speech or omitted.

3.3.1. Split speech margin

The motivation for a split speech margin in Bontok is not yet clear. Perhaps it occurs when the former and the latter speeches are somewhat loosely connected to each other.

- (9) “Away inmey akhes nen-ani,” insongfat nan
 maybe PFT.TH.go also PFT.AG-harvest PFT.TH.answer TRM
- esang ken chaicha.
 one OBL 3PL

““Maybe she went to harvest also,” one of them answered.’

“Omey yo ilaen.”
 IMPFT.TH-go 2PL see-IMPFT.TH

““You go and see.”’ (P2)

3.3.2. Omitted speech margin

When speech is interchanged several times between two participants, the speech margin tends to be omitted; if the speech chain is interrupted by a certain event, the speech margin will reappear.

- (10) “Ay ad-adchawi ngen nan Betwagan?” kasinko sinalodsod.
 ?? INTNS-far Q TRM Betwagan again-1SG PFT.TH-ask

““Is Betwagan very far?” I asked again.’

“Wen pay. Kaatna nan omayam?”
 yes still when TRM IMPFT.TH-go-NMR-2SG

‘(Mother answered,) “Yes, it is. When will you go?”’

Masapor ay wad-ay ak ischi is Chomingkho ta
 must LK EXT 1SG L3 OBL Sunday so.that

⁴ However, habitual speech margin after the speech that neutralizes highlighting is observed in text P2.

en-isolo ak isnan Lunes.”
 IMPFT.AG-teach 1SG OBL Monday

‘(I replied,) “I must have to be there on Sunday so that I will teach on Monday.”’ (P2)

3.3.3. *Pronouns identified in wider context related with speech margin*

Pronouns may be identified by a subsequent utterance. So if the immediately following direct speech provides a clue for disambiguation between two different participants who have semantic competition in successive clauses, pronouns instead of NPs may be used.

(11) Fomoknag cha pen inamin
 IMPFT.TH-go.to.work 3PL ?? PFT.TH-consume

si tilin nan kowan san aloyosna.
 ABS rice.bird TRM possession GEN friend-3SG.GEN

‘When they both went to their fieldwork, the latter’s produce was all eaten up by the rice birds.’

Khinmakhawis san kowan san aloyosna ay esa.
 PFT.TH-good TRM possession TRM friend-3SG.GEN LK one

‘But her friend’s was in good shape.’

Saet kanana en, “Ya ngag man nan taken mo? Enman
 then say-3SG COMP and what ?? TRM different Q why

khawis nan anyem?”
 good TRM harvest-2SG.GEN

‘Now the **one** (who went dancing) asked, “What is the difference? How come **you** have good produce to harvest?”’ (F4)

In the above text, participants are referred to as ‘two close friends’ and then differentiated by ‘one’ and ‘the other’. In considering only the two previous clauses, the pronoun referent in the third clause cannot be clearly identified. However, in this case the two pronouns are mutually exclusive, and subsequent reference in direct speech clarifies their identity.

4. Summary

Bontok narratives have six stages (setting, inciting incident, developing tension, climax, denouement, and climax) of plot structure; of these six stages, denouement and climax are optional. More participants are introduced in the

setting than in the other parts; fewer participants are introduced in the conclusion, denouement, and climax. This pattern is more distinctive in traditional narratives than in personal narratives.

Though the distinction between a major and a minor participant is not always obvious, a person who speaks a lot (major speech) is much more prominent than another who does not. A participant who plays an important role close to the time of the plot climax is more prominent than those who occur only in other areas of the discourse. Criteria for distinguishing major from minor participants include their total number of occurrences, relative importance of their speech roles, and the places where they occur in the discourse. A new participant may be introduced in one of four different ways. Some are introduced by background information in the setting, others by new information given by the narrator as the story proceeds. Others are introduced by means of their own speech or in another participant's speech. A participant who is introduced by means of his own speech tends to be more prominent than those introduced by other means. In referring to two parallel participants, the protagonist is usually mentioned first in a positive or neutral circumstance, and later in a negative circumstance.

When a participant is first introduced, a full NP (or a proper name) is usually used. The narrator may choose to highlight the referent by means of a full NP in place of an expected pronoun. A referent's full proper name may occur (or recur) in order to express intensification of the narrator's emotion. Usually in developing tension or at the plot climax, a referent may be over-specified. This serves to heighten tension in the narration. Often proper names are not overtly mentioned when the narration is a personal journal and where the participants are mentioned in relation to the author(s). However, in a public or official discourse, the names of most major participants are clearly mentioned.

If the participant is referred to again in the same sentence (or an immediate next sentence), a pronoun is used. In the case of an immediately subsequent third or further reference, a pronoun or a zero anaphor may occur. Plural pronouns *kami* (first person exclusive) and *cha* (third person) are often used to introduce an unspecified referent. In the personal narratives, *kami* (often including the narrator) is found at the early part of the narration. Eventually, the identity of *kami* is gradually made known from subsequent discourse. Pronouns may be identified by a subsequent utterance. So if immediately following direct speech provides a clue for disambiguation between two different participants in semantic competition in successive clauses, pronouns instead of full NPs may be used.

In the discourse climax, less overt referent marking occurs. A zero anaphor is found more often in climax and developing tension than in other places in the plot structure. This phenomenon is more distinctive in relatively long traditional narratives. If the head of a relative clause is co-referential with the participant of the relative clause, a zero anaphor occurs in the relative clause.

Speech margins in Bontok usually occur before the speech. However, when the conversation is highlighted, the margin occurs after the speech. The margin may be split by the speech or omitted altogether. When speech is interchanged several times between two participants, the speech margin tends to be omitted; if the speech chain is interrupted by a certain event, the speech margin will reappear.

Abbreviations

ABS	absolutive	LK	linker
AG	agent	NMR	nominalizer
COMP	complementizer	NEG	negator
CONT	continuous	NM	nominal marker
D1	demonstrative – near speaker	OBL	oblique
D2	demonstrative – near hearer	PAUS	pause
D3	demonstrative – far from speaker and hearer	PFT	perfective
DEF	definite	PL	plural
EXT	existential	Q	question marker
GEN	genitive	REC	reciprocal
IMPFT	imperfective	SG	singular
IN	inclusive	TH	theme
INTNS	intensive	TRM	term; nuclear nominal, either S, A or P without specifying which one
L1	locative – near speaker	.	compound gloss
L 2	locative – near hearer	??	undetermined meaning or function
L 3	locative – far from speaker and hearer		

Appendix 1. My experience at Mainit

(Translation of Text P1)

There was one time I traveled to Mainit, and this is about my experience there. When I was with Keith Benn, my job was to show the video to the fifteen villages that surround Bontoc and the eight villages that surround the municipality of Sadanga.

Now the time came for me to go to Mainit village. My baggage consisted of a generator, a beta and a player, two gallons of gasoline, ten books, and my backpack full of clothes. We had gone six kilometers from Bontoc when it began to rain.

When we were between Bontoc and Mainit, the tires of the truck got stuck because the road had become muddy and slippery. The children were crying because the rain was very strong. What I did was this: I borrowed the bolo of one of those who was riding, and I went to get a very long young pine tree, and then we raised up the tire. We did that for one hour till we got it out of the mud. All the men joined in pushing so that the truck could be lifted from the mud onto solid ground. We gave thanks because the rain had stopped, and then I said to those who were riding, "Let's pray to thank God for His help because without Him, we could not do anything."

When we reached Mainit, it was nearly evening, and there was still a kilometer and a half to travel before we reached the village. Because it was evening, I could not do anything about showing the video, so after we had finished eating at the house of the Barangay Captain, I went to the men's house to hear the stories of the old men. We were talking about God and Lomawig, but the story that they wanted was about Lomawig because he was the first they knew who was called a god, and it is only recently that Jesus Christ is called the true God in the church. Then I said to them, "Tomorrow evening you will see the true God in the video." Then an old man with

white hair told a story about the teaching of Lomawig and the way to live here on this earth, and what to do so that the harvest would be good, and so the pigs and all the animals would be good and have good health.

The next day I went to visit officials and school teachers in the village. In the afternoon I showed the story about Jesus Christ in the school. The students were happy because they understood it and because Jesus Christ spoke in Bontoc.

When the show was finished, five books of Luke were sold. In the evening I showed the video at the Anglican Church. All the men, women, young people and children came to watch.

When the show started, I went outside to guard the generator. Then I saw ten monkeys (communists) looking over the window watching. I was very afraid, and so I prayed and I went to talk to them, and they told me not to be afraid. They commanded me not to tell any soldiers that there were monkeys who were always going to that village. If I told it, they said they would kill me; we shook hands and continued the show.

After the show, I slept alone at the church without fear. I praised God that night because those monkeys bought two books of Luke in the Bontoc language.

Appendix 2. My teaching at Betwagan

(Translation of Text P2)

“The Superintendent of the school will send me to Betwagan,” I said to my mother on the seventh of June, 1967.

“Really?” my mother responded.

“Is Betwagan very far?” I asked.

“Yes, it is. When will you go?”

“I have to be there on Sunday so that I can teach on Monday.”

My companions who went to Betwagan were my mother and my cousin. We rode in the Dangwa bus that goes to Tabuk on that Sunday morning. We got off the bus in Tabrak, a place that is ten kilometers away from Bontoc. We fixed our baggage, and then we walked down until we reached the Chico River.

There was no bridge. The water was high because a typhoon had just gone by.

“We must cross the river one by one. Let’s start upstream so that we swim downward,” my mother said. My mother was the first to swim. My cousin and I watched her until she reached the other side. Then I was next. Because the water was very deep, my cousin crossed the river with his hands holding the baggage up out of the water.

We sat down for a while on the riverbank. Then we started to go downstream. “Let’s walk slowly because the path is not good,” my mother said. We moved down until we reached the crossing of the Chico River and another river coming from Betwagan. “When will we get there?” I asked my mother again.

“Let’s be patient. We will get there eventually,” answered my mother, smiling.

We followed the Betwagan River. Even though we were tired, we enjoyed the travel, because the air was cool. The water was very clear. There were lots of fish swimming together. When we went to the other side, we walked on the rice paddies. It was slippery because it was muddy. The grain of the rice we were passing by was ripe. It was full and golden. Many people were harvesting.

I asked them if Mrs. Chapasen, the head teacher was there. “Maybe she went to harvest also,” one of them said. “You go and see. Hopefully she went home to cook food for her harvesting companions.”

We reached the school. My mother went to see her. She really had gone home to cook. She invited us to eat with her and then she led us to the house that I would rent. My mother and my cousin returned to Bontoc on that same day. I went to teach the next day.

Our head teacher said, “You teach the fourth grade.” When I entered the room, only five pupils were there. “Are these my only pupils?” I asked. “There are twenty five but many did not come because it is harvest time; this is also true during rice planting season.” I still taught them even though there were only five. Then they all came when the harvest was finished.

My pupils were young men and young women. They were good. I did not have difficulty because no one troubled me. I gave my best to these pupils of mine. I always went to bed late because I prepared first what I was going to teach. I woke up at dawn and went to school early. I loved them.

But I was very sad because I did not have much time to teach. My pupils had a lot of work to do both at home and in the fields. That’s why I sent them home on time. Another thing that made me sad was that many did not attend school. They stayed at home, or they went to the fields. To make it possible for more to come, I permitted them to bring the young ones they were taking care of to school. Some of them were even carrying babies on their backs to learn. Those who were able to walk could play outside.

It was hard to stay in Betwagan back then. You had to endure the hardship. We could not go home when the water in the river had risen. It was dangerous to cross. The road was difficult. We always stumbled on the stony pathway. And if there was a meeting or seminar, they went to Sadanga Poblacion. The poblacion is the center of all schools in the municipality of Sadanga. It is very far. You have to walk nine kilometers to go there if you start from Betwagan.

Even with these hardships, I still enjoyed the three years I stayed there. My pupils were all good. The villagers were also good.

Now, when I come across my pupils from Betwagan, I don’t recognize them, but they are the first ones to talk. I am so happy to see them again. They tell me they all got married. Some have grandchildren. They also say they deeply regret that they did not finish their studies. I answer them, saying it’s their fault.

Appendix 3. Peace pact between the Golo and Yangnen tribes

(Translation of Text P3)

I am Nomi Suo from Chakan, Bontoc. During the tribal war between the Yangnen people and Golo people, the negotiation of the two tribes was brought to the government. In the beginning of the negotiation, we went to Eboli. I escorted former Governor Juno Iroda.

We went to Eboli with the Golo representatives, and the Yangnen people were already there. We went to the Provincial Capitol of Eboli to settle the case, but it was not continued because the Yangnen and the Golo tribes were confronting each other. The Yangnen tribe said they had not killed but the Golo insisted that they had killed the younger sister of Kano.

Since that negotiation had begun, there was no resolution between the Golo and Yangnen. Then Governor Mayang of Fuyo said, "Suppose we transfer the place to settle to Fuyo, my province where both the Golo and the Yangnen are all far away. Then Governor Khayaman of Eboli agreed and also Governor Iroda of Mt. Province saying "Yes, even in Limon, Fuyo."

When the scheduled date came, we waited again for the Golo tribe and escorted them to Limon. There was Bishop Sechida and Captain Gihong. I was the escort of Governor Iroda. We led forty people of Golo from Sonuk. Governor Khayaman also led the Yangnen people from Yangnen to Limon. When we met at the PC headquarters in Limon, the officials of Limon were already gathered.

This is what they did to put us in order. On one side of the table were those of us from Mt. Province. The Golo tribe sat behind us. And also the Yangnen tribe was behind Governor Khayaman. The officials were at the center around the table. The other officials of Limon were on the other side.

When ex-governor Mayang opened the meeting, he said, "Now because we gathered together here, we will settle the case." Mayor Lida of Golo said, "What shall we settle when the one who killed is not known yet? The Yangnen tribe told lies, saying they are not the ones who killed the woman, the sister of Kano, who was guarding their rice field from birds."

So there were many discussions in using all the old cultures in settling the case. Then Mayor Lida said, "Suppose you, Chong, take this sugar cane wine and say a vow, saying you will die and also your children if you are the one who killed her. Make a vow to this sugar cane wine." When Mayor Lida said those words, Governor Mayang and even Governor Khayaman of Eboli wanted Chong to make a vow if he was not the one who killed her. Chong Misu of Yangnen was afraid; he raised his right hand and said, "Truly, I am." He admitted that he was the one who killed her.

The Golo tribe said, "Things are going well for us for settlement because the killer is known." Afterward Governor Mayang said, "So now that the negotiation of this case has been made, let us drink. You shake hands and then we will make the arrangement. We will all sign. This coming week, I will butcher a cow, so you Bontok, Yangnen and Golo people, come again and we will celebrate.

And then when they were preparing the sugar cane wine to arrange the shaking of hands of Chong and Lida, Lida said, "Yes." Governor Mayang said, "Alright you meet so that we all see you two shaking hands. Have your swords meet each other so that we make a "Peace pact" between the Golo and Yangnen people. When Mayor Lida stood, Chong Misu stood also. They met and prepared to shake hands. But before their hands reached each other's, Kano, the brother of the dead woman, suddenly appeared and took out his small sword and struck at the nape of Chong's neck.

That made the people shout in confusion. The PC shot Kano. I saw the Golo tribes trying to leave, but I told them, "No one should go; otherwise you will be shot." When the crowd became more chaotic, the officials ran away. The only people left were Colonel Khoiron, Bishop Sechida and Captain Marcelo Gihong and Governor Mayang, who were shouting and trying to stop the soldiers from firing, but they could not stop because the people were in chaos.

When they were shooting over Kano who was laying down unconscious, a Golo woman went and stepped in the middle and put up her skirt. With her lower body naked, she shouted in the middle. That made the soldiers stop firing because she was

there blocking the way with her upper legs open. Only when the fire stopped, the officials came out.

After that, I said to Marcelo Gihong, "Please Captain, we should bring Kano to the Hospital so that he will not die, because he keeps bleeding." Then we carried Kano. Those Yangnen who took their bolos jumped from the second floor to the first floor to confront us with their bolos but the PC surrounded them. They were very bloody. Finally the confusion stopped.

We took Kano to the Hospital. We did not have time to eat lunch because of the confusion. In the evening, I said to Governor Iroda, "We must take Kano to Sonuk. There might be another incident in the evening; they might pursue us again, so we would rather take him home because he is our companion." Mayor Tongan said, "Yes, it is better if we take him home."

We started there in the evening at 8:00. When we reached Mt. Polis, half way to Sonuk, I asked Kano, "Why did you stab Chong while the negotiation was going on in front of the officials and when he admitted that he had killed your sister?" Then Kano told me, "This is what you should remember. Old men of long ago said, 'If they are making a 'Peace Pact', be sure to take revenge before they have the swords meet. Because if the swords meet and we (people) drink, nobody can take revenge because your stomach will burst out if you take revenge after making a peace pact.' That's why I took revenge for my sister, because the man was there, saying he was the one who had killed my sister. The old men said, 'Before the swords meet and people drink, take revenge so it will be finished.'" That was Kano's answer to me on the way.

When we arrived in Bontoc in the early morning after a long slow trip, we took him straight to the Hospital. We came to the Municipal Hall and smoked there until morning. When we went home, Mayor Tongan informed the Lanon and they celebrated. The Chakan members asked what had happened and I told them the story. Then they said, "The *ato* is ready to celebrate, so you bring out the pounding wood." That was what happened. Then we Chakan members had a celebration.

Appendix 4. The cat and the mouse

(Translation of Text F1)

Before the cat and the mouse were friends. They helped each other in all their works.

One day the cat said to the mouse, "Please, take care of my children so that I can go to find food for them and for me. Here, I will leave this food for them to eat while I'm gone."

"All right! This is really what a friend does," the mouse replied.

When the cat left, the mouse fed to her children the food that the cat had left. "Do not tell your mother," the mouse said to the cat's children. When the cat came home, all of her children were crying. She asked what the reason was. They kept silent. Whenever the cat let the mouse take care of her children, the same thing happened. She wondered why it happened.

She let the mouse take care of her children again, and then she pretended to leave. She hid and looked at them secretly. She saw what was happening. She quickly went inside. "Really, this is what you have done," the cat said with fierce anger. The mouse was very afraid. Quickly she and her children ran away, but the cat immediately caught them.

References

- Bulaay, H., E. Cangat, C. Fatongkeg, B. Lee, M. Parting, and A. Manochon. 2004. *A discourse analysis on Central Bontok Narrative*. Unpublished manuscript.
- Ekid, A. 2002. *An ethnographic analysis of speech acts in Bontoc folktales*. Unpublished MA dissertation.
- Levinsohn, S. 1992. *Discourse features of New Testament of Greek*. Dallas: SIL.
- Longacre, R. 1983. *The grammar of discourse*. New York: Plenum Press.
- McFarland, C. 1980. *A linguistic atlas of the Philippines*. Study of Languages and Cultures of Asia & Africa Monograph Series No. 15. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Pebley, C. 1997. Participant reference in Kagayanen. In *Studies in Philippine Languages & Cultures* 10:2, 21–49.
- Regt, L. 1998. *Eumecot*. Participants in Old Testament Texts. Excerpt on Hebrew discourse workshop in UK, Dec. 2003.
- Wayyas, A. and A. Anchemang. 1986. *Nan og-okhod cha Ap-apatto ken Chagchangtayan*. Manila: SIL.
- Wayyas, A. and S. Pallangchao. 1986. *Nan Alitos nan Lafaan ya nan am-ama ay nangasawa isnan Talaw*. Manila: SIL.
- Walrod, M. 1979. *Discourse grammar in Ga'dang*. Dallas: SIL.

Deictic Pronouns In Philippine Languages

Curtis D. McFarland

Waseda University School of International Liberal Studies

A comparison of deictic pronouns in seven major Philippine languages (Tagalog, Sebuano, Ilokano, Hiligaynon, Bikol, Kapampangan, and Pangasinan) reveals a complex pattern of similarities and divergences. Modern Tagalog has three degrees of deixis (near me, near you, and far), and three case forms for each pronoun. Other languages have differing numbers of pronouns and case forms. Some languages have plural forms. In some languages, including Tagalog, the third (oblique) form of deictics also expresses location/direction. In other languages the locatives are distinct, but derived from, the deictics. In some but not all languages, there are derived forms expressing similarity (Tagalog ganito 'like this') and presence (Tagalog narito 'be here'). The translational approach also reveals that the boundaries between the deictics are not sharp, in contrast, for example, to the personal pronouns, where the boundary between 1st and 2nd person pronouns is extremely sharp. We also find an interplay between the deictics and 3rd person pronouns (e.g. using ito 'this' in place of siya 'he, she'). For historical linguistics the study points to a number of avenues for changes or shifts to take place. For theoretical linguistics it raised the question of whether any model using binary processing, or assuming a central processing unit can account for this degree of linguistic change.

1. Introduction

A few years ago I was asked to teach a course on comparative Philippine grammar at De la Salle University. At the time, a number of books of stories in several Philippine languages with Tagalog translations had recently been published. Naively I thought it should be very straightforward to ask the students to use these translations, starting with the most frequent words and grammatical structures in Tagalog, to find the corresponding words and constructions in the other languages.

The comparisons turned out to be extremely complex — and therefore interesting. I felt challenged to undertake a more extensive comparison of the Philippine languages, this time using a set of New Testament translations. In previous papers I compared case marking articles and discussed areas of stability and fluidity in these languages. The present paper presents a comparison of the deictic, or demonstrative, pronouns. Since deictic pronouns are among the very highest frequency items in these languages, and indeed in any language, it might be expected to be an area of relative stability. Again this is not the case — there is a high degree of divergence, not only in the forms which express the various deictic categories, but also with regard to the categories that are distinguished. In addition, questions are raised in terms of the pragmatic

question of which deictic is used, and in fact, whether a deictic will be used at all, in a given situation.

Translated materials are not the most desirable source of data, especially materials dealing with such un-Philippine and un-modern matters as the New Testament does. On the other hand, these translations provide a wealth of comparative evidence otherwise unavailable, and provide a number of insights not attained from monolingual analysis.

The languages included in this study, along with the abbreviations used in tables and examples, are Tagalog, Bikol (BIK), Sebuano (SEB), Hiligaynon (HIL), Kapampangan (KAP), Pangasinan (PANG) and Ilokano (ILO). Diachronically there is a major boundary between the first four languages, which are Central or Southern Philippine languages, and the last three, being Northern Philippine languages (McFarland 1980, Walton 1977). Also the first four languages are more closely related, thus more similar, to one another than the last three.

In this paper, first Tagalog deictics will be discussed, then other languages will briefly be introduced. Finally, a section about translating deictics will be included.

Note that the examples are given in a multi-line format. The first line is in the language under discussion. If this language is not Tagalog, it will be identified by its abbreviation, as indicated above. If the example is not from Tagalog, then the corresponding Tagalog passage is presented in the next line of the example. The passages in Tagalog and the other languages are presented as they appear in the texts, with no corrections for pronunciation, etc. The last line is the corresponding text from the English Good News translation; in most cases this is not a literal translation of the passage in the other languages.

The verse from which the example is taken is shown at the end of the example. The abbreviations are: Mt for Matthew, Mk for Mark, Lk for Luke, Jn for John, Ac for Acts, Col for Colossians, 2Tm for 2 Timothy, and Rv for Revelation.

2. Deictics in Tagalog

2.1. Deictic pronouns in Tagalog

Deictic pronouns are generally divided into three sets expressing nearness to the speaker (*ito*), nearness to the addressee (*iyán*), and remoteness from either (*iyon*). These pronouns also occur with high frequency within texts, where there is no concrete spatial reference. There is an additional form (*ire*), variously described as denoting actual physical contact with the speaker, or as denoting nearness to both speaker and addressee, as opposed to a more general meaning for *ito*. *Ire* does not occur with high frequency in modern written Tagalog. In McFarland (1989) I found 5 occurrences of *ire* as opposed to 5577 for *ito*. In the Tagalog New Testament translation it does not occur at all. It appears that to the extent that *ire* is used by some speakers it has more of a stylistic effect than a clear distinction in meaning.

Each pronoun has several forms to express different cases. The forms cited above (*ito*, *iyán*, *iyon*) can be called the nominative case, primarily expressing the grammatical subject of a sentence. (Note: the grammatical subject may not correspond to a subject in English or other languages.)

- (1) **Ito** ang minamahal kong Anak
This is my own dear Son
 Mt 3:17
- (2) Ibibigay ko sa iyo ang lahat ng **ito**
 All **this** I will give you
 Mt 4:9
- (3) Anong tao **ito**?
 What kind of man is **this**?
 Mt 8:27

A second set of forms (*nito*, *niyan* and *niyon* or *noon*) can be called genitive. Parallel to the genitive article *ng*, these forms have a number of functions, including ‘possessor’, ‘actor complement’, and ‘object complement’. For example:

- (4) Pumunta si Jesus sa bahay ni Pedro at doo’y nakita niya ang biyanan **nito**
 Jesus went to Peter’s home, and there he saw **Peter’s [this one’s]** mother-in-law
 Mt 8:14
- (5) sinabi **nito** sa mga babae
The angel [this one] spoke to the women.
 Mt 28:5
- (6) “Huwag hahawak **nito**,” “Huwag titikim **niyan**,” “Huwag hihipo **niyon**”?
 “Don’t handle **this**,” “Don’t taste **that**,” “Don’t touch **the other**”
 Col 2:21

Between *niyon* and *noon*, *noon* is used primarily with temporal reference, *niyon* with the ordinary meaning.

- (7) Tinaga niya ang alipin ng pinakapunong saserdote, at natigpas ang tainga **niyon**.
 and struck at the High Priest’s slave, cutting off **his [that one’s]** ear.
 Mt 26:51
- (8) at ito’y gumaling -- naging tulad **noong** isa.
 and it became well again, just like **the other [that]** one.
 Mt 12:13

But

- (9) Nguni’t natutulog **noon** si Jesus.
 But Jesus was asleep [**at that time**].
 Mt 8:24

- (10) Alalahanin ninyo ang sinabi niya sa inyo **noong** nasa Galilea pa siya:
Remember what he said to you **while** he was in Galilee:
Lk 24:6

(The genitive form *nire* (corresponding to *ire*) did not occur in either McFarland (1989) or the Tagalog translation.)

The oblique forms (*dito*, *diyan*, *doon*) are primarily used to express place or destination: ‘there’, ‘to there’, etc.

- (11) **Dito** nakatira ang kanyang mga kapatid na babae, hindi ba?
Aren’t his sisters living **here**?
Mk 6:3
- (12) At papatayin namin siya bago dumating **dito**.
But we will be ready to kill him before he ever gets **here**.
Ac 23:15

They also express ‘to this’, ‘of this’, etc. as in the following examples:

- (13) napakita ang Panginoon kay Saulo at nakipag-usap **dito**
Saul had seen the Lord on the road and that the Lord had spoken **to him**
[this one].
Ac 9:27
- (14) Ang paghahari ng Diyos ay maitutulad **dito**:
At that time the Kingdom of heaven will be like **this**.
Mt 25:14
- (15) Dahil **dito**’y nagtalu-talo ang mga Judio.
[because of] **This** started an angry argument among them.
Jn 6:52
- (16) nagugulo ang kanilang isip tungkol **dito**,
They stood there puzzled about **this**
Lk 24:4

The form *dine* (corresponding to *ire*) occurred 7 times in McFarland (1989), as opposed to 2350 for *dito*. It occurred 11 times (written *dini*) in the Tagalog Bible translation (*dito* occurred 242), usually in a context in which, for example, St. Paul was in one city, writing to people in another city, thus clearly designating a place near the speaker, not near the addressee.

- (17) Sikapin mong makarating **dini** sa lalong madaling panahon.
Do your best to come to me **[here]** soon.
2Tm 4:9

Tagalog deictic pronouns are summarized in the following table.

Table 1. Tagalog deictic pronouns

ito	iyán	iyón
nito	niyan	niyon/noon
dito	diyan	doon

In Tagalog, and in the other Philippine languages, the deictics can stand alone, as in the above examples, or as demonstrative adjectives modifying nouns. Such constructions usually include some kind of ‘linker’, and the deictic can come before or after the noun. When the deictic comes first, it carries the case marking for the phrase; if the deictic comes after the noun, it will be in the nominative form — the case of the phrase will be expressed by an article at the head of the phrase. See the following examples.

- (18) **Itong** bayang nag-apuhap sa gitna ng kadiliman
The people [**this** nation] who live in darkness
Mt 4:16
- (19) hindi na ako iinom **nitong** katas ng ubas hangga’t hindi dumarating ang kaharian ng Diyos.
I will not drink **this** wine until the Kingdom of God comes.
Lk 22:18
- (20) ako’y nananalangin **dito** sa aking bahay.
I was praying [**here**] in my house
Ac 10:30
- But
- (21) Ilang ang pook na **ito**
this [**this** place] is a lonely place.
Mt 14:15
- (22) Ano ang karapatan mong gumawa ng mga bagay na **ito**?
What right have you to do **these** things?
Mt 21:23
- (23) Maghari nawa ang kapayapaan sa bahay na **ito**!
Peace be with **this** house.
Lk 10:5

2.2. Tagalog deictic extensions

A number of Tagalog words or forms are derived from the deictic pronouns. The forms *narito*, *nariyan*, and *naroon*, consisting of *na-* plus the oblique pronouns, express ‘be here’ and ‘be there’. Related forms are *nasa* ‘be at’ and *nasaan* ‘be where?’

- (24) **Narito** na ang magkakanulo sa akin.
here is the man who is betraying me!
 Mt 26:46
- (25) Sapagka’t kung saan **naroon** ang inyong kayamanan ay **naroon** din naman ang inyong puso.
 For your heart will always **be [there]** where your riches **are [are there]**.
 Mt 6:21

Ganito ‘like this’, *ganyan* ‘like that’, and *ganoon* or *gayon* ‘like that’ are abbreviations of *gaya* ‘like’ plus the genitive deictic form. With other nominal expressions the full form is used: *gaya ng*, *gaya ni*, *gaya ko*, etc.

- (26) **Ganito** kayo mananalangin: ‘Ama naming nasa langit,
This [like this], then, is how you should pray: ‘Our Father in heaven:
 Mt 6:9

In Tagalog the primary verb of movement is *pumunta* ‘go’, not related to the deictic pronouns. However there are secondary derived forms, such as *pumarito* ‘come here’, *pumariyan* and *pumaroon* ‘go there’: *puma-* plus the oblique deictic forms.

- (27) susunod po ako sa inyo saanman kayo **pumaroon**.
 I am ready to go with you wherever you **go [go there]**.
 Mt 8:19

Compare:

- (28) siya’y natakot na **pumunta roon**.
 he was afraid to **go there**.
 Mt 2:22

The Tagalog deictic derivations or extensions are summarized in the following table.

Table 2. Tagalog deictic derivations

narito	nariyan	naroon
ganito	ganyan	ganoon/gayon
pumunta (pumarito)	(pumariyan)	(pumaroon)

3. Deictics in other Philippine languages

A thorough discussion of deictic pronouns in seven Philippine languages could be the topic of a monograph. In this paper I give only a brief summary of the total constellation, as an example or illustration of the degree of divergence to be found.

3.1. Nominative deictics

The nominative deictics in the other six languages are included, along with Tagalog, in the following table.

Table 3. Nominative Deictics in seven Philippine languages

TAG	ito	iyán	iyón
BIK	ini	iyán	idto
HIL	ini	ina	yadto
SEB	kini	kana	kadto
KAP	iti/ini/deti/deni	yan/den	ita/deta
PANG	saya/aya/saraya	satan/itan/saratan	saman/iman/saraman
ILO	daytoy/dagitoy	dayta/dagita	daydiay/dagidiay

Aside from the obvious variation in forms, I note the following differences in categorization:

Belchez (1992) also has *ito* for the third form in Bikol.

Hiligaynon and Sebuano have an additional form (HIL: *amo*, SEB: *mao*) which seems to be an all-purpose deictic which can stand alone to express deixis (emphasis) without spatial reference, or can collocate with the other deictics (*amo ini*, *mao kini*, etc.)

- (29) HIL: si Juan **amo** si Elias nga magaabot.
 SEB: si Juan **mao** si Elias kansang pag-anhi gitagna.
 siya ang Elias na darating.
 John is [**that**] Elijah whose coming was predicted.
 Mt 11:14
- (30) HIL: kay **amo ini** ang imo nahamut-an.
 SEB: kay **mao** may imong gusto ug pagbuot.
 sapagka't **gayon** ang ikinalulugod mo.
this was how you wanted it to happen.
 Mt 11:26

Sebuano has an additional form *kiri* (cited by Wolff 1966 and Zorc 1987), designating near the speaker but not the addressee, as opposed to the more general meaning of *kini*. This form did not occur in the Sebuano New Testament translation.

In Sebuano deictic pronouns do not follow nouns they modify; however there is a suffix *-a* which attaches to nouns (often together with a preceding deictic) to add emphasis.

- (31) SEB: Nakasabot ba kamo **niining** mga **butanga**?
Naunawaan na ba ninyo ang lahat ng **ito**?
Do you understand **these things**?
Mt 13:51
- (32) SEB: Tan-awa ninyo **kining tawhana**!
Masdan ninyo ang **taong ito**!
Look at **this man**!
Mt 11:19

Kapampangan presents a confusing picture. First, with regard to *iti* and *ini*, Mirikitani (1971:117–118) says:

“*Iti* refers to ‘this’ which is closer to the hearer and *ini* ‘this’ which is closer to the speaker. In reference to area, *-ti* appears to denote a general area, while *-ni*, a specific area.”

Davidson (1992) cites *iti* as ‘closest’ and *ini* as ‘near us’. Both *iti* and *ini* have high frequency in the Kapampangan translation, with no clear distinction of reference.

With regard to *yan* and *ita*, we get opposite descriptions. Mirikitani (1971:43) says:

“... *iyán* ‘that’ designates something close to the listener; and *ita* ‘that over there’ designates something away from both the speaker and the listener.”

Davidson (1992) has *ita* as ‘near you’ and *iyán* as ‘far’. Once again the distinction, if it exists, is not clear from the occurrences in the New Testament translation.

Kapampangan and the other Northern languages have plural deictic forms (in each case, the last form shown in table 3 above). In Tagalog and the other Central languages plurality is indicated with the particle *mga* (/manga/).

- (33) KAP: Linapit **la deti** kang Felipe a tau Betsaida, Galilea.
Lumapit **сила** kay Felipe na taga-Betsaida, Galilea,
They [these] went to Philip
Jn 12:21

In the Pangasinan translation, the deictic pronouns each had three forms in the singular. One form with *sa-* (*saya*, *satan*, *saman*); one with *i-* or *a-* (*aya*, *itan*, *iman*); and an abbreviated form with no prefix (*ya*, *tan*, *man*). Benton (1971) does not mention the *i-* forms. Again, in the translation there is no obvious difference between the *sa-* and *i-* forms. The plural forms had only *sara-* (*saraya*, *saratan*, *saraman*).

Ilokano has also the form *daydi*, apparently a shortening of *daydiay*.

3.2. Genitive deictics

The genitive forms are summarized in the following table.

Table 4. Genitive Deictics in seven Philippine languages

TAG	nito	niyan	niyon/noon
BIK	kaini	kaiyan	kaidto/kaito
HIL	sini	sina	sadto
SEB	niini	niana	niadto
KAP	niti	nian	nita
PANG	na saya	na satan	na saman
ILO	daytoy/dagitoy	dayta/dagita	daydiay/dagidiay

This is a fairly well behaved set. We note the following differences:

Sebuano *niiri* (corresponding to *kiri*) did not occur in the translation.

The form *nini* (corresponding to *ini*) did not occur in the Kapampangan translation; *nian* occurred only twice. Kapampangan does not have plural genitive forms. In Kapampangan case marking (nominative and genitive) is shown primarily by the (redundant) personal pronouns. The nominative forms *ya* 'he, she' and *la* 'they' and the genitive forms *na* and *da* generally appear in a sentence even when there is a corresponding nominal phrase. Thus, for example, *dareti*, which may look like a genitive form, is actually a composite of *da*, the third person plural personal pronoun, and *reti*, the nominative plural deictic pronoun.

- (34) KAP: binie na karing Israelita ing gabun **da reti**
 ipinagkaloob sa kanila ang lupain ng mga **iyon**
 made his people the owners of the [**their**] land [**of those**].
 Ac 13:19
- (35) KAP: Makatalakad ya libutad **da reti** ing metung a balamu lalaki
 Nakatayo sa gitna ng mga **ito** ang isang animo'y lalaki
 and among **them** [**these**] there was what looked like a human being,
 Rv 1:13

Pangasinan does not have distinctive genitive forms. The deictic pronouns are handled like nouns and take the genitive article *na* (corresponding to *ng* in Tagalog). This applies also to the plural forms.

- (36) PANG: Dinmagop iray amin ya ombaley ed arap **na saman** ya abung.
 at nagkatipon ang buong bayan sa may pintuan ng bahay.
 All the people of the town gathered in front of the [**that**] house.
 Mk 1:33

- (37) PANG: Iner kasiy angalaan to yan too **na saraya**?
 Saan niya nakuha ang lahat ng **inyan**?
 Where did he get all **this**?
 Mk 6:2

As with personal and common nouns, Ilokano does not distinguish nominative from genitive forms.

3.3. Oblique deictics

The oblique forms in the seven languages are shown in the following table.

Table 5. oblique Deictics in seven Philippine languages

TAG	dito	diyan	doon
BIK	digdi	diyan	duman
HIL	diri	dira	didto
SEB	diri/dinhi/ari/anhi	diha/anha	didto/adto
KAP	kaniti/kanini/kareti/ kareni/keti/keni	karin/karen/ken	kanita/kareta/keta
PANG	ed saya/dia	ed satan/ditan	ed saman/diman
ILO	iti daytoy/ kadagito/ ditoy	iti dayta/ kadagita/ dita	iti daydiay/ kadagidiay/ sadiay/idiay

This set is rather less regular formally than the genitive set; yet the connections are obvious. The following differences are noted:

In the Sebuano translation, both *diri* (nearest, or near speaker only) and *dinhi* (near us, or near speaker and addressee) occurred with high frequency. Sebuano also distinguishes the *a-* set from the *di-* set, in that the *a-* set has future or unreal reference. The *a-* set can also express movement in the given direction (*adto* '(will be/go) there').

- (38) SEB: Nagpabilin siya **didto** hangtod nga namatay si Herodes.
Doon sila tumira hanggang sa mamatay si Herodes.
 [**there**] where he stayed until Herod died.
 Mt 2:15
- (39) SEB: karong adlawa **adto** ka sa Paraiso uban kanako.
 ngayon di'y isasama kita sa Paraiso.
 today you will be [**there**] in Paradise with me.
 Lk 23:43

Pangasinan and Ilokano distinguish 'to this' from 'here', etc. In Pangasinan *ed* (corresponding to Tagalog *sa*) plus the nominative form expresses 'to this', etc. In

Ilokano *iti* (Tagalog *sa*) plus the nominative singular form expresses ‘to this’; in the plural there is a distinctive set with *ka-*. For ‘there’ the Ilokano translation had two forms *sadiay* (not mentioned in Lapid 1971) and *idiay*.

- (40) PANG: Siopa kasiy makatalus **ed saya**?
sino ang makatatanggap **nito**?
Who can listen to **it [this]**?
Jn 6:60
- (41) PANG: Inaro ak na Ama lapud **saya**,
Dahil **dito**’y minamahal ako ng Ama,
The Father loves me because ... [of **this**]
Jn 10:17
- (42) PANG: **Dia** ak nepeg a bistaen.
dito ako dapat litisin.
[**here**] where I should be tried.
Ac 25:10
- (43) PANG: Tinmaynan **diman** si Jesus
Umalis **doon** si Jesus
Jesus left **that place [there]**
Mt 15:21

Kapampangan also has a double set of forms (*kaniti/keti*, etc.), but the boundary between the two seems not so clear. The ‘to this’ set has plural forms (*kareti*, etc.).

3.4. Presence (*na-*) forms

The forms corresponding to the *na-* forms (‘be at ...’) in Tagalog are summarized in the following table.

Table 6. presence forms in seven Philippine languages

TAG	narito	nariyan	naroon
BIK	uya/anion	yaon	idtoon/itoon
HIL	yari	yara	yadto
SEB	ania	anaa	atua
KAP	atyu/atilu keti/keni	atyu/atilu ken	atyu/atilu keta
PANG	wala dia/wadya/nia	wala ditan/wadtan	wala diman/wadman
ILO	adda ditoy/adtoy	adda dita	adda sadiay/idiay

The Hiligaynon and Sebuano forms are clearly derived from the oblique forms (*diri*, etc.)

The Pangasinan and Ilokano forms contain an existential (PANG: *wala* = ILO: *adda* = Tagalog *may/mayroon*) plus the oblique forms. Kapampangan behaves similarly, but with some morphophonemic adjustment. Singular *atyu* consists of *atin* (Tagalog *mayroon*) plus *-yu* in place of the third person singular pronoun *ya*. *Atilu* is *atin* plus *-lu* in place of the third person plural pronoun *la*. We are reminded that the negative form in Tagalog (as in the other Central languages) is *wala* (the negative existential).

- (44) ILO: Ta no sadino ti yan ti gamengmo, **adda** met **idiay** ti pusom.
 KAP: Uling nu man karin atyu ing pibandian yu, **atyu** naman **karin** ing pusu yu.
 Sapagka't kung saan naroon ang inyong kayamanan ay **naroon** din naman
 ang inyong puso.
 For your heart will always **be** [**there**] where your riches are.
 Mt 6:21

Compare:

- (45) ILO: **Adda** rukib dagiti musang, ket **adda** umok dagiti tumatayab,
 KAP: **Atin** lang lunggang pisuluran ding asung-gubat at **atin** lang sale ding ayup,
May mga lungga ang mga asong-gubat, at **may** mga pugad ang mga ibon,
 Foxes **have** holes, and birds **have** nests,
 Mt 8:20
- (46) ILO: Iti saan unay nga adayo kadakuada, **adda** dakkel nga arban ti baboy
 nga agsubsubsob.
 KAP: King e masiadung makarayau karin **ating** dakal a babing manyibsiban.
 Sa di kalayuan doo'y **may** isang malaking kawan ng baboy na nanginginain.
 Not far away **there was** a large herd of pigs feeding.
 Mt 8:30

3.5. Similarity (*ga-*) forms

The forms corresponding to the *ga-* forms ('like ...') in Tagalog are summarized in the following table.

Table 7. similarity forms in seven Philippine languages

TAG	ganito	ganyan	ganoon/gayon
BIK	arog/siring kaini	arog/siring kaiyan	arog/siring kaidto
HIL	subong sini	subong sina	subong sadto
SEB	sama/ingon niini	sama/ingon niana	sama/ingon niadto
KAP	anti kaniti/kanini/ antimo/makanyan	anti karin/antimo/ makanyan	anti kanita/antimo/ makanyan
PANG	onya/singa saya/sirin	onya/singa satan/sirin	onya/singa saman/sirin
ILO	kastoy	kasta	kasta

In Bikol, Hiligaynon, and Sebuano the corresponding expressions are phrases consisting of a word meaning 'like' plus the genitive deictic form. Also BIK: *siring*, HIL: *subong*, and SEB: *ingon* can stand alone to express 'like that'.

- (47) BIK: Dangan nagsurat siya nin **siring kaini**:
At lumiham siya ng **ganito**:
Then the commander wrote a letter that went **like this**:
Ac 23:25
- (48) BIK: **siring** man an ginibo ni Isaac sa aki niyang si Jacob.
ganito rin ang ginawa ni Isaac kay Jacob,
Isaac circumcised [did **like that** to] his son Jacob,
Ac 7:8

Kapampangan has *anti* 'like' plus an oblique deictic; alternatively words such as *antimo* and *makanyan*, which stand alone to express 'like that'. In Pangasinan *onya* or *singa* 'like' plus a nominative deictic; alternatively *sirin* 'like that'.

Only Ilokano has corresponding derived forms *kastoy* and *kasta*; the third member of the set (**kasdiy*) seems to be missing. *Kas* also stands alone as a conjunction and can attract the enclitic pronouns.

- (49) ILO: Kalpasanna, nagsurat iti **kastoy**:
At lumiham siya ng **ganito**:
Then the commander wrote a letter that went **like this**:
Ac 23:25
- (50) ILO: makitak dagiti magmagna a tattao, ngem **kasda** la kayo.
Nakakakita po ako ng mga taong lumalakad, nguni't **silay parang**
punungkahoy.
I can see people, but **they** look **like** trees walking about.
Mk 8:24

3.6. Derived verbs of movement

The primary verbs of movement in the seven languages are summarized in the following table.

Table 8. derived verbs of movement in seven Philippine languages

TAG	pumunta		
BIK	magdigdi	magdiyan	magduman
HIL	magkari	magkadto	magkadto
SEB	moanhi	moanha	moadto
KAP	munta/mako		
PANG	onla		
ILO	mapan/in/umay		

Bikol, Hiligaynon, and Sebuano have verbs derived from the oblique deictics as the most general expression of movement. We note the following peculiarities.

We did not observe a form *magkara* in Hiligaynon. Phrases such as *magkadto dira* did occur in the translation.

- (51) HIL: ang mga luyag **magkadto dira** halin diri indi makatabok,
ang mga narini ay hindi **makapariyan**
those who want to **cross** over [**to there**] from here to you cannot do so
Lk 16:26

The Sebuano *a-* forms lose their future sense in verbs of motion. Thus we have real tense forms such as *miadto* ‘went’.

- (52) SEB: Unya **miadto** si Jesus sa balay ni Pedro
Pumunta si Jesus sa bahay ni Pedro
Jesus **went** [**there**] to Peter’s home
Mt 8:14
- (53) SEB: Si Jesus mibiya niadtong dapita ug **miadto** sa usa sa ilang mga sinagoga.
Umalis si Jesus at **pumunta** sa sinagoga.
Jesus left that place and **went** [**there**] to a synagogue,
Mt 12:9

The Northern languages do not have such derivations, at least not as the most general expression of motion.

3.7. Time expressions

As mentioned above *noon*, one of the genitive deictic forms in Tagalog, is used primarily to express past time, either as an adverb (‘at that time’) or a conjunction (‘when ...’). The corresponding terms in the other languages are shown in the following table.

Table 9. Time expressions in seven Philippine languages

TAG	noon
BIK	kaidto
HIL	sadto
SEB	niadto
KAP	kanita
PANG	nen/nen saman/ed saman
ILO	idi

Bikol, Hiligaynon and Sebuano follow the Tagalog pattern in using the ‘far’ genitive deictic to express past time. The corresponding Kapampangan form is the ‘far’ oblique deictic. Pangasinan uses *nen*, the genitive article for personal names (corresponding to Tagalog *ni*); the stand-alone adverb is a phrase, consisting of *nen* or *ed* (Tagalog *sa*) plus *saman*, the ‘far’ nominative deictic. Ilokano has a distinct form *idi*, not directly related to the deictics.

4. Alternatives to Deictics

In working with translations, we are always looking for one-to-one correspondences. I use the phrase ‘one-to-one’ in two ways. In one sense, for a given sentence, or a piece of text, we are looking for one word or phrase in one language that uniquely corresponds to a given word or phrase in another language. In another sense, we are looking for a similar correspondence which can be found in a large number, if not most, sentences, thus constituting a gloss or cross-language definition.

When we begin looking at translations, we are immediately made aware of the fact that such one-to-one correspondences, in one or both senses, are frequently unavailable. In the case of the deictics, we note the following types of divergence.

4.1. Different Pronoun

There are many cases where, for example, where Language A uses the ‘far’ deictic, but Language B uses the ‘near me’ or ‘near you’ deictic. In other cases the situation is reversed: Language A has ‘near’ corresponding to ‘far’ in Language B.

- (54) BIK: **Iyan** [that] gabos itatao ko saimo,
 Ibibigay ko sa iyo ang lahat ng **ito** [this]
 All **this** I will give you,
 Mt 4:9

4.2. Different Case

Different languages may use different focus configurations to express the same idea, with the result that the case of the deictics will be different. To some extent this

may be systematic. For example, Hiligaynon and Sebuano exhibit a greater preference for actor focus than Tagalog and some of the other languages.

- (55) HIL: Indi bala *kamo* [nominative] makahangup *sini* [genitive] nga palaanggiran?
Hindi pa ba *ninyo* [genitive] nauunawaan ang talinghagang *ito* [nominative]?
Don't *you* understand *this* parable?
Mk 4:13

4.3. Personal Pronoun

It is quite common in Philippine languages to use a deictic pronoun in place of a personal pronoun, for example in Tagalog *ito* 'this' instead of *siya* 'he, she.' Depending on the choice, perhaps mood, of the respective translators, we can have a deictic pronoun in one language corresponding to a personal pronoun in another language.

- (56) KAP: oneng tatakut *ya* [he] karing memalen Judio
nguni't natatakot *ito* [this] sa mga Judio
but *he* was afraid of the Jewish people
Mt 14:5

4.4. Noun Phrase

In a given situation, one translator may feel that a deictic pronoun is sufficient to express the reference; another translator may feel the need to be more explicit and use a noun phrase or name instead of the deictic.

- (57) ILO: imbilin ni Pilato a maited *ti bangkay* [the body] kenkuana.
Iniutos naman ni Pilato na ibigay *ito* [this] sa kanya,
Pilate gave orders for **the body** to be given to Joseph.
Mt 27:58
- (58) PANG: Makaliliket *si Herodes* nen anengneng toy Jesus,
Tuwang-tuwa *ito* [this one] nang makita si Jesus.
Herod was very pleased when he saw Jesus,
Lk 23:8

4.5. Absence of a Deictic

In some situations, the deictic may be considered to be completely redundant and can be eliminated. For example *sinabi niya ito*: 'he said this:' versus *sinabi niya* 'he said:'. Or *itong lalaki* 'this man' versus *ang lalaki* 'the man'.

- (59) SEB: Miluhod ang *sulugoon* atubangan sa iyang agalon,
Nanikluhod ang *taong ito* sa harapan ng kanyang panginoon
The [**this**] *servant* fell on his knees before the king.
Mt 18:26

In other situations the given proposition may be paraphrased, completely eliminating the deictic reference.

- (60) SEB: Nagtuo ba kamo nga makaayo ako kaninyo?
 Naniniwala ba kayo na magagawa ko ito? [Do you believe that I can do this?]
 Do you believe that I can heal you?
 Mt 9:28

5. Conclusion

These differences have been discussed as choices in the context of translation. But they strongly suggest that the same type of choices are available and/or required in the context of composition within a single language. It has been said that every text is a translation, that is a translation from a non-verbal situation or idea into a linguistic text.

It is further suggested that each of these various choices constitutes a channel for linguistic change, leading to the kind of diversity of form and categorization discussed in this paper.

The further question is: what kind of linguistic model could describe the process whereby, in a given situation, a given set of words, including deictics, and a given structure would be chosen, rather than alternative choices. I personally cannot accept any kind of computational model, or sequence of binary choices, which produces a unique solution in every situation. I would prefer a system based on mental associations, in which a given situation would stimulate various responses in the form of linguistic expressions. For example, both *ito* and *iyon* could be stimulated, but from the sum effects of linguistic and non-linguistic stimuli, the response for *ito* would be stronger in one case and for *iyon* in other cases. This is an idea which I hope to develop in the future.

Corpus:

English:

United Bible Societies. 1994. *Good News Bible*. Glasgow: Caledonian International Book Manufacturing Ltd.

Tagalog:

Philippine Bible Society. 1973. *Mabuting Balita Para sa Ating Panahon (New Testament)*. Manila: Philippine Bible Society.

Bikol:

Philippine Bible Society. 1992. *Marahay na Bareta sa Satuyang Panahon*. Manila: Philippine Bible Society.

Hiligaynon:

Manuel Degoma, Fr. Jose S. Aripio, SSP. 2002. *Biblia sang Katilingban sang mga Kristiano*. Quezon City: Claretian Publications.

Sebuano:

Philippine Bible Society. 1973. *Maayong Balita Alang Kanimo (New Testament)*. Manila: Philippine Bible Society.

Kapampangan:

Philippine Bible Society. 1994. *Ing Mayap a Balita Biblia*. Manila: Philippine Bible Society.

Pangasinan:

Philippine Bible Society. 1983. *Maung a Balita para sayan Panaon Tayo*. Manila: Philippine Bible Society.

Ilokano:

Philippine Bible Society. 1973. *Naimbag a Damag iti Agdama a Panawen (New Testament)*. Manila: Philippine Bible Society.

References

- Belchez, Chito A. and Pamela Johnstone Moguet. 1992. *Bikol Newspaper Reader*. Kensington, Md: Dunwoody Press.
- Benton, Richard A. 1971. *Spoken Pangasinan*. Honolulu: University of Hawaii Press.
- Davidson, Alma M. and Leonardo Aquino Pineda. 1992. *Kapampangan Reader*. Kensington, Md: Dunwoody Press
- Lapid, E. Bernabe V. and Bonifacio Sibayan. 1971. *Ilokano Lessons*. Honolulu: University of Hawaii Press.
- McFarland, Curtis D. 1980. *A Linguistic Atlas of the Philippines*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- McFarland, Curtis D. 1989. *A Frequency Count of Pilipino*. Manila: Linguistic Society of the Philippines.
- Mirikitani, Leatrice T. 1971. *Speaking Kapampangan*. Honolulu: University of Hawaii Press.
- Moguet, Pamela Johnstone and R. David Zorc. 1988. *Ilokano Newspaper Reader*. Kensington, Md: Dunwoody Press.
- Sunio, Delicia, and R. David Zorc. 1992. *Hiligaynon Reader*. Kensington, Md: Dunwoody Press.

Walton, Charles, A. 1977. Philippine Language Tree. Paper presented to the Austronesian Symposium at the University of Hawaii, August 18-20, 1977.

Wolff, John U. 1966. *Beginning Cebuano Part I*. New Haven: Yale University Press.

Zorc, R. David. *Cebuano Newspaper Reader*. Kensington, Md: Dunwoody Press.

The Puzzling Case of Chabacano: Creolization, Substrate, Mixing and Secondary Contact*

Patrick O. Steinkrüger
ZAS (Centre for General Linguistics, Berlin)

It is the aim of this paper to discuss some external and internal aspects of the varieties of Chabacano in the Philippines. The historical background of Chabacano will be considered (e.g. Andaya, 1993) and selected morphosyntactic features of Zamboangueño and Ternateño will be compared with other creoles in Asia which have a Malayo-Portuguese background and even other Spanish-based Creoles in the Atlantic. Three main features are investigated, namely; the pronominal system (e.g. Ternateño shows more similarities with Indo-Portuguese and Macaŕsta than with Zamboangueño), differential object-marking with the particle kun in Zamboangueño and Ternateño, and the function and form of the preverbal TMA-markers (ta-, ya-, ay- in Zamboangueño and ta-, (y)a-, di- in Ternateño). It turns out that the Chabacano varieties are without question related to other Portuguese-based creoles in Asia and that their current structure can only be explained by a multidimensional approach considering all the different contact situations in the past and in the present.

1. Preliminary remarks

If we investigate the origin of Chabacano, we must go beyond the current facts that its lexicon consists mainly of Spanish vocabulary and that its geographical distribution is restricted to the Philippines. Instead we must go back to the early colonial contacts in Asia in the 16th and 17th centuries when Portuguese and Malay — or more precisely, restructured forms of those languages — were used by Asians and Europeans to communicate. ‘The only languages which achieved any currency as *lingue franche* in the Eastern Seas during the sixteenth and seventeenth centuries were Portuguese and Malay, and most European traders, including the English, made use of Portuguese.’ (Whinnom 1956:7) Certainly Malay was already used in Manila and Central Philippines before the arrival of the Spanish (Wolff 1989). Compared to Portuguese, Spanish never had the same impact in SE-Asia (Baxter 1996, Lipski, Müllhäusler and Duthin 1996) and its use was restricted to the Philippines and Guam.

The putative forerunner of Chabacano is a Portuguese-based pidgin from Ternate in the Moluccas. In 1521 the Spanish arrived in Tidore and in 1529, with the treaty of Saragossa, they sold the islands to the Portuguese. In 1606, when the Dutch already controlled the islands, the Spanish forces captured the former Portuguese fort and

* Acknowledgments go to John Wolff for his detailed comments on a previous version of this paper; also many thanks to Boštjan Dvořák and Peter Slomanson. Of course, all errors remain mine.

deported the Ternate Sultan and his entourage to Manila. In 1607 the Dutch came back to Ternate where they, with the help of Ternateans, built a fort in Malayo. The island was divided between the two powers: the Spaniards were allied with Tidore and the Dutch with their Ternatean allies (see Andaya 1993:140ff for more details). Taking into account that the former Spanish colony was regularly connected with other Portuguese possessions in Asia, we shouldn't be surprised why to this day Chabacano and the Luso-Asiatic creoles share many structural features: 'The similarities in grammar and syntax, and even vocabulary of the Spanish contact vernaculars in the Philippines and Indo-Portuguese, are so many [...] that we can be quite certain that Ternateño [in the Moluccas] did develop out of the common Portuguese pidgin of the Eastern Seas.' (Whinnom 1956:21)

In the following sections, I will compare three of these common structural features, adding to the argument that Chabacano and other Luso-Asiatic varieties really have a lot in common and that the Philippine features are probably quite recent.

2. Personal pronouns

The internal and paradigmatic structure of Chabacano personal pronouns shows the historical connection of the Philippines with other regions in the world, especially with Portuguese possessions in Asia. The pronouns also demonstrate that a clear cut distinction between Portuguese and Spanish during the 16th and the 17th centuries is problematic for their restructured layers.¹ For example *vosotros* 'you (Pl)' did not exist in Portuguese, but is obviously the protoform of this pronoun in the Portuguese-based creoles in Malacca, Macau and India.

The subject pronouns of the following Iberoromance-based creoles will be compared in the following table (in this order): Ternateño (T), Caviteño (C), Macaísta (M) (Santos Ferreira 1978:23), Papia Kristang (PK) (Baxter 1988:52–53), Indo-Portuguese (Norteiro) (IP) (Dalgado 1906:154–155), Papiamentu (PP) (Munteanu 1996:295), Palenquero (P) (Pérez Tejedor 2004:56) and Zamboangueno (Z). I list only the full forms, not weak or cliticized forms:

Table 1. Free subject pronouns in Iberoromance-based creoles

singular							
T	C	M	PK	IP	PP	P	Z
<i>yo</i>	<i>yo</i>	<i>iou</i>	<i>yo</i>	<i>eu</i>	<i>mi</i>	<i>yo</i>	<i>(i)yo</i>
<i>bo, tédi</i>	<i>vos</i>	<i>vôs</i>	<i>bos</i>	<i>ós</i>	<i>bo</i>	<i>bo</i>	<i>(e)bós, tu, ustéd</i>
<i>éli</i>	<i>éle</i>	<i>ele</i>	<i>eli</i>	<i>éll(a)</i>	<i>ele</i>	<i>ele</i>	<i>éle</i>

¹ Already during the 15th century, the Reconnaissance-language, which is probably the forerunner of all Portuguese-based pidgins, was in constant contact with Spanish. (See some examples in Naro (1978), pp. 322–323; 324; 325.)

plural							
T	C	M	PK	IP	PP	P	Z
<i>mihótro</i>	<i>nisós</i>	<i>nôs</i>	<i>nus</i>	<i>nós</i>	<i>nos</i>	<i>suto</i>	<i>kitá</i> (incl.) <i>kamé</i> (excl.)
<i>tédi(s)</i>	<i>vusós</i>	<i>vosôtro</i>	<i>bolotu</i>	<i>usôt</i>	<i>boso</i>	<i>enú</i> (polite), <i>utére</i>	<i>kamó, ustédes</i>
<i>lohótro</i>	<i>ilós</i>	<i>ilôtro</i>	<i>olotu</i>	<i>illôt</i>	<i>nan</i>	<i>ané</i>	<i>silá</i>

Furthermore, other Creoles in Asia with a Portuguese lexical base show the same structure, e.g. the ones in Bidau (Dili, East Timor) and Tugu (Jakarta, Indonesia). (See Baxter, 1990:11–12, and for Indo–Portuguese, the compilation in Clements, 1996:3). With the exceptions of Palenquero² and Zamboangueno, where the whole set of the plural is taken from Niger–Congo³ and Malayo–Polynesian respectively, we can assume that the proto–paradigm of Iberoromance–based creoles in Asia was the following:

Table 2. Hypothetical proto–paradigm of free subject pronouns in Iberoromance–based creoles

singular	plural
<i>yo/mi</i> ‘I’	<i>nosotros/misotros</i> ‘we’
<i>vos</i> ‘you’	<i>vosotros</i> ‘you (PL)’
<i>ele</i> ‘s/he’	<i>(e)losotros</i> ‘they’

It seems that Whinnom was right that a “[...]mixed Portuguese–Spanish Pidgin[...]” (1956:9) was the base of Chabacano. And the clear–cut terms “Spanish–based” or “Portuguese–based” do not accurately convey the linguistic situation of the 16th and 17th centuries in the Iberian colonies and maybe the Iberian Peninsula in general.⁴ All Romance–based creoles in Asia show a reflex of *–otros* or similar as a plural marker of the pronouns — even the French–based creole of Réunion in the Indian Ocean (here the paradigm is *not*, *zot* and *zot*; from Stein, 1984:66).⁵ This reminds us of the structure of plural pronouns in some Eastern varieties of Malay

² Palenquero has a totally different history from the other creoles in Asia.

³ Unlike in Asia, most Afro–Portuguese creoles show a reflex of [n] in its plural pronouns. This is certainly an African influence, since this phoneme occurs in many languages on the continent, such as in Afroasiatic plural pronouns and as a plural marker of certain noun classes in Niger–Congo (Heine & Nurse 2000:88,40). In West–Atlantic languages of Niger–Congo such as Wolof, the phoneme is even present in the whole set of plural pronouns, both subject and object pronouns (Ngom 2003:41–43).

⁴ See, for example, the relatively high number of Portuguese and even Catalan words in JudeoSpanish.

⁵ Furthermore, Réunion is the only of the French–based creoles in the Indian Ocean which shows a certain amount of words of Indo–Portuguese origin (Corne 1999: 68; 71) and the only French–based creole with an object–marker for pronouns, namely *a* ‘to’ (cf. Stein 1984: 66).

where *orang* is added to the singular pronoun fulfilling the same function. But a more striking detail is the observation that this occurs in Ternate–Manado (John Wolff, p.c.)⁶ a variety spoken in the Moluccas. Here the Portuguese-based pidgin originated which later was imported to the Philippines.

3. Differential object marking (DOM) in Chabacano

One striking feature of Chabacano morphosyntax is the marking of objects with the lexical item ‘with’ which is phonologically *kon* or *kung*. It is the element itself, and also its distribution, which attract the linguist’s interest. As for the origin of the marker, there are several theories in the literature including my own view, which is point 1 in the list below. In the following, we shall revisit some arguments and observations concerning the object marker *kon* in Chabacano.

On the origin of this marker, there are the following arguments, which are not always contradictory.

- (1) (Partial) Spanish origin: In the 16th and 17th century some comitative objects were marked with *con*, such as with the verbs *ver* ‘see’ or *matar* ‘kill’. Chabacano, which arose during this time, could have grammaticalized this marker.
- (2) Luso-Asiatic origin: In many Luso-Asiatic creoles (e.g. Malacca, Macau, Batavia/Jakarta etc.) *com* is an object marker.⁷ As in Chabacano, *ku* in Papia Kristang marks “accusative, recipient, comitative, instrumental and goal.” (Baxter 1988:167). Only the comitative in Chabacano is expressed by *kumpanyéro* and *ubán* and not by *kon*. Once again, this is not surprising, since at least until around 1800, the Philippines had constant contact with Portuguese possessions in Asia. The place of origin could be South India from where it spread over to other creoles in Asia (hypothesis by Koontz–Garboden & Clements).
- (3) Hokkien–Chinese⁸ via Malay origin: Baxter (1995) argues that at least in the Portuguese–based creole of Malacca, *ku* was functionally influenced by the Bazaar Malay *sama* ‘with’.
- (4) Philippine origin: Recently, Mauro Fernández argued (informed by pers. comm.) that *kon* has its origin in phonologically and functionally similar markers for objects before proper names as *kay* and *kiná*. But we should say that there are some functional differences of both markers and it is quite improbable that this particle is of Philippine origin. A different situation is the marking of the object pronouns in

⁶ Also other restructured forms of Malay use *orang* ‘person’ as a plural marker. The agglutinating principle of the pronouns is also found in Hokkien–Chinese which adds *-n* to the singular forms.

⁷ See also Endruschat’s (1997) analysis of *com* ‘with’ in Afro–Portuguese, where the preposition is more frequent than in Europe and marks certain objects.

⁸ Hokkien–Chinese had also a direct influence on Philippine languages like Tagalog. See, for example, the examination by Yap (1980).

Zamboangueno with the prefix *ka-* which is certainly of Philippine — more exactly of Hiligaynon — origin.⁹

It seems that explanations 1, 2 and 3 are a part of the truth and that historically, some of them conditioned one another. Especially explanation 2 seems to be very convincing; if we look at the marker itself, we see that phonologically, the marker is identical in nearly all Luso-Asiatic creoles — except in some varieties of Indo-Portuguese in which we find *para* ‘for’ — and in Chabacano. This is further evidence of a former relationship between these creoles in Asia.

Table 3. Object-markers in some Ibero-Asian creoles

creole	object-marker
Ternateño	<i>kung</i>
Zamboangueno	<i>kon, kun</i>
Macaísta	<i>ko</i>
Malacca	<i>ku</i>

A further issue is the synchronic distribution of the marker. So far, we have made some general observations concerning the application of the marker:

- (1) Human (animate) objects are most frequently marked with *kon* (examples 3–10), but inanimate, topicalized objects can also be marked (examples 1 and 12).
- (2) Recipients (nearly always humans) and (human) objects of transitive constructions are both marked with *kon*. As in Spanish, the language tends to differentiate primary and secondary objects (cf. examples 6 and 7).
- (3) Double object constructions are always ungrammatical, as in Atlantic creoles and also in Spanish.
- (4) Double marking of objects (e.g. including the theme) in ditransitive constructions is possible but rare (i.e. marked; see example 2). In Spanish this would be ungrammatical.

Examples:

- (1) Zamboangueno
- | | | | | | | |
|-------|------------|------|------------|--------|----------|--------|
| Manga | paharyadór | lang | tapwéde | bisitá | konéste | lugár. |
| Manga | paharyadór | lang | ta- pwéde | bisitá | konéste | lugár |
| PL | hunter | only | IMPFV- can | visit | OBJ.this | place |

‘The place was visited only by hunters (of birds).’

⁹ However, unlike Hiligaynon, Zamboangueno has generalized this prefix in the whole paradigm, a fact which makes it semantically more transparent.

- (2) Zamboangueno
 Yadále yo kunéste líbro kunel héente.
 Ya- dále yo kunéste líbro kunel hénte
 PRF- give I OBJ.this book OBJ.DET man
 ‘I gave this book to the man.’
- (3) Zamboangueno
 Yamirá yo el hénte yaembuná konel irúq.
 Ya- mirá yo el hénte ya- embuná konel irúq
 PRF- see I DET man PRF- hit OBJ.DET dog
 ‘I saw how the man hit the dog.’
- (4) Zamboangueno
 Konosé le konése muhér.
 Konosé le konése muhér
 know s/he OBJ.that woman
 ‘S/he knows that woman.’
- (5) Zamboangueno
 Paula taamá kun Pédro.
 Paula ta- amá kun Pédro
 Paula IMPFV- love OBJ Pedro
 ‘Paula loves Peter.’
- (6) Zamboangueno
 Si Paula talibá líbro kon Pédro.
 Si Paula ta- libá líbro kon Pédro
 A Paula IMPFV- carry book OBJ Pedro
 ‘Paula carries Peter the/a book.’
- (7) Zamboangueno
 Talibá yo un butélya de ágwa kun mi nána.
 Ta- libá yo un butélya de ágwa kun mi nána
 IMPFV- carry I ART bottle of water OBJ my mother
 ‘I carry a bottle of water to my mother.’
- (8) Zamboangueno (Forman 1972: 157)
 Yadále abíso el maga páharo konel maga peskadór.
 Ya- dále abíso el maga páharo konel maga peskadór
 PRF- give warning DET PL bird OBJ.DET PL fisherman
 ‘The birds gave a warning to the fishermen.’

- (9) Zamboangueno
 Dále komígo ágwa!
 Dále komígo ágwa
 give OBJ.1SG water
 ‘Give me water!’
- (10) Caviteño
 María yaregalá un relós cun su nóbio.
 María ya- regalá un relós cun su nóbio
 Mary PRF- give ART watch OBJ her boyfriend
 ‘Mary gave her boyfriend a watch.’
- (11) Zamboangueno (Forman 1972: 199)
 Yamirá silá síne.
 Ya- mirá silá síne
 PRF- see they movie
 ‘They saw a movie.’
- (12) Zamboangueno (Forman 1972: 157)
 Yaasé le kebrá konel básó.
 Ya- asé le kebrá konel básó
 PRF- cause s/he break OBJ.DET glass
 ‘He broke the glass.’
- (13) Zamboangueno (Forman 1972: 199)
 Yamirá le el páto.
 Ya- mirá le el páto
 PRF- see s/he DET duck
 ‘He saw the duck.’
- (14) Zamboangueno (Forman 1972:167)
 Yasaká le konel muhér konel dragón.
 Ya- saká le konel muhér konel dragón
 PRF- seize he OBJ.DET woman OBJ.DET dragon
 ‘He seized the woman from the dragon.’

4. The TMA markers in Chabacano

What makes it sometimes difficult to characterize a given feature in Chabacano as “typically creole” is the fact that some Philippine languages, like Tagalog or Cebuano, share these features with many creole languages. For example, there is no copula for equative constructions.

This is also exactly the case if we look at the preverbal TMA markers in Chabacano. As in many Philippine languages they are preverbal and behave morphophonologically as prefixes. Since some data in Schuchardt (1883) and Whinnom (1956) show that the marker could have been interrupted by other elements, we can assume that this status is recent. So it seems superficially that Chabacano follows a Philippine blueprint.

At the synchronic level, already Whinnom (1956:92–3), referring to López' description of Tagalog grammar, states that the function of the Chabacano verb with its TMA markers are like those in Tagalog, admitting that “contact–vernacular grammar is not Tagalog grammar but in every single instance represents a simplification of Tagalog practice.” (Whinnom 1956:98). Functionally, the three forms of the marked verb in Chabacano seem to coincide with those as we know them from Tagalog, where affixed finite verbs do occur also in three forms, namely as irrealis, imperfective and perfect (see Schachter 1993:1419). But the morphological system of the TMA markers in Chabacano functions quite differently from those in Philippine languages.¹⁰

As in the case of object marking, it is necessary to compare the structure of the varieties of Chabacano with other Iberoromance–based creoles in Asia and even in other areas if we assume that creolization took place during the three first centuries of Portuguese and Spanish colonization worldwide. A selection of some creoles of Asia is compiled in the following table, adding the varieties of Chabacano to the compilation by Hancock (1975):

Table 4. TMA markers in some Iberoromance–based creoles in Asia

Creole	imperfective	perfect(ive)	irrealis	completive	combinations
Zamboangueno	<i>ta</i>	<i>ya</i>	<i>ay</i>	<i>(kabá)</i>	-
Ternateño	<i>ta</i>	<i>(y)a</i>	<i>di</i>	<i>(kabá)</i>	-
Indo-Portuguese (Bombay)	<i>ta</i>	<i>ja</i>	<i>di</i>	<i>kaba</i>	(-?)
Macaísta	<i>ta</i>	<i>ja</i>	<i>lô</i>	<i>kaba</i>	(-?)
Papia Kristiang	<i>ta</i>	<i>ja</i>	<i>lo(gu)</i>	<i>kaba</i>	-
Batavia/Jakarta	<i>ta</i>	<i>ja</i>	<i>lo(gu)</i>	<i>kaba</i>	(-?)

Unlike Atlantic creoles or those in the Indian Ocean, the combination of the markers is impossible or highly restricted, as Hancock (1975) argues: “They [the tense and aspect markers] may not combine as freely as in other creoles, although further research is needed before their behaviour may be fully determined.” (p. 221) Examples of application are found in the following:

¹⁰ Following John Wolff there is little congruence between Chabacano and Philippine tense–aspect systems.

(15) Zamboangueno

Y el bátaq yaakustá na káma kun péhro y
 Y el bátaq ya- akustá na káma kun péhro y
 And DET child PRF- lay LOC bed with dog and

tadurmí.

ta- durmí
 IMPF- sleep

‘And the child lays in bed with a dog and is sleeping.’

(16) Zamboangueno

Yauyí yo talyurá el dalagíta.
 Ya- uyí yo ta- lyurá el dalagíta.
 PRF- hear I IMPF- cry DET girl

‘I heard the girl crying.’

(17) Caviteño

Diandá yo na pláza.
 Di- andá yo na pláza
 IRR- go I LOC market/place

‘I will go to the market.’

(18) Zamboangueno

Yaatrasá era le na bus.
 Ya- atrasá era le na bus
 PRF- miss MOD s/he LOC bus

‘He would have missed the bus’.

(19) Zamboangueno

El péhro tabuská konel palakáq.
 El péhro ta- buská konel palakáq
 DET dog IMPFV- search OBJ.DET frog

‘The dog is looking for the frog.’

(20) Ternateño

Kel péhro amurdé kung kel muhér.
 Kel péhro a- murdé kung kel muhér
 DET dog PRF- bite OBJ DET women

‘The dog bit the woman’

(21) Ternateño

Takedá éle na Maníla.
 Ta- kedá éle na Maníla
 IMPFV- live s/he LOC Manila

‘S/he lives in Manila.’

(22) Ternateño

Kung yo takedá ríko, dimerká yo kása.
 Kung yo ta- kedá ríko di- merká yo kása
 COMPL I IMPFV- become rich IRR- buy I house

‘If I were/became rich, I would buy a house.’

(23) Ternateño

Akresé yo na Barra.
 A- kresé yo na Barra
 PRF- grow.up I LOC Ternate

‘I grew up in Ternate’.

5. The different contact situations of Chabacano

It should be clear by now that the modern varieties of Chabacano are not the result of one single contact scenario, but rather the product of a much more complex scenario involving different contact situations in the past and in the present.

5.1. Creolization

Ternateño arrived probably as a creolized language in the Philippines (see Whinnom 1956 and Rafael 1978): “[...]the community in which the contact vernacular became creolized, that is the Ternateño community [in the Moluccas; P.S.]” (Whinnom 1956:10) or: “[...]the first of the Spanish contact vernaculars in the Eastern Seas arose in Ternate, and had already achieved creolization before the evacuation of the island by garrison and Christian inhabitants.” (Whinnom 1956:10). This is unlike the case of Zamboangueno, where creolization took place in situ (this is the assumption e.g. of Warren 1981 and more recently by Mauro Fernández, p.c.). Here, escaped slaves, originating from the Visayas and also Malayic-speaking islands, contributed to the heavy substrate influence of Hiligaynon (especially the pronoun system) and also Cebuano in the language (for all these features see Frake 1971, Bowen 1971, Bunye and Yap 1971 and Wolfende 1971). We can doubt this theory, though, since all varieties of Chabacano are structurally alike and all of them share typical features with the other Luso-Asiatic Creoles. However, Warren’s analysis would explain why Zamboangueno shows so many elements from Hiligaynon (especially pronouns) and Cebuano.

In addition, since the varieties of Chabacano share many structural features, it is quite probable that they were formerly in contact with each other. This is at least the case for Cavite and Zamboanga via navigation. Maybe an older creole in Zamboanga

was overlaid by a more recent one; the language in Zamboanga is phonologically and lexically closer to modern Spanish or Caviteño than Ternateño.

So, finally, what are the substrate languages of Chabacano? It is quite problematic to apply a mere sociolinguistic definition of a substrate as formulated by Holm (1988:5): “Usually those with less power (speakers of the substrate) are more accommodating and use words from the language of those with more power (the superstrate), although the meaning, form, and use of these words may be influenced by the substrate languages.” Did the substrate language speakers of Chabacano really have less power? We would need more information about the social impact of the substrate during the process of creolization to determine that. At the grammatical and lexical level, we have at least three different substrates of Chabacano:

- (1) Malayic (also in its restructured form)
- (2) Visayan languages (only for Zamboangueno)
- (3) Hokkien Chinese (influences on lexicon and grammar, e.g. object marking)

5.2. Intertwining

In the case of Zamboangueno, some scholars have even claimed that it is a mixed or intertwined language:

There are languages that justify a type of ML [mixed language] comprising mixed creoles. An example is Chavacano, also known as Zamboangueno or Mindanao Creole Spanish [...] widely spoken on the island Mindanao in the Philippines [...]. The source of most lexical and grammatical morphemes is Spanish Creole, but the syntactic framework and relations between categories (for instance animacy and definiteness, aspect) tend to follow an Austronesian blueprint. Austronesian grammatical morphemes include plural markers, animate definite articles, past-tense existential verbs, and transitive derivation markers, while the pronoun system is mixed: The singular pronouns are all based on Spanish Creole. In the plural, Austronesian forms seem to co-exist with Spanish Creole forms (e.g. second person *kamo* alongside *ustedes*), but the first person plural, Austronesian forms, which mark the exclusive/inclusive distinction (*kamí*, *kitá*) seem to be preferred. (Matras and Bakker 2003:11)

Apart from the sociolinguistic setting — usually an exactly bicultural search of identity — also structurally, this assumption seems to be highly problematic. The criticism of the above quotation consists of the following points:

- (1) Do most grammatical morphemes really have a Spanish creole source? This is only the case for the TMA markers of the verb; all the rest is Austronesian.
- (2) The relations between the categories of animacy, definiteness and aspect is already known in Spanish. It is not necessarily Austronesian.

- (3) Other things are not entirely clear: With the term “animate definite article” the authors probably refer to *si* before proper names as subjects. However, past tense existential verbs are not necessarily of Austronesian origin. The only ones known to me are *nuay* (< Span. *no hay*) and *estaba* (< Span. *estaba* ‘was, stayed’). The latter is actually only used as a locative of the past.
- (4) As for the personal pronouns, in the singular, Austronesian forms are also in use; instead of *tu*, many young speakers in Zamboanga use the Visayan *ka* after verbs.

Nevertheless, it is true that Chabacano has in some respect a mixed character as a result of language contact in the Philippines, but this fact is not an argument against its original structure as a creole. Besides, code switching (as opposed to intertwining) with English, as in example 24, is quite widespread among young speakers in Zamboanga:

- (24) Zamboangueno
- | | | | | | | | | |
|------|---------|----------|-----------|--------|---------|---|-----|------|
| Dále | kumígo | ditúyo | cellphone | number | para | I | can | call |
| Dále | kumígo | ditúyo | cellphone | number | para | I | can | call |
| Give | OBJ.1SG | POSS.2SG | cellphone | number | so.that | I | can | call |
- you later.
 you later
 you later

‘Give me your cellphone number so that I can call you later.’

Also in Ternate, many speakers mix their language with Tagalog or they switch. But there is a big difference between old and young speakers and this is one piece of evidence that this intensive contact with Philippine languages — Tagalog in Ternate and Cebuano in Zamboanga — is probably not very old.¹¹

5.3. Secondary contact or adstrate

We can assume that the ‘Philippine character’ of Chabacano is a more recent phenomenon which we call ‘secondary contact’ or ‘adstrate’. As for Ternateño, the impact of the neighbouring language Tagalog began in the late 19th century. By the 20th century nearly all Ternateños had become bilingual. In the case of Zamboangueno, the situation is somewhat different, since Visayan languages participated in the formation of the language (personal pronouns, derivational affixes, morphophonology, discourse particles etc.). Therefore Zamboangueno currently shows a much more ‘Philippine character’ than Ternateño does. Also here, Chabacano native speakers have become the minority (under fifty percent of the current population).

¹¹ Some older people in Ternate remember their grandparents who were born in the 1870s or 1880s and only knew Chabacano and some Spanish. Consequently, we can assume that this intensive contact began around 1900.

This last fact, combined with the influence of mass media and education, has led to further secondary contact; all varieties of Chabacano are exposed to English and Tagalog, and in Zamboanga, to Cebuano as well. Monolingual speakers of Chabacano nowadays hardly exist. One result is massive borrowing and code-switching (see example 1). For this reason, it is not surprising that we find quite a lot of ‘Philippine features’ in all the varieties of Chabacano at many different levels, for example:

- phonetics and phonology (e.g. processes like palatalization and intonational patterns)
- morphophonology (e.g. affixation)
- morphology (e.g. derivation in Zamboangueno; see Steinkrüger 2003)
- syntax (e.g. VSO basic word order, P2 movement and the use of the potentive mode in Zamboangueno; see Rubino 2005)
- pragmatics (e.g. application of discourse particles like *ba* or *daw*)

6. Some conclusions concerning the different contact situations of Chabacano

The case of Chabacano is still an interesting challenge for contact linguistics and its structure can only be explained by a multidimensional approach. But again, concerning the Philippine features of Chabacano, if we look at the structural similarities of Chabacano with other Portuguese-based creoles in Asia, we should ask how old or how recent the features of Philippine origin are exactly. There is also more investigation needed on the interaction of both Iberoromance languages in Southeast-Asia during the 16th and 17th centuries; apart from the pronouns, are there more elements in the Portuguese-based creoles which show the potential influence of Spanish? More details on the importance of Hokkien-Chinese during the formation of the creoles is also needed. Finally, the results of 500 years of contact with European languages should be considered in greater detail in Austronesian studies (e.g. in the handbook by Adelaar and Himmelmann 2005) and not just in creole studies.

Abbreviations

A	agent	MOD	modal
ART	article	OBJ	object-marker
COMPL	complementizer	PL	plural
DET	determiner	POSS	possessive
IMPFV	imperfective	PRF	perfect
IRR	irrealis	SG	singular
LOC	locative		

References

- Adelaar, Alexander and Nikolaus P. Himmelmann, eds. 2005. *The Austronesian Languages of Asia and Madagascar*. London and New York: Routledge.
- Andaya, Leonard Y. 1993. *The World of Maluku: Eastern Indonesia in the Early Modern Period*. Honolulu: University of Hawai'i Press.

- Baxter, Alan N. 1988. *A Grammar of Kristang (Malacca Creole Portuguese)*. Canberra: Australian National University.
- Baxter, Alan N. 1990. Notes on the Creole Portuguese of Bidau, East Timor. *Journal of Pidgin and Creole Languages* 5(1): 1–38.
- Baxter, Alan N. 1995. Um Importante Sincretismo no Português Crioulo de Malacca: a Preposição Multifuncional *ku*. In *Miscelânea de Estudos Lingüísticos, Filológicos e Literários in memoriam Celso Cunha*, ed. by Celine da Cunha Pereira and Paulo Roberto Pereira, 15–33. Rio de Janeiro: Nova Fonteira.
- Baxter, Alan N. 1996. Portuguese and Creole Portuguese in the Pacific and Western Pacific rim. In *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, ed. by Stephen A. Wurm, Peter Mühlhäusler and Darrell T. Tryon, 299–338. Berlin and New York: Mouton de Gruyter.
- Bowen, J. Donald. 1971. Hispanic Languages and Influence on Oceania. In *Current Trends in Linguistics* (vol. 8), 938–52. The Hague: Mouton.
- Bunye, Maria Victoria R. and Elsa Paula Yap. 1971. *Cebuano Grammar Notes*. Honolulu: University of Hawai'i Press.
- Chan Yap, Gloria. 1980. *Hokkien Chinese Borrowings in Tagalog*. Canberra: Australian National University.
- Clements, J. Clancy. 1996. *The Genesis of a Language. The Formation and Development of Korlai Portuguese*. Amsterdam/Philadelphia: Benjamins.
- Corne, Chris. 1999. *From French to Creole*. Westminster: University Press.
- Dalgado, Sebastião Rodolpho. 1906. Dialecto Indo-Português do Norte. *Revista Lusitana* 9(3–4): 142–228.
- Endruschat, Annette. 1997. Die Präposition *com* im Afroportugiesischen: Versuch einer neuen Interpretation ihrer spezifischen Verwendung. In *Afrolusitanistik – eine vergessene Disziplin in Deutschland?*, ed. by Ruth Degenhardt, Thomas Stolz and Hella Ulferts, 243–278. Bremen: Universitätsverlag.
- Forman, Michael L. 1972. *Zamboangueño Texts with Grammatical Analysis: A Study of Philippine Creole Spanish*, PhD dissertation at Cornell University.
- Frake, Charles O. 1971. Lexical Origins and Semantic Structure in Philippine Creole Spanish. In *Pidginization and Creolization of Languages*, ed. by Dell Hymes, 223–242. Cambridge: Cambridge University Press.
- Francisco, Michael, C. 2002. *Cavite en Siglo 19*. Cavite: De La Salle University.

- Hancock, Ian F. 1975. Malacca Creole Portuguese: Asian, African or European? *Anthropological Linguistics* XVII: 211–236.
- Heine, Bernd and Derek Nurse (eds.). 2000. *African languages*. Cambridge: Cambridge University Press.
- Holm, John. (1988). *Pidgins and Creoles. Vol. 1: Theory and Structure*. Cambridge: Cambridge University Press.
- Koontz-Garboden, Andres J. & J. Clancy Clements. (2002). *Some theoretical implications of adpositions in Spanish and Portuguese-based creoles*. Paper presented at the Society of Pidgin and Creole Languages conference, 4–6th January 2002 San Francisco.
- Lipski, John, Peter Mühlhäusler and F. Duthin F. 1996. Spanish in the Pacific. In: *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, ed. by Stephen A. Wurm, Peter Mühlhäusler and Darrell T. Tryon, 271–298. Berlin and New York: Mouton de Gruyter.
- Lipski, John M. 1988. Philippine Creole Spanish: Assessing the Portuguese Element. *Zeitschrift für Romanische Philologie* 104: 25–45.
- Lipski, John M. 1992. New Thoughts on the Origins of Zamboangueno (PCS). *Language Sciences* 14(3): 197–231.
- López, Cecilio. 1965. The Spanish Overlay in Tagalog. *Lingua* 14: 467–504.
- Matras, Yaron and Peter Bakker. 2003. *The Mixed Language Debate. Theoretical and Empirical Advances*. Berlin and New York: Mouton de Gruyter.
- Munteanu, Dan. 1996. *El Papiamento, Lengua Criolla Hispánica*. Madrid: Gredos.
- Naro, Anthony J. 1978. A Study on the Origins of Pidginization. *Language* 54(2): 314–347.
- Ngom, Fallou. 2003. *Wolof*. München: Lincom.
- Pérez Tejedor, Juna Pabla. 2004. *El Criollo de Palenque de San Basilio: Una Visión Estructural de su Lengua*. Bogotá: Universidad de los Andes.
- Rafael, Vicente L. 1978. From Mardicas to Filipinos: Ternate, Cavite in Philippine History. *Philippine Studies* 26: 343–362.
- Riego de Dios, María. 1989. A Composite Dictionary of Philippine Creole Spanish (PCS). *Studies in Philippine Linguistics* 7(2): 1–210.
- Rubino, Carl. 2005. The Potentive Mode in Zamboangueno. Paper presented at the *International Creole Conference*, June 2005, Leipzig (Germany).

- Santos Ferreira, José dos. 1978. *Papiá Cristám di Macau. Epítome de Gramática Comparada e Vocabulário*. Macau: Fundação Macao.
- Schachter, Paul. 1993. Tagalog. In *Syntax* (vol. 2), ed. by Arnim v. Stechow et al., 1418–1430. Berlin: Walter de Gruyter.
- Schuchardt, Hugo. 1883. Kreolische Studien IV: Über das Malaiospanische der Philippinen. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien* 105: 111–150.
- Stein, Peter. 1984. *Kreolisch und Französisch*. Tübingen: Niemeyer.
- Steinkrüger, Patrick O. 2003. Morphological Processes of Word Formation in Chabacano (Philippine Spanish Creole). In *Phonology and Morphology of Creole Languages*, ed. by Ingo Plag, 253–268. Tübingen: Niemeyer.
- Steinkrüger, Patrick O. 2004. s.v. Philippine Spanish Creoles. In *Encyclopedia of Linguistics*, ed. by Philipp Strazny. New York: Routledge.
- Steinkrüger, Patrick O. (2007). Notes on Ternateño (a Philippine Spanish Creole). *Journal of Pidgin and Creole Languages*. 22:2, 367–378.
- Thompson, Roger M. 2003. The Spanish Overlay. In: idem. *Filipino English and Taglish*, 59–66. Amsterdam and Philadelphia: Benjamins.
- Warren, James Francis. 1981. *The Sulu Zone, 1768–1898*. Singapore: Singapore University Press.
- Whinnom, Keith. 1956. *Spanish Contact Vernaculars in the Philippine Islands*. Hong Kong: Hong Kong University Press.
- Wolfenden, Elmer Paul. 1971. *Hiligaynon Reference Grammar*. Honolulu: University of Hawai'i.
- Wolff, John U. 1989. Malay Borrowings in Tagalog. In *Southeast Asian History and Historiography* (Festschrift for D.G.E. Hall), ed. by C.D. Cowan and O.W. Wolters, 345–367. Ithaca and London: Cornell University Press.

The Sambalic Languages of Central Luzon

Roger Stone
SIL International

This paper examines the Sambalic languages of Central Luzon in order to ascertain the degree to which the languages are lexically similar to each other, to Ivatan, and to the languages of wider communication (LWCs); specifically, Kapampangan, Ilokano and Tagalog. The study compares 268 lexemes from eleven languages, and computes cognate percentages within seven grammatical categories. The data presented in the paper is used to hypothesize subgroupings within the Sambalic language family. While no attempt is made to reconstruct the past history of the Ayta and Sambalic languages, it is hoped that this paper will help lay a foundation for other linguists to do such a study.

1. Introduction

Little published data exists on the Sambalic languages and virtually no published data exists on the Ayta languages. This study aims to suggest possible subgroupings within the Sambal language family based on lexical similarity. It has also been hard to determine what languages belong in the same subgrouping with the Sambalic languages. Many have included the Bashiic languages with Sambal under the Northern Philippines branch (Rubino 2005, Rubrico 2005, Gordon 2005, Zorc 1977), although Blust (1991) suggests there is no convincing evidence to make this claim. Many have included Kapampangan directly with Sambal in the Central Luzon languages (Rubino 2005, Rubrico 2005, Gordon 2005), while others (Llamzon 1978) have rejected this hypothesis. Because of these questions, Kapampangan and Ivatan (Bashiic) are included in this comparative study of the Sambal languages. The languages of wider communication (LWCs) Tagalog, and Ilokano are also included in the study.

The methodology used to compare these languages has much in common with McFarland's (1974) work on the Bikol languages and Zorc's (1977) work on the Bisayan Dialects in that comparisons are made by grammatical category. Seven grammatical categories are studied, namely nouns, verbs, adjectives, pronouns, particles conjunctions and interrogatives. Cognate percentages are determined separately for each grammatical category.

2. The Sambalic Language Family

The following languages are generally accepted to form the Sambalic language family, although Ayta Magbukun, which has not been included in this study, is likely to also belong to the group. Note also that there are a few alternate spellings for some of the languages. (See, for example, Gordon, 2005).

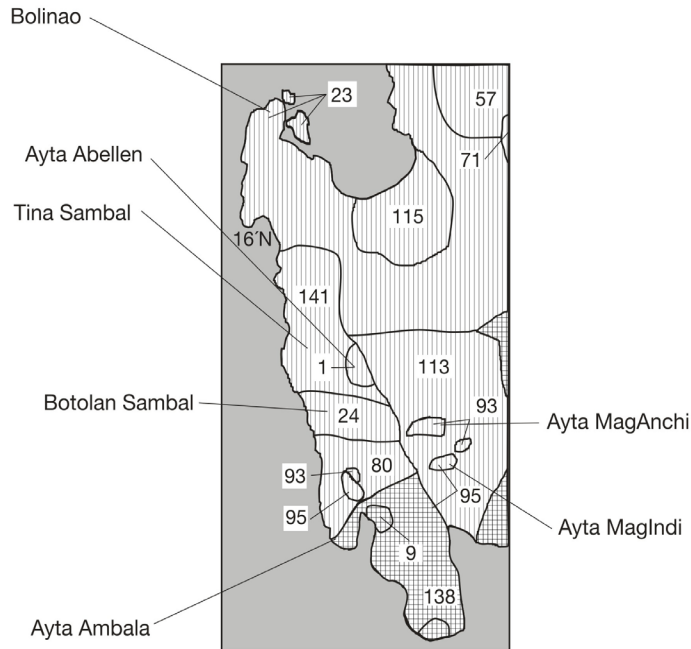


Figure 1. Sambalic Languages of Central Luzon

2.1 Ayta Abellen

The speakers of Ayta Abellen are located in the mountainous western part of the Tarlac province of Luzon with the area reaching into the province of Zambales even on the western side of the Zambales mountains (Nitsch 1998). There are about 3500 speakers of the language (Stone 2005).

2.2 Ayta Ambala

Ayta Ambala is spoken by more than 2,000 Ayta living in the provinces of Bataan and Zambales (Ramos 2005).

2.3 Ayta MagAntsi

Ayta MagAntsi is spoken by approximately 4,200 Ayta in a number of sitios (small villages) in the provinces of, Tarlac, Pampanga and Zambales. (Storck and Storck 2005).

2.4 Ayta MagIndi

The speakers of Ayta MagIndi live in the province of Pampanga in the areas of Nabuklod, Floridablanca, Pasbul, Planas, Kamias, and other areas. The Ayta MagIndi speakers number approximately 5000 (Green 1991).

2.5 Bolinao

The Bolinao language is spoken by approximately 50,000 people (Gordon 2005) located in the municipalities of Bolinao and Anda, Pangasinan (Persons 1998).

2.6 Botolan Sambal

Botolan Sambal is spoken by 32,867 people (Gordon 2005) who live around the municipality of Botolan in the Zambales province. There is also a group of Aytas living in the upland areas of Zambales who speak this language. (Antworth 1979)

2.7 Tina Sambal

Tina Sambal is spoken by 70,000 people (Gordon 2005) located in the Northern half of the Zambales province and in two barangays (villages) across the border in Pangasinan (Goschnick 2005).

3. Data

Initial survey for the Ayta languages was done by Kurt and Margaret Storck in the early 1980s. The data collected by them and others laid the foundation for the delineation of language boundaries for the 5 Ayta languages. However, that survey data was never published and a review of it made this researcher realize that it would be better to base an analysis on more current data, including the Storcks' outstanding work on MagAntsi (2005) rather than the old wordlists collected many years ago. But even having access to the latest lexicography work done for each of these languages was still deemed inadequate, so for most of the language data presented here, the dictionary data was verified with a resident linguist or face to face with a native speaker of the language. (See Appendix 2 for Data Acknowledgements.) It is still acknowledged, however, that despite these efforts, there may be inaccuracies in the data.

The findings of this study, as well as the data collected (which is included here in appendix 1), are intended to help modify the language groupings and lexical similarity percentages presented in the Ethnologue (Gordon 2005) which is based on the older survey data.

4. Method for Determining Cognate percentages

For determining which lexical forms are cognate with each other, syllolexicostatistics was used. Syllolexicostatistics is a hybrid of syllostatistics and lexicostatistics and is currently being developed by Noel Mann (Mann 2005). The process involves the following steps:

- 1) Split root words into their constituent syllables.
- 2) Compare syllables for apparent cognicity.
- 3) Categorize syllable pair into one of three categories as follows:

Category 1: Probable match of cognate syllables; onset and rhyme elements match – that is, they map in a predictable or recurrent manner.

Category 2: Possible match of cognate syllables; either onset or rhyme elements match

Category 3: Unlikely match of cognate syllables; onset and rhyme do not match

- 4) Use the following chart to determine cognicity. The chart shows the minimum requirements for two words to be regarded as cognates.

Table 1. Cognicity Criteria

Syllables	Category 1	Category 2	Category 3
1	1	0	0
2	1	1	0
3	1	1	1
4	2	1	1
5	2	2	1
6	2	2	2

After determining cognicity as per the method outlined above, the data was inputted into the computer program WordSurv, which then organized, grouped, and computed cognate percentages for each language. Below are the results and observations for the seven grammatical categories analyzed.

5. Results by Grammatical Category

Presented here are the cognate percentages for the languages, grouped initially according to grammatical category. Composite cognate percentages are included at the end. Within the results tables, the bold section represents the Sambalic languages.

5.1 Nouns

129 common nouns were selected for analysis. This is the largest set of words gathered and analyzed for any of the grammatical categories included in this study.

Almost one third of all lexical items here have a unique form that is repeated among Sambalic languages but not cognate with the same lexical item in any of the LWCs studied.

The cognate percentages of Tagalog, Kapampangan, and Ilokano nouns with nouns in the Sambal languages range between 38–51%, whereas within the Sambal languages, the nouns are between 63–87% cognate. Ivatan nouns are less than 30% cognate with Sambal language nouns.

Table 2. Noun Cognate Percentages

	Bol	T.S.	Abel	B.S.	Ants	Indi	Amba	Kap	Iva	Ilo	Tag
Bolinao	100%	78%	66%	71%	63%	64%	65%	48%	26%	43%	43%
Tina Sambal	78%	100%	67%	75%	64%	64%	66%	48%	28%	42%	43%
Ayta Abellen	66%	67%	100%	80%	73%	70%	69%	45%	26%	44%	41%
Botolan Sambal	71%	75%	80%	100%	78%	78%	77%	47%	28%	44%	43%
Ayta MagAntsi	63%	64%	73%	78%	100%	87%	72%	49%	25%	38%	41%
Ayta MagIndi	64%	64%	70%	78%	87%	100%	76%	51%	25%	38%	41%
Ayta Ambala	65%	66%	69%	77%	72%	76%	100%	45%	27%	42%	45%
Kapampangan	48%	48%	45%	47%	49%	51%	45%	100%	19%	32%	43%
Ivatan	26%	28%	26%	28%	25%	25%	27%	19%	100%	30%	26%
Ilokano	43%	42%	44%	44%	38%	38%	42%	32%	30%	100%	38%
Tagalog	43%	43%	41%	43%	41%	41%	45%	43%	26%	38%	100%

5.2 Verbs

In researching verbs, only root forms were considered, in order to avoid differences due to affixation. 48 lexical entries were analyzed. For these verbs where there is a cognate form common to a majority of the Sambal languages, 38% do not have a corresponding cognate form in the regional LWCs.

Within the verbs, high cognate percentages for Ayta languages with each other can be seen (73–88%) as well as between Bolinao and Tina Sambal (73%). But here, Botolan Sambal groups naturally with the Ayta languages since the cognate percentages are much higher with the other Ayta languages (73–85%) than with the coastal lowland counterparts of Tina Sambal (52%) and Bolinao (52%).

Kapampangan verbs are 44–52% cognate with Sambal language verbs. Ivatan verbs are not more than 27% cognate with any of the Sambal language verbs.

Table 3. Verb Cognate Percentages

	Bol	T.S.	Abel	B.S.	Ants	Indi	Amba	Kap	Iva	Ilo	Tag
Bolinao	100%	73%	56%	52%	54%	58%	54%	48%	25%	42%	40%
Tina Sambal	73%	100%	54%	52%	52%	56%	50%	50%	27%	40%	42%
Ayta Abellen	56%	54%	100%	83%	88%	77%	77%	50%	27%	40%	44%
Botolan Sambal	52%	52%	83%	100%	85%	73%	77%	48%	25%	38%	42%
Ayta MagAntsi	54%	52%	88%	85%	100%	83%	79%	52%	27%	40%	42%
Ayta MagIndi	58%	56%	77%	73%	83%	100%	75%	54%	27%	40%	40%
Ayta Ambala	54%	50%	77%	77%	79%	75%	100%	44%	25%	42%	42%
Kapampangan	48%	50%	50%	48%	52%	54%	44%	100%	25%	25%	40%
Ivatan	25%	27%	27%	25%	27%	27%	25%	25%	100%	23%	17%
Ilokano	42%	40%	40%	38%	40%	40%	42%	25%	23%	100%	33%
Tagalog	40%	42%	44%	42%	42%	40%	42%	40%	17%	33%	100%

5.3 Adjectives

For the 36 adjectives selected, Sambal adjectives exhibit less similarity with each other than for most of the other grammatical categories. Lexical similarity ranges from 44% to 72% within the Sambalic languages.

Ayta MagIndi borrows heavily from Kapampangan (53% cognate) and its adjectives are more cognate with Kapampangan than with two of the other Sambalic languages.

Kapampangan adjectives, however, do not have high cognate percentages with other Sambalic languages; its percentages range between 28% and 42%. Ivatan cognate scores with Sambalic languages are especially low, ranging from 6% to 17%.

Table 4. Adjective Cognate Percentages

	Bol	T.S.	Abel	B.S.	Ants	Indi	Amba	Kap	Iva	Ilo	Tag
Bolinao	100%	58%	50%	58%	47%	50%	47%	33%	11%	33%	33%
Tina Sambal	58%	100%	56%	64%	50%	44%	56%	31%	6%	19%	28%
Ayta Abellen	50%	56%	100%	64%	50%	44%	53%	28%	11%	17%	31%
Botolan Sambal	58%	64%	64%	100%	67%	53%	72%	36%	11%	19%	42%
Ayta MagAntsi	47%	50%	50%	67%	100%	61%	61%	33%	11%	14%	36%
Ayta MagIndi	50%	44%	44%	53%	61%	100%	53%	53%	17%	19%	39%
Ayta Ambala	47%	56%	53%	72%	61%	53%	100%	42%	11%	17%	47%
Kapampangan	33%	31%	28%	36%	33%	53%	42%	100%	17%	22%	58%
Ivatan	11%	6%	11%	11%	11%	17%	11%	17%	100%	11%	11%
Ilokano	33%	19%	17%	19%	14%	19%	17%	22%	11%	100%	19%
Tagalog	33%	28%	31%	42%	36%	39%	47%	58%	11%	19%	100%

5.4 Pronouns

Sambalic languages have nearly identical pronoun systems. The differences that exist between the Sambalic languages are predominantly due to phonological changes. The differences between Sambalic pronouns and those of neighboring LWCs are greater and cannot be summarized so easily.

Sambalic languages have a fronted pronoun paradigm. This paradigm also exists in Ilokano, Kapampangan, and Ivatan, although many of the forms in these latter languages have multiple differences in comparison to the Sambalic languages.

For focused 1st person exclusive pronouns, there is a split distribution among Sambalic languages. Tina Sambal, Bolinao, and Ayta Ambala use *kami*, whereas all other languages use *kayi*. For focused 2nd person plural pronouns, Tina Sambal and Bolinao use *kamo* whereas the other languages use *kao* or *kawo*.

For possessive 1st person exclusive pronouns, there is a split distribution. Ayta MagAntsi and Botolan Sambal both have a *nawen* form, whereas the rest of the Sambalic languages except for MagIndi have *mi*. MagIndi's *yan* form is unique to all languages researched.

For possessive 2nd person plural pronouns, there is also a split distribution. Tina Sambal and Bolinao both have a *moyo* form, whereas all other Sambalic languages have *yo*.

For direction/beneficiary focus pronouns, Sambal pronouns have a *kVn-* prefix. The vowel, however, is not uniform. For instance, the languages along the coast (Bolinao, Tina Sambal, and Botolan Sambal) all have the vowel *o* as in *kon-*. But the inland languages (Ayta) all have *kan-*. Since this is a predictable phenomenon, pronouns with both of these prefixes were counted as cognates of each other.

An interesting phonological difference is that some languages have the /s/ phoneme in their pronouns, while others don't. MagIndi and Bolinao have pronoun forms with the phoneme /s/, while all other Sambal languages use /h/ in its place.

In terms of cognate percentages for pronouns, the Sambal languages all have high cognate percentages with each other, ranging from 81–97%. Kapampangan has several forms that at first glance do not seem to be cognate with the Sambal languages, but a further analysis using the syllolexical statistics criteria defined previously yields the minimum requirements for being considered cognate. Kapampangan pronouns are between 69% and 75% cognate with Sambal languages. Ivatan pronouns are only between 31% and 38% cognate with Sambal languages.

Of the 32 pronouns analyzed, there are six forms that are repeated among Sambal languages which are not cognate with the LWCs included in this study.

Table 5. Pronoun Cognate Percentages

	Bol	T.S.	Abel	B.S.	Ants	Indi	Amba	Kap	Iva	Ilo	Tag
Bolinao	100%	97%	88%	84%	81%	81%	84%	72%	31%	69%	28%
Tina Sambal	97%	100%	91%	88%	84%	84%	88%	72%	34%	69%	31%
Ayta Abellen	88%	91%	100%	91%	94%	94%	91%	75%	34%	66%	31%
Botolan Sambal	84%	88%	91%	100%	97%	91%	84%	69%	38%	63%	34%
Ayta MagAntsi	81%	84%	94%	97%	100%	94%	88%	72%	38%	59%	34%
Ayta MagIndi	81%	84%	94%	91%	94%	100%	91%	72%	34%	59%	31%
Ayta Ambala	84%	88%	91%	84%	88%	91%	100%	75%	34%	66%	31%
Kapampangan	72%	72%	75%	69%	72%	72%	75%	100%	31%	59%	28%
Ivatan	31%	34%	34%	38%	38%	34%	34%	31%	100%	22%	31%
Ilokano	69%	69%	66%	63%	59%	59%	66%	59%	22%	100%	25%
Tagalog	28%	31%	31%	34%	34%	31%	31%	28%	31%	25%	100%

5.5 Particles

The eight Sambalic particles studied are very similar; in general exhibiting only minor pronunciation differences. Once again the Ayta languages have high cognate percentages with each other while Bolinao and Tina Sambal have several mutual particles which are not cognate with the Ayta languages.

The one exception is the particle denoting surprise on account of the speaker (*pala* in Tagalog). Each Sambal language has unique forms for this particle.

Kapampangan has more particle cognates with Tagalog (50%) than with the Sambal languages (< 40%). Ivatan also scores less than 40% cognate with the Sambal languages.

Table 6. Particle Cognate Percentages

	Bol	T.S.	Abel	B.S.	Ants	Indi	Amba	Kap	Iva	Ilo	Tag
Bolinao	100%	63%	50%	50%	38%	38%	63%	25%	38%	13%	25%
Tina Sambal	63%	100%	38%	38%	25%	25%	38%	38%	25%	25%	13%
Ayta Abellen	50%	38%	100%	88%	75%	75%	75%	25%	25%	25%	13%
Botolan Sambal	50%	38%	88%	100%	75%	75%	75%	25%	25%	25%	13%
Ayta MagAntsi	38%	25%	75%	75%	100%	75%	63%	25%	25%	25%	13%
Ayta MagIndi	38%	25%	75%	75%	75%	100%	63%	38%	25%	25%	25%
Ayta Ambala	63%	38%	75%	75%	63%	63%	100%	25%	38%	25%	25%
Kapampangan	25%	38%	25%	25%	25%	38%	25%	100%	38%	25%	50%
Ivatan	38%	25%	25%	25%	25%	25%	38%	38%	100%	25%	38%
Ilokano	13%	25%	25%	25%	25%	25%	25%	25%	25%	100%	25%
Tagalog	25%	13%	13%	13%	13%	25%	25%	50%	38%	25%	100%

5.6 Conjunctions

In analyzing the conjunction systems of these languages, it is easily observed that the most basic conjunctions are virtually uniform among the Ayta languages. Several Ayta languages (Abellen, MagAntsi, and MagIndi) have identical conjunction sets for the seven conjunctions researched. In the three cases where there is a split distribution among Sambal languages, it is accounted for by Tina Sambal and Bolinao having common lexical entries that are different from Botolan Sambal and the Ayta languages.

The LWCs Ilokano, Kapampangan, and Tagalog have had little influence on the current use of connectors in Ayta and the other Sambalic languages. Among the LWCs, Kapampangan conjunctions are 29% or less cognate with Sambal languages. Other than *ta* 'because', Ivatan has no conjunctions in common with any of the Sambal languages.

Table 7. Conjunction Cognate Percentages

	Bol	T.S.	Abel	B.S.	Ants	Indi	Amba	Kap	Iva	Ilo	Tag
Bolinao	100%	57%	29%	29%	29%	29%	29%	14%	0%	14%	14%
Tina Sambal	57%	100%	57%	57%	57%	57%	43%	29%	14%	29%	29%
Ayta Abellen	29%	57%	100%	86%	100%	100%	86%	29%	14%	29%	29%
Botolan Sambal	29%	57%	86%	100%	86%	86%	71%	29%	14%	29%	43%
Ayta MagAntsi	29%	57%	100%	86%	100%	100%	86%	29%	14%	29%	29%
Ayta MagIndi	29%	57%	100%	86%	100%	100%	86%	29%	14%	29%	29%
Ayta Ambala	29%	43%	86%	71%	86%	86%	100%	29%	0%	14%	29%
Kapampangan	14%	29%	29%	29%	29%	29%	29%	100%	0%	14%	14%
Ivatan	0%	14%	14%	14%	14%	14%	0%	0%	100%	14%	0%
Ilokano	14%	29%	29%	29%	29%	29%	14%	14%	14%	100%	0%
Tagalog	14%	29%	29%	43%	29%	29%	29%	14%	0%	0%	100%

5.7 Interrogatives

Unlike other grammatical categories, the six interrogatives studied show very little similarity between Sambalic languages. With the exception of Botolan Sambal's similarity with MagAntsi (67%), Bolinao (67%) and Tina Sambal (83%), there are no combinations over 50% cognate. Even Ayta language interrogative systems compared with each other exhibit very limited similarity.

Kapampangan's interrogatives are 50% cognate with Tagalog and Ivatan but not very cognate with Sambal languages (33% or less). Ivatan, on the other hand, is surprisingly 50% cognate with both Botolan Sambal and Ayta MagAntsi in addition to Kapampangan.

Table 8. Interrogative Cognate Percentages

	Bol	T.S.	Abel	B.S.	Ants	Indi	Amba	Kap	Iva	Ilo	Tag
Bolinao	100%	67%	17%	67%	50%	17%	17%	17%	33%	33%	33%
Tina Sambal	67%	100%	33%	83%	50%	17%	17%	17%	33%	50%	33%
Ayta Abellen	17%	33%	100%	33%	17%	17%	17%	0%	0%	33%	0%
Botolan Sambal	67%	83%	33%	100%	67%	33%	33%	33%	50%	50%	33%
Ayta MagAntsi	50%	50%	17%	67%	100%	50%	33%	33%	50%	33%	33%
Ayta MagIndi	17%	17%	17%	33%	50%	100%	33%	17%	17%	17%	0%
Ayta Ambala	17%	17%	17%	33%	33%	33%	100%	17%	17%	17%	17%
Kapampangan	17%	17%	0%	33%	33%	17%	17%	100%	50%	17%	50%
Ivatan	33%	33%	0%	50%	50%	17%	17%	50%	100%	17%	50%
Ilokano	33%	50%	33%	50%	33%	17%	17%	17%	17%	100%	17%
Tagalog	33%	33%	0%	33%	33%	0%	17%	50%	50%	17%	100%

5.8 Composite Cognate Percentages

Below are the total composite cognate percentages for all 266 lexical items. There is a distinct gap between Sambal languages and non Sambal languages. The 7 Sambal languages are all 60% or higher in lexical similarity with each other. The Ayta languages are all 70% or more lexically similar with each other. Botolan Sambal is 75% or more lexically similar with all the Ayta languages.

Kapampangan is the most similar of the 3 LWCs with Sambal languages, ranging between 45% and 52% in lexical similarity with Sambalic languages. Ilokano and Tagalog are between 37% and 43% cognate with Sambalic languages. Ivatan is less than 27% cognate with Sambalic languages.

Table 9. Composite Cognate Percentages

	Bol	T.S.	Abel	B.S.	Ants	Indi	Amba	Kap	Iva	Ilo	Tag
Bolinao	100%	76%	62%	66%	60%	60%	61%	47%	24%	43%	38%
Tina Sambal	76%	100%	64%	70%	61%	60%	62%	48%	25%	41%	38%
Ayta Abellen	62%	64%	100%	79%	75%	70%	70%	45%	24%	41%	37%
Botolan Sambal	66%	70%	79%	100%	80%	75%	76%	47%	26%	41%	40%
Ayta MagAntsi	60%	61%	75%	80%	100%	83%	73%	49%	25%	37%	38%
Ayta MagIndi	60%	60%	70%	75%	83%	100%	73%	52%	25%	37%	37%
Ayta Ambala	61%	62%	70%	76%	73%	73%	100%	46%	25%	39%	41%
Kapampangan	47%	48%	45%	47%	49%	52%	46%	100%	22%	31%	42%
Ivatan	24%	25%	24%	26%	25%	25%	25%	22%	100%	24%	23%
Ilokano	43%	41%	41%	41%	37%	37%	39%	31%	24%	100%	31%
Tagalog	38%	38%	37%	40%	38%	37%	41%	42%	23%	31%	100%

6. Discussion and conclusions

6.1. Sambalic languages

In trying to postulate lexical similarity subgroupings within the Sambalic language family, several observations are intuitive. Ayta Abellen and Botolan Sambal appear to belong in the same grouping. Abellen has the highest cognate scores with Botolan Sambal (79%) and Botolan Sambal has its 2nd highest score with Abellen. Their geographic proximity make this conclusion almost predictable.

The northern coastal languages of Tina Sambal and Bolinao also form a group, being 76% cognate with each other. Since these languages are lowland, coastal languages and geographically close, this result also is not surprising.

The southern Ayta languages of MagIndi and MagAntsi also form a group, having the highest composite cognate scores of any two languages in this study (83%).

The mystery language in this study is Ayta Ambala. It has the highest cognate percentage with Botolan Sambal (76%) but that is not very remarkable since that is almost the lowest cognate percentage Botolan Sambal has with any Ayta language. So it is hard to know whether Ayta Ambala belongs with Ayta Abellen and Botolan Sambal in the north or with Ayta MagIndi and Ayta MagAntsi to the east. Another possibility is that emerging research on the Ayta Magbukun language will reveal a possible subgrouping with that language.

Botolan Sambal has high cognate percentages with all the Ayta languages and also fairly high similarity with the coastal lowland languages. This makes one wonder if it was either the original parent of all these other languages or maybe that a form of it is currently emerging as a trade language for Sambal speakers talking to someone from another group. Historical reconstruction is needed to answer these types of questions.

6.2. Sambalic Languages compared with others

In terms of the LWCs, Kapampangan is the most lexically similar with Sambalic languages, however it is still at least 10% less similar than the Sambalic languages are

with each other. Kapampangan also scores in the same range of lexical similarity with Tagalog as with Sambalic languages. While the gap in lexical similarity percentages between Kapampangan and Sambalic languages makes it difficult to conclude they belong to the same microgroup, Kapampangan's higher scores with Sambalic languages than other LWCs suggest macrogrouping.

As for Ivatan, while it is certainly premature to say that all Bashiic languages bear no link with the Central Luzon language family, it can be deduced that Ivatan either is fairly far removed from the Central Luzon language family or it has undergone much lexical change to reach the point where not even 30% of its lexicon is recognizable as cognates with any Sambalic language.

References

- Antworth, Evan. 1979. *A grammatical sketch of Botolan Sambal*. Manila: Linguistic Society of the Philippines.
- Blust, Robert. 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30(2), 73–106.
- Elgincolin, Sotero B. 1998. *English-Tina Sambal-Pilipino dictionary*. Manila: SIL.
- Enriquez, S.R. and M.O. Guzman. 1998. *English-Tagalog vocabulary*. Quezon City: Marran.
- Goschnick, Hella. 2005. Tina Sambal grammar notes. Ms.
- Green, Joanne and Roger Green. 1991. The grammar of Ayta MagIndi. Ms.
- Gordon, Raymond G., Jr. (ed.), 2005. *Ethnologue: Languages of the world*, Fifteenth edition. Dallas, Tex.: SIL International. Online version: <http://www.ethnologue.com/>
- Hidalgo, Cesar A. 1998. *Ivatan-Filipino-English dictionary: A Cultural Dictionary of Batanes*. Manila: Academics Foundation.
- Llamzon, Teodoro A. 1978. *Handbook of Philippine language groups*. UNESCO. Quezon City: Ateneo de Manila University Press.
- Mann, Noel. 2005. Syllostatistics. Ms.
- McFarland, Curtis Daniel. 1974. The dialects of the Bikol area. Ph.D. Dissertation, Yale University.
- Nitsch, Wilhelm. 1998. Some Ayta Abellen Grammar. Ms.
- Persons, Gary and Diane Persons. 1998. *Miirgu-irgu Atamo (Let's Dialogue): A Topical Phrasebook in Bolinao, Ilocano, Filipino, and English*. Manila: SIL Philippines.

- Ramos, Marciana Extra, Melchora Estoque, Meriel Padilla Quitco, and Hennie Lai Yee Chiu. 2005. A preliminary grammar sketch of Ayta Ambala. Ms.
- Rubino, Carl. Philippine language tree. [online] retrieved November 1, 2005, from <http://iloko.tripod.com/philtree.html>
- Rubino, Carl. 2000. *Ilocano dictionary and grammar*. Honolulu: University of Hawai'i Press.
- Rubrico, Jessie Grace U. The languages of the Philippines. [online] retrieved November 1, 2005, from http://www.languageinks.org/onlinepapers/fil_lang.html
- Stone, Roger. 2005. Ayta Abellen Ethnologue report. Ms.
- Storck, Kurt and Margaret Storck. 2005. *Ayta Mag-Antsi—English dictionary*. Manila: SIL Philippines.
- Turla, Ernesto C. *English to Kapampangan dictionary*. [online] retrieved 2005 from <http://www.maxpages.com/diccionario>
- Zorc, R. David. 1977. The Bisayan dialects of the Philippines: Subgrouping and reconstruction. *Pacific Linguistics*. C.44. 305–320.
- Zorc, R. David. 1982. Micro- and macro-subgrouping: criteria, problems, and procedures. In Carle et al, (eds.), *Gava', Studies in Austronesian Languages and Cultures Dedicated to Hans Kähler*. Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, Band 17:305–320. Berlin: Dietrich Reimer Verlag.

Appendix 1 – Wordlist Data

Table 10. Nouns (129 Total)

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
afternoon	mahilim	hapon	ma?apon	yabina	mahilim	mallim	ka?awat	makoyab	gatpanapon	hapon	apon
all	kagana?an	la?at	naubuh	sarba	kaganawan	amin	kaganawan	atavo	sabla?	lahat	halban
ankle	tikid	hakuñan	bukluñ	ñalwanalwan	bukoñ?bukoñ	mukod	buwi	kasindan	bukoñ?bukoñ	bukoñ?bukoñ	balogo?balogo?
ash	tubog	abo	abo	abonsyan	abo	dapo	abu	avo	abu	abo	aboh
back	gulot	bokot	bokot	gorot	bokot	likod	bukut	dichod	gulut	likod	bokot
banana	ha?a	sa?a	ha?a	batag	ha?a	saba	sa?a	vinyevéh	sagin	sagin	batag
belly	bitoka	bitoka?	bitoka	tyan	tiyan	tiyan	bituka	vodik	atian	ti?an	tyan
bird	manokmanok	manokmanok	manok	manokmanok	manokmanok	billit	manuk	manomanok	ayup	ibon	manok?manok
blanket	owih	owis	uwih	olis	owih	olis	uwis	ayob	ulas	kumot	ili
blood	daya?	daya?	daya	daya?	daya?	dara	daya	raya	daya?	dugo	daya?
body	laman	nawini	lawini	lalaman	nawini	bagi	lawini	karakohan	katawan	katawan	lalaman
bone	bot?o	bot?o	but?o	bot?ol	but?o	tulan	but?u	to?han	butul	buto	botol
brain	utik	otik	utik	otik	itik	utik	utik	otik	utak	utak	otok
breakfast	pamigat	almohal	mamoah	aramosal	almohal	pamigat	almusal	riagin	panaba	almusal	armosal
breast	huho	soso?	hoho	soso?	nuno?	suso	suso	soso	susu	suso	soso
butterfly	gopa?pa?	palo?	palo?palo	kaibañban	palo?palo?	kolibañban	palu	kodibañban	taluban	paruparo	kalibañban
carabao	damwag	damolag	damwag	damolag	damowag	nuañ	dulig	pagad	damulag	kalabaw	damolag
cheek	piñhipiñih	piñih	piñi	pispis	piñih	piñipig	piñi	pisnyi	pisñi	pisñi	piñpiñ
chest	pagaw	dibdib	pagaw	kirip	nibnib	barokoñ	pagaw	kalañan	salu	dibdib	kolop
child	anak	anak	anak	anak	anak	ubbiñ	anak	motdih	anak	bata	anak

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
chicken	manok	manok	manok	manok	manok	manok	manok	manok	manok	manok	manok
chin	baba	baba?	baba?	ba?ba?	baba?	timid	baba	tomid	baba?	baba	ba?ba?
cloud	uwip	olap	ginim	gonim	li?im	ulip	lem	dimdim	biga?	ulap	lolom
coconut	oŋot	oŋot	oŋot	oŋot	oŋot	inyog	uŋut	nyoy	uŋut	niyog	oŋot
crow	owak	owak	owak	owak	owak	owak	uwak	owak	wak	uwak	owak
day	maŋa?amot	allo	allo	awro	allo	aldaw	ma?amot	araw	aldo	araw	awlo
deaf	ti?i?	tolokon	tikkin	tirik	ti?ik	umil	tikin	makotiŋ	maklak	biŋi	tolok
debt	utaŋ	otaŋ	utaŋ	otaŋ	otaŋ	utaŋ	utaŋ	gatos	utaŋ	utaŋ	otaŋ
deer	oyha	oyha	oyha	oyha	oyha	ogsa	uwisa	agsah	usa	usa	oyha
dog	aho	aho	aho	aso	aho	aso	asu	tsito	asu	aso	aso
dream	taynip	taynip	taynip	taynip	taynip	tagtagainip	taynip	tayanip	tinap	panaginip	taynap
dust	tu?apok	gabot	tuwapok	ta?pok	alikaok	tapok	alikaok	ahbik	alikaok	alabok	gabok
ear	talina?	tik	ti:k	totolyan	tolih	lapayag	tik	tadinya	balugbug	ta?iŋa	totolyan
earthquake	lay?on	dayon	layon	rayon	layon	ginginid	layun	nini	ayun	lindol	layon
egg	talay	okoy	poga	salay	okoy	itlog	ukuy	otioy	ebun	itlog	oybon
eggplant	taloŋ	taloŋ	balansenas	tarom	talom	taron	balansenas	vahosa	balansenas	taloŋ	talom
elbow	hiko	hiko	hiko	siko	hiko	siko	siko	sicho	siko	siko	hiko
eye	mata	mata	mata	mata	mata	mata	mata	mata	mata	mata	mata
face	lupa	lupa	lupa	ropa	lupa	rupa	lupa	daŋoy	lupa	mukha	lopa
father	bapa	bapa	tatay	tatay	tatay	tataŋ	bapa	ama	ibpa?	ama	ama?
feather	polokpok	baniban	habot	sabot	habot	dotdot	sabot	boboh	bulbul	balahibo	habot
feces	taka?	takya?	taka	takli?	taka?	tak?ki	takya	tatsi	takla?	ta?i	taka
finger	galamay	galamay	galamay	garamay	galamay	ramay	galamay	kakamay	taliri?	daliri	galawagaw

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
fire	apoy	apoy	apoy	apoy	apoy	apoy	apuy	apoy	api?	apoy	apoy
fish	ikan	kina?	kina?	kona?	kona	ikan	kina?	amoy	asan	isda?	ikan
fly	lanjw	lanjw	lanjw	lanjo	lanjw	ŋilaw	lanjw	nanid	lanjo	lanjw	lanjo
fruit	dawa	tagiy	tagiy	boŋa	boŋa	boŋa	tagiy	asi	boŋa	boŋa	boŋa
frog	pahina?	palanŋka?	togak	pasaj	pahina?	tukak	pasinga?	palakak	tugak	palaka?	talakba?
hair	labok	labok	habot	sabot	habot	bu?ok	sabot	vo:k	buak	buhok	habot
hand	gamit	gamit	gamit	gamit	gimit	ima	gamit	rapan	gamat	kamay	gamot
head	o:	olo	ulo	olo	olo	ulo	ulu	oho	buntuk	ulo	olo
heart	puho	poho	biki?	poso?	puso?	puso	pusu	tawol	pusu?	puso	poso
heel	bu?ih	bowi?	bo?i?	bu?id	bo?eh	mukod	buwi	tovin	sakuj	sakonj	bo?or
house	baiy	bali	bali	bali	bali	balay	bali	vahay	bale	bahay	bali?
husband	ahawa	ahawa	ahawa	asawa	ahawa	asawa	asawa	kakovot	asawa	asawa	ka?amba
knee	tu?ol	tagul	to?ol	to?od	to?ol	tumiŋ	tu?ul	to:d	tud	tuhod	to?or
leftside	odi	kaliwa?	uki	wiri	oki	kanigid	odi	do ori na	kaili	kaliwa	wili?
lightening	kilat	kimat	kimat	kimat	kimat	kimat	kilat	tsidat	kildap	kidlat	kimat
lip	labi	labi?	labi	labi?	labi?	bibig	labi	vivi	labi?	labi	labi
liver	agtay	agtay	agtay	agtay	agtay	dalim	agtay	atay	ate	atay	agtay
lungs	baya?	baya?	baya	baya?	baga?	bara	baya?	apwaw	baga?	baga?	baya?
man	laki	liyaki?	lalaki?	lalaki	lalaki	lalaki	lilaki	mahakay	lalaki	lalaki	lalaki
mat	amak	amak	amak	apay	amak	ikamin	dasay	apin	dase	banig	apay
monkey	baki?	bakulaw	baki?	baki?	baki?	bakis	bakulaw	tsorjo	bakis	ungoy	bako?
moon	buwan	boyan	buwan	bulan	bowan	bulan	buwan	vohan	bulan	buwan	bulan
morning	mahanib	mahambak	hanib	boklas	mahambak	bigat	maranun	mavikas	abak	umaga	boklah
mother	indo	indo?	indo	ina	nanay	ina	indu	ina	indu?	ina	indo?

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
mosquito	ilik	yamok	ilik	ilik	ilik	lamok	ilik	tamonij	yamuk	lamok	ilok
mouth	bibiy	bibiy	bibiy	bibiy	bibiy	niwat	bibiy	nos	asbuk	bibig	boboy
mud	pita?	pita?	pita	pita?	pita?	pitak	burak	hotah	burak	putik	dotak
name	nanan	nanan	nanan	nanan	nanan	nagan	lagyu	nanan	lagyu?	nanan	nanan
neck	li?iy	biŋ	li?iy	li?iy	li?iy	tiŋid	biŋ	lagaw	batal	li?ig	lo?oy
needle	kadayim	kadayim	karayim	kadayim	karayim	dagom	karayom	dayim	karayum	karayom	karayom
night	madiglim	yabi	yabi	yabi	yabi	rabi?i	yabi	ahip	benji	gabi	yabi?
noon	ugto	ogto	ugto	ogto	ogto	aldaw	ugtu	payarawmin	ugtu	taphali	kaogtowan
nose	balonoh	balonoh	balunoh	balonoh	balonoh	agon	balunus	momodan	aruŋ	ilon	a?lon
palm	daokap	palad	dalipapa	dalokap	dawokap	dakolap	daiipapa?	rapan	palad	palma	palal
peanut	mani?	mani?	mani	mani?	mani?	mani?	mani	mani	mani?	mani?	mani?
pig	baboy	baboy	baboy	babuy	baboy	baboy	babuy	bago	babu?	baboy	baboy
pillow	onan	ponan	ulunan	alonan	onan	pujan	ulunan	hanpan	ulununan	unan	alonan
rain	udan	abagat	uran	rapig	oran	tudo	uran	timoy	uran	ulan	abagat
rainbow	bowakaw	bolalapanaw	buwanapanaw	kabonlalakaw	kabullalapanaw	balalayaw	buwanapanaw	ranyiran	pinapanari	bahaghari	kabonlalapaw
rat	dagih	baki?	dagih	ibot	baki?	ba	dagis	karam	dagis	daga	bolaki
rib	tagyan	tagyan	tagyan	tagyan	tagyan	paragpag	tagyan	taglan	tadyan	tudyuhin	tagyan
rice ¹	biyah	biyah	biyah	byas	boya	bagas	biyas	paray	abias	bigas	boyah
river	kabaturan	yawog	ilog	ilog	bala	karayan	ilog	ahson	ilog	ilog	kabatwan
root	yamot	yamot	wakat	yamot	yamot	ramot	uyat	yamot	yamut	ugat	yamot
roof	atip	atip	atip	atip	bubon	atip	bubon	atip	bubunan	bubon	atop

¹ husked

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
rope	lubil	yobil	yubil	robir	yobil	tali	yubil	pinospos	lubid	lubid	yobil
salt	ahin	ahin	ahin	asin	ahin	asin	asin	asin	asin	asin	asin
sand	lanhi?	bohanjin	balah	boyanjin	kapati?an	darat	balas	anay	balas	buhanjin	boyanjin
shadow	anino	anino	anino	alina	anino	anniwawan	aninu	anino	alino	anino	anino
shoulder	bantaw	bantaw	bantaw	abaya	abaya	abaga	bantaw	pakoh	pago	balikat	abaya
sibling	katojno	patil	patil	busat	patil	kabsat	patil	kakthih	kapatad	kapatid	talakaka
skin	luti?	balat	balat	katat	katat	kudil	balat	kodit	balat	balat	katat
sky	lanjit	lanjit	lanjit	lanjit	lanjit	lanjit	lanjit	tohos	banua	lanjit	lanjit
smoke	anoh	ahok	ahuk	asok	ahok	asok	asuk	ahob	asuk	usok	asok
snake	bikat	otan	utan	olay	utan	ulig	utan	boday	ubinyan	ahas	ulay
soup	habaw	habi?	habaw	sabaw	habaw	sopas	sabi?	asoy	sabo	sabaw	habaw
spider	giganj	pidaj?	giganj	giganj	giganj	lawwalawwa	giganj	hahawa	babaguar?	gagamba	giganj
spit	luda?	topay	dula	loda?	dula?	topra	tupay	tipa	lura?	lura	lora?
star	bitu?in	bitiwin	bitiwin	bitu?in	bitu?in	bitu?in	bitiwin	vitohin	batuin	bitu?in	bitu?on
stone	dapah	bato	bato	bato	bato	bato	bato	bato	bato	bato	bato
sweat	haynit	haynit	haynit	saynit	haynit	lipit	saynit	inalindin	pawas	pawis	haynot
potato ²	kamoti	kamoti	kamotsi	kamoti	kamoti	kamoti	kamuti	wakay	kamuti	kamoti	kamoti
tail	ikoy	ikoy	ikoy	ikoy	ikoy	ipos	ikuy	ipos	iki?	buntot	ikoy
tear ³	lowa?	lowa?	luwa	luwa?	luwa?	luwa	luwa	ho:	lua?	luha	lowa?
teeth	nipin	nipin	gogot	nipin	nipin	nipin	lasi	nyipin	ipan	nipin	nipon

² sweet

³ from crying

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
termite	anag	anag	anag	anag	anag	anag	anag	anag	ane	anag	anag
thigh	paʔa	paʔa	paʔa	paʔa	paʔa	lupʔpo	paʔa	pa:	puad	pigi	paʔa
thorn	diwih	dyiwi	diwi	duwi	dowi	siʔit	diwi	tolok	suksuk	timik	dowi
throat	bikʔlaw	biklaw	halo	bokraw	boklaw	karabukob	salu	titihnan	akmulan	lalamunan	boklaw
thunder	kudol	kimat	kilat	kodor	korol	gurrud	kilat	adiy	duldul	kulog	korol
toe	galamay	galamay	galamay	garamay	galamay	ramay	galamay	kakamay	taliriʔ	daliri	galawagaw
tongue	dila	dila	dila	dilaʔ	dilaʔ	dila	dila	rida	dilaʔ	dilaʔ	dila
tree	kayo	kayo	kayo	kayo	kayo	kayo	kayu	kayo	pun kayuʔ	kahoy	kayo
trousers	pantalón	halwal	hanwal	botarga	halwal	pantalón	sanwal	salavini	salol	pantalón	pantalón
turtle	pagʔonj	pagʔonj	pagʔonj	pagonj	pagʔonj	pagʔonj	pagʔonj	iranj	pauʔ	pagonj	pagʔonj
urine	duday	dyodyay	duray	omiʔ	doray	isbo	duray	pitig	imiʔ	ihi	omiʔ
vein ⁴	oyat	oyat	iyat	oyat	iyat	urat	uyat	oyat	uyat	ugat	oyat
wall ⁵	liqlin	dinqin	lolog	riqrin	liqlin	dinqin	liqlin	gadagada	dalig	pader	padir
water	lanom	lanom	lanim	ranom	lanom	danom	lanim	danom	danum	tubig	lanom
waterfall	halughog	bihay	bihay	bosay	bihay	burayok	bisay	komayasakas	puntuʔ	talon	bosay
wife	ahawa	ahawa	ahawa	asawa	ahawa	asawa	asawa	kakovot	asawa	asawa	kaʔambalin
wing	pakpak	palakpak	puwal	pakpak	pakpak	payak	palakpak	panyid	pakpak	pakpak	pakpak
winnow	bilawo	tatap	bitsay	tatap	bilawo	bigao	igo	bilaʔut	tatap	bilaʔo	tatap
woman	babayi	babayi	babayi	babayi	babayi	babai	babayi	mavakis	babai	babaʔi	babayi

⁴ blood

⁵ of a house

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
worm	ulil	bolati	uwil	bolati	uwil	igʔgis	uwil	ohid	bulati	uod	olol
year	taʔon	taʔon	taʔon	taʔon	taʔon	tawin	taʔun	awan	banua	taʔon	taʔon

Table 11. Verbs (48 Total)

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
answer	hibat	kitab	kibat	ubat	tubay	sunbat	kitab	atbay	pakibat	sagot	obat
bathe	paliʔu	loyoʔ	palyu	maluyoʔ	liyo	digos	luyu	ryus	diluʔ	ligo	lyoʔ
bite	kayat	kayat	kayat	kayat	kayat	kayat	kayat	sonyit	ket	kaqat	kayat
burn	ulam	olam	ulam	iwik	puʔol	uram	ulam	susuhan	silaban	sunog	poʔol
bury	ilbinj	ilbinj	ilbinj	ilbinj	tabon	tabon	libinj	vuvun	ikutkut	baʔon	tabon
buy	haliw	haliw	haliw	haliw	gataj	gataj	haliw	manadiw	salif	pabili	haliw
call	inat	baʔig	inat	bawʔgin	inat	ayab	big	tawag	aus	tawag	haʔway
choose	piliʔ	piliʔ	piliʔ	piliʔ	pili	pili	piliʔ	pidi	piliʔ	piliʔ	piliʔ
climb	manik	dakriy	unik	oliʔ	molif	uli	dakey	kayab	ukiak	akyat	oliʔ
count	bilan	bilan	bilan	bilan	bilan	bilan	bilan	vida	bilan	bilan	bilan
cough	kukoʔ	kokoʔ	koko	koʔkoʔ	kukoʔ	uyik	kuko	gugu	kukuʔ	ubo	koʔkoʔ
cry	tanjih	tanjih	tanjih	akis	tanjih	sapit	tapis	tomanyis	tanjis	iyak	tanjih
defecate	takaʔ	takyaʔ	takaʔ	takif	takaʔ	takʔki	takyaʔ	tatsi	taklaʔ	taʔi	takaʔ
dig	kali	kalih	kali	kotkot	kali	kali	kali	kadi	kulkul	hukay	kotkot
drink	inom	inim	inim	inom	inom	inom	inum	inum	inum	inom	inom
eat	kan	kan	ikan	kan	kan	manan	ikan	kan	kan	kaʔin	kan
fly	lompad	lompad	lompad	lompad	lompad	tayab	lompad	sayap	sulagpo	lipad	lompar

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
forget	liwa	liwan	liwa	liqwan	liwa	lipat	liwa	wayak	kaliqwan	limot	liqwan
give	biay	am?i?	bi	bi	mam?ih	itid	bi	turuh	bie	bigay	bi
hear	lini?	lini?	lini?	iqar	lingi?	dinqig	lini?	adnjy	damdam	kinig	loqo?
hide	tayo	tago?	tago	adi	tago?	limming	tagu?	tayu	salikut	tago	ari?
kill	patiy	pati	patsi	pati	pati	patay	pati	diman	pate	patay	pati
laugh	kili?	kayli?	kaili?	ka?lis	ka?ili	katawa	kayli?	ayak	kaili	tawa	omlih
lie ⁶	boqkok	boqkok	boqkok	bula?	boqkok	ulbod	lawig	dada?ay	laram	kasinuqalinan	kalotoy
pay	bayad	bayad	bayad	bayad	bayad	bayad	bayad	paga	bayad	bayad	bayar
plant	tanim	tanim	tanim	tanim	tanim	mula	tanim	muha	tanam	tanim	mola
pull	goloy	goloy	guloy	goroy	goloy	guyod	binunay	palan	igut	batak	bira
play	dagaw	daragaw	dagaw	ikap	dagaw	ay?ayam	kapal	yayam	aluq	laro?	ikap
push	tudon	doron	dulin	toror	tulak	duron	turon	pasosoy	tulak	tulak	tolak
repeat	oman	oman	oman	uman	uman	ulit	uman	pirwah	uman	ulit	olit
run	mayiw	lokho?	uwayu	palayo	powayo	taray	uwayu	yayu	pulayi	takbo	olayo?
see	kit	akit	ikit	kit	kit	kita	akit	vuya	akit	kita	kit
sew	tahi?	tayi?	taynib	tayi?	tayi?	da?it	tayi?	kapanhimib	tayi?	tahi?	tayi?
sit	ikno?	ikno?	ikno?	tikri?	ikno?	tugaw	iknu?	disna	lukluk	upo	toklo?
sleep	toloy	toloy	tuloy	ilik	tuloy	turog	biw	itsih	tudtud	tulog	ma?lok
smell	da?ip	da?ip	da?ip	aqot	da?ip	aqut	dip	aqot	bau	amoy	aqot
squeeze	pihpih	pispis	pihpih	pispis	pihpih	pispis	pispis	pitos	paslan	siksik	poloh
stand	idij	idij	irij	idij	irij	takder	irij	atnik	tikdo	tayo	odog

⁶ false

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
steal	takaw	takaw	takaw	takaw	takaw	takaw	takaw	takaw	tako	takaw	takaw
suck	hiphip	sipsip	hiphip	sipsip	hiphip	susop	sipsip	somosoh	sipsip	sipsip	hophop
swim	tañoy	lanjy	kaway	tañoy	tañoy	lanjy	kaway	yawat	kawε	lanjy	tañoy
throw	tapon	halo?	tapon	bantak	tapon	bato	tapun	pagsid	ugε	tapon	bantak
vomit	hoka?	hoka	huka	soka	hoka	sarwa	suka	mutawta	suka	suka	hoka
walk	kodanj	alyako?	bita	lalako	owako	magna	bita	mayam	lakad	lakad	lalako?
wait	aqad	injan	igjan	ta?gan	anti?	uray	itiñ	nanaya	aya	hintay	ta?gan
wash	oyah	laba	uyah	uyas	oyah	buggo	huyas	uyas	uas	hugas	oyah
winnow	tatap	tatap	tatap	tatap	tatap	ta?ip	tatap	wakwak	tatap	tahip	tatap
work	obda	trabaho	trabaho	obra	trabaho	trabaho	obra	trabahu	obra	trabaho	trabaho

Table 12. Adjectives (38 Total)

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
big	laki	hmak	hilay	?alaki	hilay	dak?hl	ragul	rakoh	dagul	laki	hi?ban
bitter	pa?it	pa?it	pa?it	pa?it	pa?it	pa?it	pa?it	akpad	payit	pa?it	pa?it
black	ñitit	w/lin	ñitsit	ñisit	ñitit	ñisit	uayan	vahuj	tulin	itim	ñitit
cold	lay?ip	layip	lay?ip	ray?ip	lay?ip	lami?is	lip?it	hanibnib	dimla	lamig	layop
deep	lah?	lah?	lah?	?ararim	lah?	unig	lah?	rahim	lalam	lalam	lalo
difficult	?idap	?idap	?irap	?idap	?irap	riqat	sakit	sadit	sakit	hirap	?irap
dirty	dinat	yanat	rimik	nikasama?	yanat	rugit	dinat	rudit	dinat	rumi	yanat
dull	mo?mo	purul	porol	kamil	purul	mudil	purul	ngarib	purul	purul	porol
far	ta?añ	dayo?	rayu?	?adayo?	rayo?	adayo	dayu	rayi	dayu?	layo	rayo?
fat	taba?	taba?	taba?	taba?	taba?	lukmig	taba?	tava	taba?	taba?	taba?

fragrant	baṅoh	baṅo?	baṅoh	baṅlu	baṅlo	baṅi?	asdip	baṅlu	baṅo	baṅloh
good	ṅid	hampat	hampat	bista	imbag	ṅid	pia	?ap	buti	labah
heavy	biyat	biyat	byat	byat	dagsin	biyat	rarahmit	bayat	bigat	byat
hot	?amot	?umot	?omot	?mot	puḍot	?umut	kuhat	pali?	?imit	a?mot
hunger	bitil	bitil	lonoh	bitil	bisin	layan	aptiṅ	danup	gutom	bitil
long	kadaṅ	kadaṅ	karaṅ	?andro	atid?og	karaṅ	anaru	kaba?	haba	anloh
many	laboṅ	dila?	l?at	?abaw	adu	laki	aru	dakal	rami	lako?
narrow	hapiy	kayik	kapit	?akipot	akikid	kapit	idid	kiput	kipot	kipot
near	haliy	dani?	rani	?adani	asidig	dani	asṅin	lapit	lapit	rani?
new	bayo	bayo	bayo	ba?yu	baro	bayu	vayu	bayu	bago	bayo
red	tibya	odit	orit	o?dit	labaga	udit	vaya	lutu?	pula	tibya?
sharp	tadim	tadim	tarim	tadim	tadim	tarim	tarim	taram	talim	talas
skinny	biṅ	biṅ	toklod	biṅ	kuttoṅ	ibiṅ	golaṅ	payat	payat	a?boṅ
small	kandi	kayik	nanawa	da?ikliṅ	bassit	biṅ	didikiy	lati?	li?it	kalog
smooth	lino?	kinis	ṅolinṅinṅ	dalonot	lamuyot	siliṅ	whas	kinis	kinis	yamo?
sour	ahim	asim	ṅiho	alsim	alsim	aslam	napa	aslam	?asim	?alhom
straight	toynoṅ	toynoṅ	toynoṅ	to?rir	lintig	tinek	talinṅ	tulid	tuwid	toynoṅ
strong	hikaw	lakas	hikaw	ksaw	piḡsa	lakas	mayit	sikan	lakas	kohaw
sweet	lam?ih	tam?ih	ayaṅ	sam?it	sam?it	tam?is	unawnas	yumu	tamis	hamit
thick	kuḡpa	kodpa	kuḡpa	kobpal	puskol	kudpa	tukpuh	kapal	kapal	korpah
thin	inpih	nipis	impih	impis	inpih	impis	taripis	inpih	nipis	impih
weak	kapiy	ina?	ayna	kapoy	kapoy	ina	kaha	ina?	hina	kapoy
wet	baha?	baha?	baha	albit	basa	basa?	vasa	basa?	basa	to?pakon
white	puti?	poti?	putsi	poti?	puraw	puti	idak	puti?	puti?	poti?
wide	way	lawag	ak?wal	?alawaṅ	lawa	kuwal	wvuṅ	lapad	lapad	lawañ
yellow	holyaw	dilaw	hilyaw	silyaw	ki?aw	?ilyaw	yuhama	dilo	dilaw	holyaw

Table 13. Pronouns (32 Total)

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
F1D	hikita	hakita	hikita	si?ta	hita	sita	sikita		ikata		hita
F1E	hikayi	hakami	hikay	si?kami	hikayi	sikami	sikay	yamin	ikami		hikami
F1I	hikitawo	hakitamu	hikitamo	si?tamo	hitamo	sitayo	sikitamo	yatin	tamu		hitamo
F1S	hiko	haku	hiko	si?ko	hiko	siak	siku	yakin	yaku		hiko
F2P	hikawo	hakawu	hikaw	si?kamo	hikawo	sikayo	sikau	inio	ikayu		hikawo
F2S	hika	haka	hika	si?ka	hika	sika	sika	imo	ika	ikaw	hika
F3P	hila	hila	hila	sira	hila	isuda	sila	sira	ila		hila
F3S	hiya	hiya	hiya	sya	hiya	isuna	siya	sia	iya		hiya
Foc1D	kita	kita	kita	ata	ta	ta	kita		kata		ta
Foc1E	kayi	kami	kay	kami	kayi	kami	kay	kami	kami	kami	kami
Foc1I	kitawo	tamu	kitamo	atamo	tamo	tayo	kitamu	ta	ikatamu	tayo	tamo
Foc1S	ako	aku	ako	ako	ako	ak	aku	ako	ku	ako	ako
Foc2P	kawo	kawu	kaw	kamo	kawo	kayo	kau	kamo	kayu	kayo	kamo
Foc2S	ka	ka	ka	ka	ka	ka	ka	ka	ka	ka	ka
Foc3P	hila	hila	hila	sara	hila	da	sila	sira	la	sila	hila
Foc3S	ya	ya	ya	ya	ya	na	ya	sia	ya	siya	ya
Pos1D	ta	ta	ta	ta	ta	ta	ta		ta		ta
Pos1E	mi	mi	na?in	mi	nawin	mi	yan	namin	mi	namin	mi
Pos1I	tawo	tamu	tamo	tamo	tamo	tayo	tamo	ta	tamu	natin	tamo
Pos1S	ko	ku	ko	ko	ko	ko	ku	ko	ku	ko	ko
Pos2P	yo	yu	yo	moyo	yo	yo	yu	mio	yu	ninyo	moyo
Pos2S	mo	mu	mo	mo	mo	mo	mu	mo	mu	mo	mo
Pos3P	la	la	la	ra	la	da	la	da	da	nila	la

Pos3S	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
Dir1D	kanta	kanta	kanta	kanta	kanta	kanta	kanta	kanta	kanta	kanta	kanta	kekata	kekata	kekata	kekata	kekata	kekata	kekata
Dir1E	kammi	kanyan	kanna?in	komi	konnawin	kaniami	kanyan	kantamu	diamin	kekami	amin	kekami	kekami	kekami	kekami	kekami	kekami	kekami
Dir1I	kantawo	kantamu	kantamo	kontamo	kontamo	kaniatayo	kantamu	diatin	kekati	kekati	atim	kekati	kekati	kekati	kekati	kekati	kekati	kekati
Dir1S	kaŋko	kaŋku	kaŋko	kaŋko	kaŋko	kaniak	kaŋku	kaŋku	kaŋku	kanaku	akin	kanaku	kanaku	kanaku	kanaku	kanaku	kanaku	kanaku
Dir2P	kanyo	kayu	kamoyo	komo	komo	kaniayo	kamuyu	dimio	kekayu	inyo	komoyo	kekayu	kekayu	kekayu	kekayu	kekayu	kekayu	kekayu
Dir2S	kammo	kamu	kamo	komo	komo	kaniam	kamu	dimo	keka	iyoy	komo	keka	keka	keka	keka	keka	keka	keka
Dir3P	kanla	kanla	kalla	konra	konla	kaniada	kalla	dira	karëla	kanila	konla	karëla	karëla	karëla	karëla	karëla	karëla	karëla
Dir3S	kana	kana	kana	kona	kona	kaniana	kana	dia	kaya	kaniya	kona	kaya	kaya	kaya	kaya	kaya	kaya	kaya

Table 14. Particles (8 Total)

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
surprise	manayti	alud	awod	anamain	awod	gayam	pata	sawon	pala	pala	payti
still/yet	po?	pun	pon	it	pon	pay	pun	pa	pa	pa	ot
now	ana	yna	ana	ana	ana		ana	na	na	na	ana
desire	dayi	hana	dayi	sapa	dayi	koma	dayi	sana	sa?	sana	komon
only	biŋat	dili?	tana	boŋatana	boŋat	la?in	kay	voyvoh	mu?	lang	boŋat
also	met	mit	it	anamaet	it	mit	it	pa	rin	din	anamaot
it is said	kano	kanu	kano?	kano	kano	kano	kanu	kono	kanu	daw	kano
again	oman	uman	oman	uman	uman	manin	uman	mirwa	nanaman	ulit	anaman

Table 15. Conjunctions (7 Total)

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
and	boy	boy	boy	tan	boy	kin	buy	kano	amponj	at	tan
because	ta	kasi?	ta	bana?	ta	ta	ta	ta	ulij	dahil	ta
but	noba	nowa?	no?a	bale?	piro	ɲim	nuwa?	amna	onej	pero	kot
if/when	no	nu	no	no	no	no	nu	an	nuj	kuj	no
or	o	o	o	ono	o	win?no	o	anmana	o	o	o
so that	ta?omon	ta?imin	ta?imin	pigaw	imin	tapno	imin	tapian	baŋ	upañ	pijaw
therefore	kaya?bay	kaya?	kabay	kanya?	kaya?	ɲarud	kaya	dawa	inia?	kaya	kanya?

Table 16. Interrogatives (6 Total)

	Abel	Amb	Anchi	Bol	BS	Ilo	Indi	Iva	Kap	Tag	TS
when?	makano	makano	makano	ka?no	makano	ka?ano	makano	an?maŋu	kapilan	ka?ilan	makano
where?	way?ihtiw	ha?anto?	anto	adi	ayri	ayan	antu	dinu	nu	nasa?an	ayti
who?	hinya	hita?	hino	si?no	hino	sinno	sisabit	sino	ninu	sino	hino
what?	aya	hita?	hino	ani	anya	anya	sabet	aŋu	nanu	ano	anya
how?	way?omon	pakapakun	paŋno	pa?no	paŋno	kasano	parasa?antu	maypaŋu	makananu	pa?ano	pa?no
why?	takit	owta?	awta	mayin	inta?	apay	awta	unta	inta?	bakit	anoŋkot

Appendix 2 – Data Acknowledgements

Ayta Abellen	Data files of Wilhelm Nitsch and Roger Stone, interview with Rodante Capiendo and Emilio Laurzano
Ayta Ambala	Data files of Mercy Ramos and Hennie Chiu, SIL survey files
Ayta MagAntsi	<i>Ayta MagAntsi Dictionary</i> (Stork and Stork, 2005), correspondence with Kurt and Margaret Storck
Ayta MagIndi	Data files of Roger and Joanne Green, SIL survey files, interview with numerous Ayta MagIndi in Nabuklod, Pampanga
Botolan Sambal	Data files of Char Houck, <i>Grammatical Sketch of Botolan Sambal</i> (Antworth, 1979), interview with Joelito De La Cruz
Bolinao	Correspondence with Gary Persons
Ilokano	<i>Ilokano Dictionary and Grammar</i> (Rubino, 2000), interview with Evangeline Stone and Mary Jane Mercado
Ivatan	<i>Ivatan-Filipino-English Dictionary</i> (Hidalgo, 1998) and correspondence with Ginny Larson
Kapampangan	Correspondence with Michael Raymon Manaloto Pangilinan, and <i>English to Kapampangan Dictionary</i> , (Turla)
Tagalog	<i>An English-Tagalog Vocabulary</i> (Enriquez and Guzman, 1998)
Tina Sambal	<i>English-Tina Sambal-Pilipino Dictionary</i> (Elgincolin, 1998) and correspondence with Hella Goschnick

Ergative Control of Syntactic Processes in Southern Sinama*

Douglas Trick
SIL International

This paper documents the high level of ergative control of five syntactic processes in Southern Sinama (southwestern Philippines). It begins with an explanation of ergativity in reference to morphology (with data from Southern Sinama which illustrates morphological ergativity). It then introduces and explains each of the following syntactic operations, demonstrating that they exhibit an ergative-absolutive pattern in Southern Sinama: relativization, clefting, WH-question formation, equi-NP deletion, and second-position cliticization. That is, it is the O argument of a transitive clause which controls these syntactic processes. This contrasts with most other Philippine languages in which control of these syntactic processes is distributed more or less evenly between the A argument and the O argument.

1. Introduction

In recent years, Philippine languages have been analyzed as being morphologically ergative.¹ Although most of the world's languages that display morphological ergativity also display a strong nominative pattern of syntactic control, this is not the case with Philippine languages. Available studies of syntactic processes in Philippine languages indicate that in transitive clauses, control is more or less evenly distributed between the two syntactically required arguments, exhibiting neither a dominant nominative pattern nor a dominant ergative pattern of syntactic control.² Southern Sinama,³ however, appears to be an exception to this general pattern for Philippine languages

* This paper was presented at the Tenth International Conference on Austronesian Linguistics (10-ICAL), January 17–20, 2006, at Puerto Princesa, Palawan, Philippines.

¹ For theoretical discussions, see Payne (1982), de Guzman (1988), Gerdts (1988), Kroeger (1993), Mithun (1994), Brainard (1994), and others. For ergative analyses of Philippine languages, see Walton (1986) for Sama Pangutaran, Hodder (1999) for Mayoyao Ifugao, Pebley and Brainard (1999) for Kagayanen, Gault (1999) for Sama Bangingi', and Brainard and Behrens (2002) for Yakan.

² See Schachter (1976, 1977) and Kroeger (1993) for Tagalog, and Brainard (1994, 1996) for Karao.

³ Southern Sinama (or Sama Southern) is spoken by the Sama people of the province of Tawi-Tawi in the Sulu archipelago of the southwest Philippines. It is estimated that there are about 100,000 speakers of Southern Sinama in Tawi-Tawi, and an additional 100,000 on the north and east coasts of Sabah, Malaysia.

There are several distinct Sama languages within the Sama-Bajaw, or Samalan, subgroup. Southern Sinama is most closely related to Central Sinama and Pangutaran Sinama. These and other Sama-Bajaw languages are related to the Barito languages of southeast Borneo

in that it displays a high degree of syntactic ergativity.⁴ In this paper, we will show that the majority of syntactic processes occurring in Southern Sinama are controlled exclusively by S, the single argument of an intransitive clause, and O, the more patient-like argument of a transitive clause.

The paper will begin with a general explanation of ergativity (in reference first to morphology and then to syntactic control), followed by a description of case marking morphology in Sinama. Five major syntactic processes will then be investigated, establishing that each of them has an exclusive ergative pattern of control.

2. Explanation of ergativity

Every language has ways of expressing states or events, some that involve only one argument (e.g., *She is sleeping*) and others that involve two or more arguments (e.g., *She helped me*). Generally, single-argument states and events are expressed by intransitive clauses; other states and events are expressed by transitive clauses. For this discussion, arguments are assumed to be NPs that bear a grammatical relation to the verb and thus are grammatical relations. Following Dixon (1979, 1994), these arguments are defined as:

- (1) 'S', the single argument of an intransitive clause;
- (2) 'A', the more agent-like argument of a transitive clause (in general, the one initiating the action);
- (3) and 'O', the more patient-like argument of a transitive clause (in general, the one affected by the action).

Thus, in the sentence *She is sleeping*, *she* is the S argument. In the sentence *She helped me*, *she* is the A argument, and *me* is the O argument.

Languages have ways of encoding these different arguments, or grammatical relations, typically using one or more of the following three formal devices:

- (1) word order (e.g., English; cf. *Ben helped Tim* and *Tim helped Ben*. In these basic transitive English clauses, the A argument precedes the verb and the O argument follows the verb.)

(Blust 2005). The term 'Sama' refers to the Sama people or to their language; 'Sinama' refers specifically to the language and will be used in the remainder of this paper.

Research in Sinama was carried out by the author under the auspices of the Summer Institute of Linguistics during the period of September 1987 to January 2006. Approximately four years of that time were spent resident in the village of Tubig Sallang, Bongao, Tawi-Tawi. About 100 texts of various genre were collected, paradigms were elicited, and a dictionary of some 3300 entries was compiled. These data are the basis for the results presented here.

I would like to express sincere appreciation to my primary Sama language research associates, Mr. Himpun Pallong (deceased) of Bongao and Mr. Nasaruddin Sambas of Simunul. I would also like to thank Dr. Sherri Brainard for her helpful comments on an earlier version of this paper.

⁴ Other research suggests that the Sama language family as a whole, including Sama Pangutaran (Walton 1986), Sama Bangingi' (Gault 1999), and Yakan (Brainard and Behrens 2002), exhibit a high degree of syntactic ergativity.

- (2) case marking (e.g., English; cf. *She helped me* and *I helped her*. English has one set of pronouns that refers to the A argument (often called subject pronouns) and a separate set that refers to the O argument (often called object pronouns).)⁵
- (3) agreement (e.g., English present tense; cf. *He helps us* and *We help him*. Note the verbal suffix *-s* that occurs on English present tense verbs when A is 3rd-person singular (as in *He helps us*). When O is 3rd-person singular (as in *We help him*), the *-s* suffix does not occur. Thus, in this restricted environment, the form of the verb agrees with the number of A (but it is unaffected by the number of O).)

Now, for many of the world's languages, the formal device used to encode S is the same as that used to encode A, but not O. Many of these languages have a case system, in which the nominative case is used for S and A, and the accusative case is used for O. For example, we can refer to the English pronoun set *I, he, she, we, they* as nominative, and the set *me, him, her, us, them* as accusative; the nominative set is used for S (e.g., *She is sleeping*) and A (e.g., *She helped me*), and the accusative set is used for O (e.g., *She helped me*). By convention, a language in which the same formal device is used to encode S and A, but not O, is said to be 'nominative-accusative' (or its shortened term, 'nominative'); a language in which the same formal device is used to encode S and O, but not A, is said to be 'ergative-absolutive' (or its shortened term, 'ergative').⁶ In a nominative language, S and A are said to be nominative, and O is accusative. In an ergative language, S and O are said to be absolutive, and A is ergative.

3. Ergativity in Sinama morphology

In Sinama, pronouns, common nouns, and proper nouns all have ergative case marking. This pattern is seen most clearly when pronoun referents occur. Sinama has three pronoun sets: absolutive, ergative, and oblique.⁷ As illustrated in (1)–(3), S and O are absolutive, and A is ergative.⁸

⁵ English no longer has a complete set of personal pronouns to distinguish A from O; the A set includes *I, you, he, she, we, you, they*, while the O set includes *me, you, him, her, us, you, them*. That is, *you* (2nd-person singular or plural) is used both for A and O.

⁶ It is considerably more precise to speak of a nominative vs. ergative system within a given language, rather than referring in a general way to the language itself as being nominative or ergative. A language may exhibit a nominative-accusative pattern in some features, but an ergative-absolutive pattern in others.

⁷ See the appendix for a table of Sinama pronouns. See Trick (1997:126–127) for data demonstrating morphological ergativity in Sinama.

⁸ The orthography of Sinama consists of 17 consonants and 5 vowels: *b* [b], *d* [d], *g* [g], *h* [h], *j* [dʒ], *k* [k], *l* [l], *m* [m], *n* [n], *ng* [ŋ], *p* [p], *r* [r], *s* [s], *t* [t], *w* [w], *y* [j], *a* [ə][e], *e* [e][ɛ], *i* [i][i], *o* [o], *u* [u]. Glottal stop is a phoneme. It is represented by *h* when it occurs syllable-finally (e.g., *lumah-na* [lu.məʔ.nə] 'his/her house') and by a hyphen when it occurs syllable-initially between morphemes (e.g., *mag-adjal* [mæg.'ʔəd.dʒəl] 'to cook'). It is not represented when it occurs intervocally (e.g., *pütu* [pi.'ʔi.tu] 'to come here') or word-initially (e.g., *eroh* [?'ɛ.roʔ] 'dog').

- (1) Tuli akú gana-gana.
 tuli akú gana-gana
 sleep 1SG.ABS later
 S

‘I (S) will sleep later.’

- (2) Tabangan-na akú.⁹
 tabang -an -na akú
 help -CL -3SG.ERG 1SG.ABS
 A O

‘S/he (A) will help me (O).’

- (3) Tabangan-ku iyá.
 tabang -an -ku iyá
 help -CL -1SG.ERG 3SG.ABS
 A O

‘I (A) will help him/her (O).’

When S or O is encoded by a full NP (whether a common noun or a proper noun), it has no case marking (4)–(7).

- (4) Tuli anak-anak gana-gana.
 tuli DUP- anak gana-gana
 sleep DIM- child later
 S

‘The child (S) will sleep later.’

- (5) Tuli si Ben gana-gana.¹⁰
 tuli si Ben gana-gana
 sleep PM name.person later
 S

‘Ben (S) will sleep later.’

Geminate consonants occur and are represented as a sequence of two identical segments (e.g., *addat* [ʔəd.dət] ‘custom’). Geminate vowels also occur, though with relatively low frequency. In general, the orthography does not represent geminate vowels; however, in a few cases where there may be ambiguity, an acute accent indicates geminate vowels (e.g., *pasód* [pə.'so.od] ‘to enter’).

⁹ Word order in Sinama tends to be VS and VOA; however, when an ERG pronoun occurs with a verb which is not prefixed with *ni-*, the ERG pronoun is bound to the right side of the verb and, thus, must precede O.

¹⁰ In Sinama, all personal names are preceded by the personal marker *si* (regardless of the grammatical relation or semantic role).

- (6) Tabangan-ku anak-anak.
 tabang -an -ku DUP- anak
 help -CL -1SG.ERG DIM- child
 A O

‘I (A) will help the child (O).’

- (7) Tabangan-ku si Ben.
 tabang -an -ku si Ben
 help -CL -1SG.ERG PM name.person
 A O

‘I (A) will help Ben (O).’

When A is encoded by a full NP (whether a common noun or a proper noun), it is preceded by *leh*, and the verb is obligatorily affixed with the agreement affix *ni-* (8)–(11).¹¹

- (8) Nitabangan anak-anak leh mastal.
 ni- tabang -an DUP- anak leh mastal
 AGR- help -CL DIM- child ERG teacher
 O A

‘The teacher (A) will help the child (O).’

- (9) Nitabangan akú leh si Ben.
 ni- tabang -an akú leh si Ben
 AGR- help -CL 1SG.ABS ERG PM name.person
 O A

‘Ben (A) will help me (O).’

- (10) Nitabangan anak-anak leh si Ben.
 ni- tabang -an DUP- anak leh si Ben
 AGR- help -CL DIM- child ERG PM name.person
 O A

‘Ben (A) will help the child (O).’

¹¹ Traditionally, verb agreement refers to an affix on the verb that indicates number, gender, case, person, or tense; furthermore, agreement usually occurs on verbs in both intransitive and transitive clauses. In Sinama, it appears that when A is a *leh*-marked phrase (which is obligatory when A is not a pronoun, and optional when it is a pronoun), the verb is prefixed with *ni-*. For want of a better term, we are currently calling *ni-* an agreement affix because of its cooccurrence with *leh*.

- | | | | | | | |
|------|------------|----|-------------|-----|----|-------------|
| (11) | Nisampak | si | Ben | leh | si | Wahid. |
| | ni- sampak | si | Ben | leh | si | Wahid |
| | AGR- slap | PM | name.person | ERG | PM | name.person |
| | | O | | A | | |

‘Wahid (A) will slap Ben (O).’

Note that the marker *leh* marks A only; it never marks S or O, as in *Tuli *leh si Ben* ‘Ben (S) will sleep’ or *Tabangan-ku *leh si Ben* ‘I (A) will help Ben (O)’.

4. Syntactic processes and patterns of control in Sinama

The previous section shows that in Sinama, case marking of S, A, and O follows a consistently ergative pattern. This section will demonstrate that syntactic control in Sinama also displays a high degree of syntactic ergativity. Specifically, it will show that S and O, and only S and O, are the syntactic control for relativization, clefting, WH-question formation, equi-NP deletion, and second-position cliticization.

4.1. Relativization

Relativization is a process by which a NP is modified by a subordinate clause. The subordinate clause is the relative clause, and the NP that it modifies is its head noun. Sentences (12) and (13) are independent clauses; in (14), the sentence in (13) functions as a relative clause. Note that in the clause which is relativized in (14), the O argument (referring to the rope) has been deleted (being coreferential with the head noun of the main clause).¹² Sentence (15) demonstrates that the A argument cannot be the head of a relative clause.

- | | | | | |
|------|------|------------|----------|--------|
| (12) | Bey | tandah-ku | | lubid. |
| | bey | ta- ndah | -ku | lubid |
| | PPFV | NCTRL- see | -1SG.ERG | rope |
| | | | A | O |

‘I saw the rope.’

- | | | | | | | |
|------|------|------------|-----|------------|-------|--------|
| (13) | Bey | nikottob | leh | anak-anak | lubid | itu. |
| | bey | ni- kottob | leh | DUP- anak | lubid | itu |
| | PPFV | AGR- cut | ERG | DIM- child | rope | D1.ABS |
| | | | A | | O | |

‘A/The child cut this rope.’

¹² In the examples throughout sections 4.1 through 4.4, notations are made to indicate which argument is deleted in the various syntactic processes.

- (14) Bey tandah-ku lubid ya bey nikottob
 bey ta- ndah -ku lubid ya bey ni- kottob
 PPFV NCTRL- see -1SG.ERG rope NMZ PPFV AGR- cut

∅ leh anak-anak.
 ∅ leh DUP- anak
 ∅ ERG DIM- child
 O=∅ A

‘I saw the rope which a/the child cut.’

- (15) *Bey tandah-ku anak-anak ya bey nikottob(-na) lubid ∅.¹³

‘I saw the child who cut the rope.’

Sentence (16) may appear on the surface to illustrate that an A argument may be deleted; however, note that in this sentence, the relativized clause is an antipassive construction. (That is, it has been detransitivized;¹⁴ the deleted referent is S, the single required argument of an intransitive clause.) As such, the structure of the relative clause in (16) is very similar to that in (18), which is derived from the clearly intransitive clause of (17).

- (16) Bey tandah-ku anak-anak ya bey
 bey ta- ndah -ku DUP- anak ya bey
 PPFV NCTRL- see -1SG.ERG DIM- child NMZ PPFV

ngottob ∅ lubid.
 ngaN- kottob ∅ lubid
 INTR- cut ∅ rope
 S=∅

‘I saw the child who cut a/the rope.’

- (17) Bey nengge anak-anak.
 bey ngaN- tengge DUP- anak
 PPFV INTR- stand DIM- child
 S

‘The child stood.’

- (18) Bey tandah-ku anak-anak ya bey
 bey ta- ndah -ku DUP- anak ya bey
 PPFV NCTRL- see -1SG.ERG DIM- child NMZ PPFV

¹³ The parentheses around the ergative pronoun *-na* indicate that this sentence is ungrammatical with or without the ergative pronoun.

¹⁴ Note the *ngaN-* intransitive prefix on the verb.

nengge		Ø.
ngaN-	tengge	Ø
INTR-	stand	Ø
		S = Ø

‘I saw the child who stood.’

4.2. Clefting

A cleft construction is one in which a NP is deleted from the main clause and reappears as a fronted head noun; in Sinama, this head noun is cross-referenced on the nominalized verb. Cross-linguistically, the structure of relative clauses and cleft constructions tends to be similar.

Cleft constructions in Sinama follow an ergative pattern of syntactic control in that only S (19)(20) and O (21)(22) may be the head noun.¹⁵ The head noun precedes the nominalized clause. The argument in the nominalized clause that is coreferential with the head noun is obligatorily absent. If the head noun is a pronoun, the pronoun is from the oblique class.¹⁶

(19)	Bey	nengge	si	Ben.
	bey	ngaN-	tengge	si Ben
	PPFV	INTR-	stand	PM name.person
				S

‘Ben stood.’

(20)	Si	Ben	ya	na	bey	nengge	Ø.
	si	Ben	ya	na	bey	ngaN-	tengge Ø
	PM	name.person	NMZ	LK	PPFV	INTR-	stand Ø
							S = Ø

‘Ben is who stood.’

(21)	Bey	nikottob	lubid	itu	leh	anak-anak.
	bey	ni-	kottob	lubid	itu	leh DUP- anak
	PPFV	AGR-	cut	rope	D1.ABS	ERG DIM- child
				O		A

‘A/the child cut this rope.’

¹⁵ The syntactic control may also be a recipient, an instrument, or a beneficiary that has been promoted to O.

¹⁶ For example, (20) would be: *Íya ya na bey nengge*. ‘He/She is who stood.’

- (22) Lubid itu ya bey nikottob Ø leh anak-anak.
 lubid itu ya bey ni- kottob Ø leh DUP- anak
 rope D1.ABS NMZ PPFV AGR- cut Ø ERG DIM- child
 O = Ø A

‘This rope is what a/the child cut.’

As (23) shows, the A argument cannot be the head of a cleft construction; however, if the transitive clause is changed to an antipassive construction so that the A argument becomes an S argument, then the argument can be the head of a cleft (24).

- (23) *Anak-anak ya bey nikottob(-na) lubid itu.

‘A/The child is who cut this rope.’

- (24) Anak-anak ya bey ngottob lubid Ø.
 DUP- anak ya bey ngaN- kottob lubid Ø
 DIM- child NMZ PPFV INTR- cut rope Ø
 S = Ø

‘A child is who cut (or cut at) a rope.’

If an oblique NP is to become a head noun, it must first be promoted to O (i.e., direct object), in which case it loses its oblique case marker and is cross-referenced on the verb by an appropriate affix. In (25)–(28), an oblique recipient is promoted to become the head of a cleft construction; in (29)–(32), an oblique beneficiary is promoted to become head of a cleft.

- (25) Nipamuwan búk leh mastal ni anak-anak.
 ni- pangaN- buwan búk leh mastal ni DUP- anak
 AGR- TR- give book ERG teacher to DIM- child
 O A OBL

‘A/The teacher will give the book to a/the child.’

- (26) Búk itu ya na nipamuwan Ø leh mastal
 búk itu ya na ni- pangaN- buwan Ø leh mastal
 book D1.ABS NMZ LK AGR- TR- give Ø ERG teacher
 O = Ø A

ni anak-anak.
 ni DUP- anak
 to DIM- child
 OBL

‘This book is what a/the teacher will give to a/the child.’

- (27) Nibuwanan anak-anak búk leh mastal.
 ni- buwan -an DUP- anak búk leh mastal
 AGR- give -VI DIM- child book ERG teacher
 O (promoted RECIP) A

‘A/The teacher will give the child a book.’

- (28) Anak-anak ya na nibuwanan búk Ø leh mastal.
 DUP- anak ya na ni- buwan -an búk Ø leh mastal
 DIM- child NMZ LK AGR- give -VI book Ø ERG teacher
 promoted RECIP = Ø A

‘A/The child is to whom a/the teacher will give a book.’

- (29) Adjal-na keyk itu ma kau.
 adjal -na keyk itu ma kau
 cook -3SG.ERG cake D1.ABS LOC 2SG.OBL
 A O OBL

‘She will bake this cake for you.’

- (30) Keyk itu ya adjal-na Ø ma kau.
 keyk itu ya adjal -na Ø ma kau
 cake D1.ABS NMZ cook -3SG.ERG Ø LOC 2SG.OBL
 A O = Ø OBL

‘This cake is what she will bake for you.’

- (31) Adjalan-na kow keyk.¹⁷
 adjal -an -na kow keyk
 cook -VI -3SG.ERG 2SG.ABS cake
 A O (promoted BEN)

‘She will bake you a cake.’

- (32) Kau ya adjalan-na Ø keyk.
 kau ya adjal -an -na Ø keyk
 2SG.OBL NMZ cook -VI -3SG.ERG Ø cake
 A promoted BEN = Ø

‘You are for whom she will bake a cake.’

¹⁷ In (31) and (32), we use a pronoun to further exemplify that this argument is marked as absolutive in a transitive clause and as oblique in a cleft construction.

4.3. WH-question formation

A WH-question (also called an information question or a content question) is one which contains a pro-form (e.g., English *who*, *what*, *where*, *why*, *when*). In many languages, this pro-form occurs in a clause-initial position, resulting in a gap at the position where the questioned argument occurs in the non-question form (cf. *He will give the book to you* and *What will he give \emptyset to you?*).

In Sinama, WH-question formation follows an ergative pattern of syntactic control: S and O may be questioned; A may not. In the following sentences, (33) and (34) show that S may be the questioned element of a WH-question. Sentences (35) and (36) show that O may be the questioned element, and (37) and (38) show that once an oblique NP is promoted to O (i.e., direct object), it also may be questioned.

- (33) Bey nengge anak-anak.
 bey ngaN- tengge DUP- anak
 PPFV INTR- stand DIM- child
 S

‘The child stood.’

- (34) Siyan bey nengge \emptyset ?
 siyan bey ngaN- tengge \emptyset
 who PPFV INTR- stand \emptyset
 S = \emptyset

‘Who stood?’

- (35) Bey pamuwan búk leh danda ni anak-anak.
 bey pangaN- buwan búk leh danda ni DUP- anak
 PPFV TR- give book ERG female to DIM- child
 O A

‘A/The woman gave the book to a/the child.’

- (36) Iyan bey pamuwan \emptyset leh danda ni anak-anak?
 eyyan bey pangaN- buwan \emptyset leh danda ni DUP- anak
 what PPFV TR- give \emptyset ERG female to DIM- child
 O = \emptyset A

‘What did a/the woman give to the child?’

- (37) Bey nibuwanan anak-anak búk leh danda.
 bey ni- buwan -an DUP- anak búk leh danda
 PPFV AGR- give -VI DIM- child book ERG female
 O (promoted RECIP) A

‘A/The woman gave the child a book.’

- (38) Siyan bey nibuwanan Ø búk leh danda?
 siyan bey ni- buwan -an Ø búk leh danda
 who PPFV AGR- give -VI Ø book ERG female
 promoted RECIPIENT = Ø A

‘To whom did a/the woman give a book?’

Sentences (39)–(41) show that A cannot be the questioned element unless the transitive clause changes to an antipassive construction. Following antipassivization, A becomes S, and the argument is now eligible to be questioned.

- (39) *Siyan bey pamuwan buk Ø ni anak-anak?
 *A = Ø

‘Who gave the book to a/the child?’

- (40) Bey akú muwan búk ni anak-anak.
 bey akú ngaN- buwan búk ni DUP- anak
 PPFV 1SG.ABS INTR- give book to DIM- child
 S

‘I gave a book to the child.’

- (41) Siyan bey Ø muwan búk ni anak-anak?
 siyan bey Ø ngaN- buwan búk ni DUP- anak
 who PPFV Ø INTR- give book to DIM- child
 S = Ø

‘Who gave a book to the child?’

4.4. Equi-NP deletion

Equi-NP deletion is a syntactic process involving a main clause and a complement clause, in which an argument in the main clause is coreferential with one in the complement clause, and one of the coreferential arguments is deleted (usually the coreferential argument of the complement clause). The process is illustrated in the English sentences in (42)–(46).

- (42) I want the book.

- (43) I will sleep.
 S

- (44) I want to sleep.
 S = Ø

(45) I will give the book to her.
A O

(46) I want to give the book to her.
A = \emptyset

The verb *want* can take either a NP complement, as in (42), or a clause complement, as in (44) and (46). In (44), the one ‘wanting’ and the one ‘sleeping’ are coreferential. Similarly, in (46), the one ‘wanting’ and the one ‘giving’ are also coreferential. In both (44) and (46), the coreferential argument of the complement clause is deleted. Note that in (44), the deleted argument is S (the one sleeping), and in (46), the deleted argument is A (the one giving). Thus, in English, equi-NP deletion operates on a nominative pattern of syntactic control, since it is S or A (and not O) that is deleted.¹⁸

Although equi-NP deletion operates on a nominative pattern of control in most languages of the world, including Philippine languages, Sinama is an exception to this near universal pattern in that equi-NP deletion operates on an exclusive ergative pattern of control.¹⁹ That is, only S of an intransitive clause (49) and O of a transitive clause (51) are deleted under coreference, never A (52).²⁰

(47) Kabilahian-ku búk.
ka- bilahi -an -ku búk
INV- want -CL -1SG.ERG book

‘I want the book.’

(48) Tuli akú.
tuli akú
sleep 1SG.ABS
S

‘I will sleep.’

(49) Kabilahian-ku tuli \emptyset .
ka- bilahi -an -ku tuli \emptyset
INV- want -CL -1SG.ERG sleep \emptyset
S = \emptyset

‘I want to sleep.’

¹⁸ In English, it might appear that O or IO (indirect object) may also be deleted if the clause has first been made passive, as in *She wants to be given the book* (IO = \emptyset ?) or *The puppy wants to be given to the little girl* (O = \emptyset ?). In fact, though, a passive clause is a single-argument construction having only S. In the clause *She was given the book*, the pronoun *she* is S. Likewise, *puppy* in *The puppy was given to the little girl* is S.

¹⁹ This has also been demonstrated for Yakan (Brainard and Behrens 2002:161–163).

²⁰ For a fuller discussion of equi-NP deletion in Sinama, see Trick (1997).

- (50) Nilinganan akú leh si Ben.
 ni- lengan -an akú leh si Ben
 AGR- call -CL 1SG.ABS ERG PM name.person
 O A

‘Ben will call me.’

- (51) Kabilahian-ku nilinganan Ø
 ka- bilahi -an -ku ni- lengan -an Ø
 INV- want -CL -1SG.ERG AGR- call -CL Ø
 O = Ø

leh si Ben.
 leh si Ben
 ERG PM name.person
 A

‘I want Ben to call [me].’

- (52) *Kabilahian si Ben nilinganan akú Ø.

‘Ben wants to call me.’²¹

4.5. Second-position cliticization

A clitic is a form that has some features of an independent word but that is bound to another word (known as the host).²² In Sinama, when a host element such as *ley* ‘past perfect’, *bey* ‘past perfective (PPFV)’, or *maha* ‘negator (NEG)’ occurs clause-initially, and S or O is also a pronoun, the pronoun will move to the left of the verb into the second position of the clause, as in (54) and (56). This is not, however, the case with A (57)(58).²³

- (53) Nengge iyá.
 ngaN- tengge iyá
 INTR- stand 3SG.ABS
 S

‘She will stand.’

²¹ Sentence (52) is grammatical with the meaning ‘Ben wants that I will be called (by someone else)’. That is, the complement clause is passive, and not transitive. The deleted argument is not coreferential with *Ben*.

²² Other features of clitics are: 1) they are phonologically unstressed, 2) they usually attach to the edges of words (i.e., outside of derivational or inflectional affixes), 3) they function at the phrase or clause level, often having grammatical rather than lexical meaning.

²³ This pattern of ergative control for second-position clitics has also been noted for Sama Bangingi’ (Gault 1999) and Yakan (Brainard and Behrens 2002:127–131).

- (54) Bey iyá nengge.
 bey iyá ngaN- tengge
 PPFV 3SG.ABS INTR- stand
 S

‘She stood.’

- (55) Tabangan-na akú.
 tabang -an -na akú
 help -CL -3S.ERG 1S.ABS
 A O

‘She will help me.’

- (56) Maha akú tabangan-na.
 maha akú tabang -an -na
 NEG 1S.ABS help -CL -3S.ERG
 O A

‘She will not help me.’

- (57) *Maha-na akú tabangan.
 A O

‘She will not help me.’

- (58) *Maha-na tabangan akú.
 A O

‘She will not help me.’

5. Conclusion

The data presented here demonstrate that in addition to morphological ergativity, Sinama exhibits a high degree of syntactic ergativity. Specifically, S and O, and only S and O, control not only relativization, clefting, and WH-question formation, as in many Philippine languages, but also equi-NP deletion and second-position cliticization.²⁴ In each of these syntactic operations, S and O pattern the same way, and A patterns differently. Other Philippine languages demonstrate syntactic ergativity with respect to some of these processes, but to date published results have not documented syntactic ergativity to this degree.

²⁴ Preliminary research indicates that imperatives, reflexivization, and reciprocalization operate on a nominative-accusative pattern of syntactic control, but these processes are beyond the scope of this paper.

Abbreviations

A	more agent-like argument in transitive clause	LOC	location
ABS	absolutive	NCTRL	no-control
AGR	agreement affix	NEG	negator
BEN	beneficiary	NMZ	nominalizer
CL	verb classifier	O	more patient-like argument in transitive clause
D1.ABS	demonstrative, near, absolutive	OBL	oblique
DIM	diminutive	PM	personal marker
DU	dual	PPFV	past perfective
DUP	reduplication affix	RECIP	recipient
ERG	ergative	S	single argument of intransitive clause
INTR	intransitive	TR	transitive
INV	involuntary	VI	valence increaser
LK	linker		

Appendix - Sinama Personal Pronouns

Person	Number	Absolutive		Ergative		Oblique	
1	sing	<i>akú</i>	1SG.ABS	<i>-ku</i>	1SG.ERG	<i>áku</i>	1SG.OBL
	pl	<i>kamí</i>	1PL.ABS	<i>-kami</i>	1PL.ERG	<i>kami</i>	1PL.OBL
2	sing	<i>kow</i>	2SG.ABS	<i>-nu</i>	2SG.ERG	<i>kau</i>	2SG.OBL
	pl	<i>kam</i>	2PL.ABS	<i>-bi</i>	2PL.ERG	<i>kaam</i>	2PL.OBL
DU	sing	<i>kitá</i>	DU.SG.ABS	<i>-ta</i>	DU.SG.ERG	<i>kita</i>	DU.SG.OBL
	pl	<i>kitabí</i>	DU.PL.ABS	<i>-tabí</i>	DU.PL.ERG	<i>kitabí</i>	DU.PL.OBL
3	sing	<i>iyá</i>	3SG.ABS	<i>-na</i>	3SG.ERG	<i>iya</i>	3SG.OBL
	pl	<i>sigá</i>	3PL.ABS	<i>-sigá</i>	3PL.ERG	<i>sigá</i>	3PL.OBL

References

- Blust, Robert. 2005. The linguistic macrohistory of the Philippines: Some speculations. *Current issues in Philippine linguistics and anthropology, Parangal kay Lawrence A. Reid*, ed. by Hsiu-chuan Liao and Carl R. Galvez Rubino, 31–68. Manila: Linguistic Society of the Philippines and SIL Philippines.

- Brainard, Sherri. 1994. *Voice and ergativity in Karao*. Unpublished dissertation. Eugene: University of Oregon.
- Brainard, Sherri. 1996. Why the 'focused NP' is not the subject in Philippine languages: Evidence from Karao. *Philippine Journal of Linguistics* 27.1/2.1–47.
- Brainard, Sherri. 2003. Karao texts. *Studies in Philippine Languages & Cultures* 13.
- Brainard, Sherri, and Dietlinde Behrens. 2002. *A grammar of Yakan*. Manila: Linguistic Society of the Philippines Special Monograph Issue 40:1.
- de Guzman, Videia P. 1988. Ergative analysis for Philippine languages: An analysis. *Studies in Austronesian linguistics*, ed. by Richard McGinn, 323–345. Ohio University Monographs in International Studies, Southeast Asia Series, No. 76.
- Dixon, R.M.W. 1979. Ergativity. *Language* 55.59–138.
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Gault, JoAnn Marie. 1999. *An ergative description of Sama Bangingi?*. Manila: Linguistic Society of the Philippines, Special Monograph Issue 46.
- Gerds, Donna. 1988. Antipassives and causatives in Ilokano: Evidence for an ergative analysis. *Studies in Austronesian linguistics* (Ohio University Monographs in International Studies, Southeast Asia series, 76), ed. by Richard McGinn, 295–321. Athens, Ohio: Ohio University Center for International Studies.
- Hodder, Barbara M. 1999. Functions of cleft constructions in expository and hortatory discourse in Mayoyao Ifugao. *Philippine Journal of Linguistics* 30.1/2.31–74.
- Kroeger, Paul R. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford, California: CSLI Publications (Center for the Study of Language and Information).
- Mithun, Marianne. 1994. The implications of ergativity for a Philippine voice system. *Voice: Form and function* (Typological Studies in Language 27), ed. by Barbara Fox and Paul J. Hopper, 247–277. Amsterdam: John Benjamins.
- Payne, Thomas. 1982. Role and reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog. *Studies in Language* 6.75–106.
- Pebley, Carol Jean, and Sherri Brainard. 1999. The functions of fronted noun phrases in Kagayanen expository discourse. *Philippine Journal of Linguistics* 30.1/2.75–121.
- Schachter, Paul. 1976. The subject in Philippine languages: Topic, actor, actor-topic, or none of the above? *Subject and topic*, ed. by Charles N. Li, 493–518. New York: Academic Press.

Schachter, Paul. 1977. Reference-related and role-related properties of subjects. *Syntax and semantics 8: Grammatical relations*, ed. by Peter Cole and Jerrold M. Sadock, 279–306. New York: Academic Press.

Trick, Douglas. 1997. Equi-NP deletion in Sama Southern. *Philippine Journal of Linguistics* 28.125–144.

Walton, Charles. 1986. *Sama verbal semantics: Classification, derivation and inflection*. Manila: Linguistic Society of the Philippines.