Bit Personal Pronouns in a Northern Mon-Khmer Context

Nathan Badenoch

Center for Southeast Asian Studies, Kyoto University

This paper presents the personal pronoun system of the Bit language, as spoken in Luang Namtha province of Laos. The system differentiates between masculine, feminine in the second and third person pronouns. The dual is formed through a morphological process on the feminine form of each. Third person pronouns are also used as clitics indicating definiteness or familiarity, a function that is not seen in other languages of the region. These characteristics are discussed in the context of other Khmuic and Palaungic pronoun systems. Looking at the group, one can observe that there has been a semantic shift within the pronoun systems in which a natural gender has become grammaticalized as a marker for intimacy and number.

0.   Introduction

Bit is spoken in northern Laos and Yunnan by approximately 2,750 people. The vast majority of speakers is in Laos (2,372 according to the 2015 census), and they typically speak Lao and Khmu on a daily basis in addition to their own language. The Bit language has been known since Lefevre-Pontalis recorded 78 words in the late 1800s, from Bit speakers in Phongsaly. Despite some phonological and lexical differences, this is clearly the same language as that spoken today in Luang Namtha. Kosaka (1999) has published 305 words, also from Phongsaly, and presented a discussion of the affiliation of Bit within the northern Mon-Khmer branch based on both phonological and lexical comparison. Gao (高 2004) has published a book on the Buxing language of China, a closely related dialect of Bit, presenting basic information about the phonology and syntax of the language, including a wordlist of more than 2,000 items.

The Bit are more commonly known as Khabit, although in Laos this name is no longer politically acceptable because of the *kha* element that is now considered to be derogatory. In the official ethnic classification they are now known as Lao Bit. They themselves prefer the name *psi:ŋ*, which means 'person', and is the source of the Chinese *buxing*. This word is cognate with *ksi:ŋ* in Ksingmul, where it also means 'person'. In daily life, the Bit often use the terms *kbɛt* and *khabɛt* when speaking within their own community. The meaning of /bit/ - /bɛt/ is unclear.

The data presented in this paper was collected over the period of 2011-2015 in Ban Bompiang in Luang Namtha province of Laos. Lexical items are drawn from an

annotated lexicon currently being compiled by the author, while examples are extracted from stories and conversations recorded in the village.

The phoneme inventory of Bit is introduced in Table 1 below (Badenoch 2015). Bit has 10 simple vowels, all contrasting for length, and four diphthongs.

Table 1: Bit vowels

short			long			diphthong		
i	ɨ	u	ii	ɨɨ	uu	iə	ɨə	uə
e	ə	o	ee	əə	oo			
ɛ		ɔ	ɛɛ		ɔɔ			
a		ɒ	aa		ɒɒ		aɨ	

The consonant system includes 21 phonemes (Table 2), all of which can occur in word and syllable initial position. With the exception of the voiced stop series, all occur as coda as well.

Table 2: Bit consonants

initial					final				
p [p]	t [t]	k [k]	c [tɕ]	ʔ [ʔ]	p [p]	t [t]	k [k]	c [tɕ]	ʔ [ʔ]
ph [pʰ]	th [tʰ]	kh [kʰ]							
b [b]	d [d]		ɟ [ɟ]						
	s [s]			h [h]		s [yʰ]			h [h]
m [m]	n [n]	ŋ [ŋ]	ɲ [ɲ]		m [m]	n [n]	ŋ [ŋ]	ɲ [ɲ]	
	l [l]					l [l]			
	r [r]					r [r]			
w [w]		y [j]			w [w]		y [j]		

Bit is a disyllabic, iambic language, with many onset clusters and a rich system of morphology.

1.   Bit personal pronoun system

The Bit personal pronoun system is presented in Table 2.

Table 3: Bit personal pronouns

	singular	dual	plural inclusive	plural exclusive
1	yɔː	daː	ʔiː	ɟiə
2	paː (f)	rpaː		piə
	meː (m)			
3	koː (f)	rkoː		keː
	ŋɒː (m)			
	ʔɒː (i)			

With respect to synchronic description, noteworthy characteristics of this system are:

- The first person distinguishes between singular, dual, plural inclusive and plural exclusive.
- The second person distinguishes singular, dual and plural, in addition to masculine (m) and feminine (f).
- The third person is typologically similar to the second person in its differentiation between singular, dual and plural, but marks a third distinction, adding an inanimate (i) pronoun. The inanimate does not have dual or plural forms.
- Dual forms can be analyzed as the feminine form with an /r/ pluralizing prefix.

In this system, it is interesting to note that the second person feminine forms have a common *p-* element, while the third person feminine has a common *k-* element, which are shared across the singular, dual and plural. From this point of view, the masculine and inanimate forms are divergent, and seem to be marked within the system. With regards to the rhyme, it is noted that the first person plural exclusive and second person plural have a common *-iə*.

2. Feminine base for dual forms in 2p and 3p

As noted above, several structural features of the Bit pronoun system are noteworthy. This following analysis will focus second and third person forms, paying particular attention to the gender distinction and the morphology of the dual forms.

As noted above, the second and third person share an *r-* element in the dual forms. For the second person dual, this form can be analyzed as *r-paː*, reflecting an *r-* pluralizing prefix with the feminine form *paː* 'you (fem)'. The third person dual *rkoː* is formed the same way, with the plural *r-* prefixed to *koː* 'she'. It should be stressed that *rpaː* and *rkoː* are the only forms possible for the dual positions in the second and third person; that is, **rmeː*, **rŋɒː* and **rʔɒː* are not possible. This means that the expression of dual in these slots takes the feminine form as the

default, modifying it with a straightforward and transparent (but historical) morphological operation.

Thus, while the second and third person dual forms are based on the feminine, there is no sense of a gender distinction in actual usage so,

> (1)   tɛʔ   bi: rpa:ʔ
>   do   what   2p-dual
>   "What are you two doing?"

is used identically with two people of any sex. Similarly,

> (2)   rko: waʔ   tɛʔ kwa: peɲ   ʔah
>   3p-dual   to   do-travel-shoot   NEG
>   kreh   ʔandaɨ
>   get   anything
>   "Those two went hunting but didn't get anything."

can be used with men, women or a combination of the two.

In Bit, plural /r/ is most commonly used as an infix denoting plural in adjectives and statives. For example, *krdi:ŋ* 'large (pl) < *kdi:ŋ* 'large', *srŋaam* 'beautiful (pl) < *sŋaam* 'beautiful', referring to two or more people or things. In this sense, the /r/ should be considered non-singular, as there is no distinction between dual and plural. However, this morpheme is limited to /CrC/ position in a minor syllable, so infixing is done after an interim prefixing step, *trdoʔ* 'small (pl)' <*tdoʔ < doʔ* 'small'. Prefixing is a common part of Bit morphology, and will be taken as internal evidence for the proposal of an interim step in deriving the plural form of CVC adjectives and statives.

Aside from these personal pronoun forms, use of the plural /r/ morpheme as a prefix is limited to expressive morphology. For example, in *rsɨ:l rsɨ:l* 'describing many people are jumping up and down' < *sɨ:l* 'describing up and down motion' and *rok-hok* 'describing a small space left by objects that have fallen out or disappeared' < *hok* 'describing a small open space'[1], the r- prefix is attached to an expressive base.

The use of the *r-* prefix is interesting within the Bit system because it puts the prefixed dual forms within the expressive phonology of the language. If the expressive phonology is distinct from the prosaic, as suggested by Diffloth (1979), this makes the feminine duals marked morphophonologically, even as the feminine is semantically the default. From the standpoint of the second and third person pronouns, it seems as if the Bit pronoun system has at least a formal orientation towards the feminine.

---

[1] This is by way of rhyme copy and insertion.

3. Grammaticalization of the third person pronoun

In addition to their function as third person pronoun, *ko:* and *ŋɒ:* are used in its singular form in several other related ways. First is as a prefix to peoples' names, marking personal acquaintance and gender. Second, is as a definite article, marking nouns that have already been referred to in discussion.

3.1 Names: Marking direct acquaintance and gender

People who are known to the speakers have *ko:* or *ŋɒ:* prefixed to their names, depending up on their gender. For females, /ko:/ is shortened to [k-], while /ŋɒ:/ is shortened to syllabic [ŋ̩], as in the names *kɲa:* 'Nya (girl's name)' < *k-ɲa:* and *ŋ̩me:k* 'Mek, boy's name' < *ŋ-me:k*. This form is only used when refering to the person in the third person. When the person is addressed directly, no prefix is used.

This form is carried into names that involve kinship terminology. In Bit, parents are referred to as the parents of their oldest child. For example, the woman who is the mother of Bət is called, in both referential and address situations, as

(3)   meʔ   ŋbə:t
         mother   masc-Bət

and the father of the girl Pim, as

(4)   ʔo:ɲ   kpim
         father   fem-Pim

The use of *k-* and *ŋ-* is in principle obligatory if the referent person is known to the speakers. These prefixes may even be used with kinship terms, as in:

(5)   yɔ: ci: ʔu:c   sɔ:k   kmeʔ   yɔ:

         1-SG   IRR   return   look for fem-mother   1-SG

         "I am going to go back and look for my mother"

In storytelling especially, the names of animals, humans and important spirits, the central characters, are always prefixed with the appropriate form: *ŋrɔ:k* 'the toad', *kna:ŋ lun kna:ŋ la:* 'Ms Lun and Ms La', adn *ŋpɲa: ʔin* 'Indra'. The usage is so common that it could be argued that the noun with the prefixed pronoun could be considered a proper noun, for example *ŋrwa:y* 'the tiger' could be understood as Mr. Tiger.

3.2   Marking definiteness

Related to this usage, the third person singular pronouns are used as a clitic indicating definiteness or familiarity. When a noun is already known from the preceding discussion, *ko:* and *ŋɒ:* are placed in front of the noun, often in their reduced forms *k-* and *ŋ-*.

Bit is a head-initial language, so a personal pronoun following a noun indicates possession, as in

> (5)   *meʔ*   *ko:*
>   mother   3-SG-fem
>   'her mother'

> (6)   *mʔuə*   *ŋɒ:*
>   fish   3-SG-masc
>   'his fish'

However, the third person pronoun is commonly found preceding an animate noun. This marks a definite referent, something known in a specific context or having been referred to in the preceding discussion,

> (7)   *ko:*   *meʔ*
>   3-SG-fem   mother
>   'the mother'

> (8)   *ŋɒ:*   *mʔuə*
>   3-SG-masc   fish
>   'the fish'

where both the speaker and listener are familiar with the mother and the fish from the preceding discussion, or from some larger shared body of knowledge. The gender of non-human animates is discussed below. The third person plural *ke:* can also be used in this way for animates as well, for example

> (9)   ke: bi: mraʔ
>   3-PL   person-old
>   'the old people'

Similarly, a third person pronoun preceding an adjective marks a specific item or object, bearing the qualities of that adjective, such as *ŋɒ: kdi:ŋ* 'the big one' referring to fish (usually masculine), or *ko: kdi:ŋ* 'the big one' referring to an insect (insects are feminine by default). In this usage, however, the inanimate third person is found commonly as well, *ʔɒ: kdi:ŋ* 'the big one' referring to a rock.

Used with verb phrases,

(10)   ŋɒː   ɲoʔ   rʔoː   hə:
        3-SG-masc   be at   down there   that
        'that guy down there'

(11)   ʔɒː   keː   dɔːm   ŋiʔ
        3-SG-inan   3-PL   take away   throw away
        'the one they took and threw away'

the pronoun functions to indicate a specific topic, modified by a verb phrase. Constructions such as these are often used with deictics *hee* 'this' and *hǝǝ* 'that', as in (10) above, further deepening the specification with a spatial reference.

4.   'Who' and gender marking

The third person singular pronoun is also used in the interrogative 'who', 'what', which', in its reduced forms /k-/ and /ŋ-/. Like the previous usage with the stative, all three forms can be used, depending upon the gender of the object under question, as in Table 4.

Table 4: Interrogative pronouns

Common reduced form	Underlying form	
kbiː	koː biː	'who' (default and fem)
ŋbiː	ŋɒː biː	'who' (masc.)
ʔbiː, biː	ʔɒː biː	'what, which' (inanimate)

In these constructions, the *biː* element is a general interrogative particle, also used in interrogatives such as *ɲaːm biː* 'when', *nɛːw biː* 'how' and *sɨ biː* 'how many'.

Here again, it should be noted that with regards to animate objects, the default form for 'who' is the feminine form *kbiː*. The masculine form is used only when one is sure that the person under question is masculine.

5.   Fauna, flora and semantic gender

The question word 'who' default usage of the feminine form, seems to be part of a larger principle in the Bit linguistic world that defaults animates as feminine. For example, when referring to animals, the personal pronoun used as default is also the feminine. Generally speaking, with large mammals such as deer, the sex is often clear and the appropriate form is used. However, particularly with insects, the sex is not usually clear and *koː* is used. Normally, neither masculine nor feminine forms of the third person pronoun are used with flora. Instead, the inanimate form *ʔɒː* is used. As seen in the examples above, a fish may be referred to as masculine, but from observations and discussions with native speakers, there is a tendency to take the feminine as the default for animates.

One interesting exception to this rule is in storytelling, when flora are often treated as animates because they appear as characters in the stories. Narrators quote thoughts, speech and actions of flora as if they were human. In this case, the gender depends upon the flora or fauna's role in the story. For example there is a well-known story of a tree that falls in love with a woman, and *ŋɒ:* is used when refering to this character in the story. This is necessary because in the stories animals and trees often mate with humans, although in these cross-relationships the female is usually the human.

6. Bit and Bumang/Khang

To begin a regional consideration of the Bit personal pronoun system, a look at other Austroasiatic languages spoken in the mountains of the Laos, Vietnam, China border area provides useful context. Although not much is known about the classification of these many and small languages, there seems to be good evidence that Bit and Bumang/Khang are closely related to each other. Bumang/Khang is a group of languages consisting of varieties spoken in Vietnam (Khang: 13,800 speakers according to the 2008 census) and Yunnan (Bumang: 200 speakers according to Dao 2007). Providing compelling evidence to demonstrate these relationships is an urgently needed contribution, but is beyond the scope of this paper. Suffice it to say here that Bumang/Khang share a number of phonological and lexical innovations with Bit, and the pronoun system is one of these points of interest. Bumang/Khang has undergone significant restructuring of the canonic word, reducing the older disyllabic structure to a mostly monosyllabic one, reducing onset clusters and developing tones in the process.

The pronouns are not identical, but similarities are worth exploring in some detail here. Table 5 shows the Bumang pronoun system, which will serve as a point of reference, using the original notation of Dao (2007):

Table 5: Bumang pronouns

	singular	dual	plural inclusive	plural exclusive
1 p	da^{55}	da^{55} bɯa^{24}	nɔ^{24}	jia^{55}
2 p	pa^{55} (f)	pa^{55} bɯa^{24}	mət^{24}	pia^{55}
	mi^{55} (m)	mi^{55} bɯa^{24}		
3 p	ku^{33} (f)	ku^{33} bɯa^{24}	ki:	
	ŋa^{55} (m)	ŋa^{55} bɯa^{24}		

From the start, we find a first person singular form that is different from Bit, but corresponds to the Bit first person dual form. Moving to the dual, we see that the Bumang dual is marked not with an pluralizing prefix as in Bit, but rather with the word *bɯa^{24}*, throughout the all three persons and genders. This word means 'two' < PMK *ɓa:r, giving a dual form of 'I-two', 'you-two', 'she-two', and 'he-two',

and is cognate with the Bit word *buər*, which means 'two' (but is preserved only in a few elaborate expressions, the native Bit numeral system having been almost completely replaced with Tai forms). Because Bumang word shape has been so drastically restructured, loss of minor syllables one of the notable changes, it is likely these dual forms were created to compensate for the loss of morphology in the language.

In the plural, the first person differentiates between inclusive and exclusive, like Bit. The exclusive form corresponds directly with the Bit form, but the inclusive form is different. In Bit, there is a form *nɔʔ*, which means 'companion', and this is likely cognate with the Bumang form *nɔ²⁴*. The 24 tone in Bumang regularly corresponds to final -ʔ in Bit[2], providing phonological evidence for cognacy. From the semantic point of view, the Bit *nɔʔ* implies a familiar, 'in-group' relationship that maps nicely to the idea of inclusive. The second and third person plural forms are also clearly cognate with the Bit forms.

However, our main interest here lies in the second and third person singular forms that differentiate for gender. The Bumang forms are very similar to the Bit forms, with just a slight difference in vowel quality. The second person forms are very similar, differing only in the height of the vowel (Bumang /i/, Bit /e/) in the masculine form. In the third person, the Bumang feminine form *ku³³* slightly higher than the Bit *koː*, and the masculine form *ŋa ⁵⁵* slightly more fronted than the Bit *ŋɒː*. There is no inanimate third person form in Bumang.

The Khang forms (Ueda ed. 2003) are given in Table 6, maintaining the original notation. The tones are not relevant to the discussion here.

Table 6: Khang personal pronouns

	singular	**plural**
1	zɔ¹	mot²dia¹
		mo¹zɔ¹
2	mi¹	puɯŋ³ʔaj ⁴
	ʔaj ⁴	puɯŋ³ʔem¹
	ʔem¹	puɯŋ³luaŋ ⁴
	luaŋ ⁴	
3	ko¹ (f)	ke¹
	ŋa¹ (m)	puɯŋ³ʔaj ⁴ni¹
		puɯŋ³ʔem¹ni¹

This data is from a rapid survey, so many details, including the lack of a dual form, cannot be confirmed. The second person seems to be in the process of restructuring, but the *mi¹* form is present. With the second person forms we find *ʔay⁴* and *ʔem¹*, the Tai Dam words for 'father' and 'mother'. Interestingly, according to this data

---

[2] Bumang 24 tone also corresponds to several other coda in Bit, including many final –r words, as in the /bɯə²⁴/ - /buər/ pair.

the *pa* form has fallen out of the system, but this needs to be confirmed with more fieldwork. The semantic difference between the other forms is not clear at this point. However, the gender distinction is preserved in the *ko¹/ŋa¹* pair in the third person.

With regards to the third person forms, the *ko:/ŋʋ:* is also interesting because the plural form in Bit, Bumang and Khang all retain the k- element, but with an unrounded vowel /i~e/. At this point, there is no evidence of Bumang or Khang grammaticalization of third person pronouns as in Bit. Again, however, this is perhaps not surprising given the degree of restructuring that the Bumang/Khang word has undergone. With the loss of morphology, the dual distinction was lost, and there was no compensation.

The first person singular form *zɔ¹* is cognate[3] with the Bit *yɔ:*. This suggests that the Bumang form *da⁵⁵* replaced the older *yɔ:/zɔ¹* form at some point, which is not difficult to imagine, considering that a non-singular form is often used as the singular form in casual or intimate speech in many languages in the region. Furthermore, it should be noted that the first person cognate *yɔ:/zɔ¹* in Bit-Khang is not seen elsewhere, suggesting another innovation for this group.

7.   Khmuic *me/ba* semantics

While the third person pronouns, marking gender, in Bit are not found beyond its close relative Bumang/Khang, a complex set of markings using the second person pronoun is widely attested. In Khmuic languages one finds personal pronoun cognates with the Bit *me:* and *pa:*, but the semantics of this pair differs across languages.

7.1 Khmu pronouns

The Khmu personal pronouns[4]   are shown in Table 7 below. Note that the feminine-masculine distinction is made in the second and third person singular. In the second person, the forms are direct cognates with Bit, while in the third person the forms are not related, even if the general scheme is the same.

Table 7: Khmu pronouns

	singular	dual	plural
1	ʔoʔ	ʔaʔ	ʔiʔ
2	ba: (fem)	sba:	bɔ:
	me: (masc)		
3	na: (fem)	sna:	nɔ:
	gə: (masc)		

---

[3] The Khang /z/-Bit /y/ correspondence is regular.
[4] Khmu forms cited are from Suvilai (2002), representing the Eastern Khmu dialect that is more conservative in its preservation of voiced initials.

The second and third person duals are formed by prefixing /s-/ to the feminine. The plural forms are also derived from the feminine form of the singular pronouns, with a shared vowel alternation of *aː/ɔː*. The vocalic relationships across the forms in this system are much more regular than in Bit.

There is one problem with this comparative perspective on the second person feminine Bit *paː*. Khmu and Bit both preserve historical voiced stops, which means that a Khmu voiced bilabial in the second person singular feminine would normally give the same voiced stop in Bit.

7.2 Phong, Ksingmul and Mlabri: The shifting semantics of me/ba

A *me/ba* distinction is evident in other Khmuic languages, such as Phong, Ksingmul and Mlabri[5], but the distinction is not one of gender. In these languages, the cognate forms are not used to mark gender or person, but rather to mark T-V distinctions as known in the Indo-European languages, of respect or familiarity, in the singular. As these languages are not well-known, the full pronoun systems are provided below for Phong Laan, Phong Khami, as well as Ksingmul and Mlabri.

Phong is a group of languages, spoken in Huapanh province of Laos. Population is unclear. Ethnologue gives 1,000 speakers of 'Phong-Kaniang', but this underestimates the full Phong-speaking population. Government census data is unreliable because typically Khmuic Phong languages are confused with Vietic Pong languages. At this point in time, it seems that there are at least three main groups of Phong languages: Laan, Khami-Piat and Cepuang-Phen[6].

In Phong Laan, there is a distinction between familiar and respectful in the second and third persons, as shown in Table 8. Of interest here is the second person, where *baː* the respectful form and *mɔː* is the familiar form.

Table 8: Phong Laan Pronouns

	singular	dual		plural	
		inclusive	exclusive	inclusive	exclusive
1	ʔaɲ	bərʔaː	bərʔiː	ʔay	ʔiː
2	baː (respectful)		bərbaː		bay
	mɔː (familiar)				
3	briah (respectful)				ʔãh
	niː (familiar)				

---

[5] Ksingmul and Phong data are from the author's own fieldnotes (Jan-Feb 2014)

[6] Based on author's work on Laan, Khami and Cepuang, and in discussion with Gérard Diffloth.

Furthermore, the feminine respectful is combined with *ba:r* 'two', reduced to *bər-*, to create the second person dual. This reflects the Bit and Khmu principle of using the *p-* form as the defaul. This is carried forward into the plural as well. The *bər-* morpheme is then used to create a first person dual that distinguishes inclusive and exclusive. The lack of a third person dual form suggests that the system is in flux.

Phong Khami[7] has a second person familiar *mɨ:* and respectful *bia*, reflecting a similar situation to Phong Laan, as shown in Table 9. It is not clear if the respectful *bia* form is cognate with *ba:*, or if it is a form of a plural. According to my informants, there is no distinction in the third person.

Table 9: Phong Khami Pronouns

	singular	plural	
		inclusive	exclusive
1	ɲɔ:	mot ʔia	mot ɹi:
2	bia (respectful)		mot bia
	mɨ: (familiar)		
3	thə̃mal ʔah		ʔah

Ksingmul is spoken on both sides of the Lao-Vietnam border, in the area of Xiang Khor district of Huapanh province in Laos. As shown in Table 10, there are respectful-familiar pairs in both second and third person. The second person has *mih* for the respectful and *ba:* for the familiar, or even rude. These forms interestingly show the same final –h on m- form, but a simple long vowel on the b-form. In passing it is worth noting as well that the dual form is derived, in this case from the plural, in conjunction with the *biər* 'two'.

Table 10: Ksingmul Pronouns

	singular	dual	plural	
			inclusive	exclusive
1	ʔaɲ	ʔa: biər	ʔa:	ʔɛ:
2	ba: (familiar)	bɛ: biər		bɛ:
	mih (respectful)			
3	ʔɨn (familiar)			ʔnɛ:
	nu: (respectful)			

Richel (1995) identified the peculiar Mlabri pronoun system, where the plural forms are derived from the dual, shown in Table 11 below for Mlabri B[8]. The

---

second person forms have *mɛh* as the singular, and *bah* for the dual. The plurals are *bah jum* and *bah tiʔ* , 'you-DL group' and 'you-DL hand', respectively.

Table 11: Mlabri personal pronouns

	singular	**dual**	**plural**
1	ʔoh	ʔah	ʔah jum
2	mɛh	bah	bah jum, bah tiʔ
3	ŋ̥ɛʔ, ʔat tiʔ	ʔat bɛr	ʔat tiʔ

Ito (伊藤 2013) has introduced data for Mlabri A, which is similar in structure, forming the plural from the dual, but by combination with the word *tʰr̩ːŋ* 'five', thus giving *bah tʰr̩ːŋ* in the second person. According to Ito, the *mɛh* form is used as a familiar reference between males and females of a similar age group. Moreover, he has found an inclusive/exclusive distinction in the third person (which he calls 同集団 'same group' and 別集団 'different group'): *ʔah bɛr tʰʁŋ* and *jum ɲʌʔ*, respectively (伊藤 2015).

Accordingly, it seems fair to argue that not only has Mlabri taken the original female *ba*, for dual, but this has been projected into the plural as well. Ito (伊藤 2013) concludes that the dual is the unmarked form of the non-singular pronoun in Mlabri. The final *–h* of the Mlabri first and second person, singular and dual forms is noteworthy. This is the conservative section of the system, and it is probably the case that there was a leveling, with *–h* spreading from the first person to the second person.

From this look at the Khmuic *me/ba* issue, it is clear that there are several different types of marking going on with the use of these two forms. Mlabri only marks number with these forms, the other languages mark biological or social factors.

8. Palaung pronoun paradigm

Looking further into the Palaungic languages, we find a pronoun system that has a remarkably transparent paradigm of forming person and number in the non-singular. In the Palaungic languages, the *me/ba* distinction has been lost in the singular, with *me*-cognate forms such as *mɔj*, *məj*, *maj*, and *mɛ* for the second person singular, and the *ba* forming moving into the dual. In fact there is no differentiation made anywhere in the singular. The Mlabri *mɛh/bah* pair introduced aligns with the Palaungic system, with regards to the semantic significance of the initial m-/b- elements.

What has caught interest in the Palaung pronoun system is the transparent and regular system of marking number and inclusive/exclusive with the onset and vowel combinations in the non-singular. The process of leveling that has undergone in both the onsets and rhymes throughout the Palaung pronoun system is remarkable. Janzen and Janzen explained the non-singular system in their 1972

work, where the onset consonants mark person (/y-/ first person exclusive, /ʔ-/ first person inclusive, /b-/ second person, /g-/ third person), while vowels mark number (/-ai/ dual, /-e/ plural). The full non-singular system is shown in Table 12 below for Pule, adapted from Janzen and Janzen (1972):

Table 12: Pule non-singular personal pronoun template

	1 person				2 person		3 person	
	exclusive		*inclusive*					
Dual	y	ai	ʔ	ai	b	ai	g	ai
Plural	y	e	ʔ	e	b	e	g	e

This basic scheme holds true for all Palaung languages, as shown in Deepadung (2013). With the *ba* form moving into the dual and plural, this calls for some attention to the third person /g-/ as well. Recalling the Bit-Bumang/Khang third person pronouns, it seems probable that a reflex of the third person female (Bit *koː*, Bumang/Khang *kuː*) has become the default from and undergone the same shift into the dual and plural. The evidence to support this is weak, as the third person has undergone a significant amount of restructuring in these languages. The presence of a voiced velar in the Khmu third person singular /gəː/ seems relevant, but this is the masculine form. Nonetheless, amidst the dynamism that is evident in the pronoun systems, one might be able to suggest that the feminine forms of the second and third person, because of some default value within an older system, have moved out of singular into the non-singular. One explanation for such a shift could be aninterim step to an intimacy marking which then came to mark number.

9. Khasi footnote

Here we must take note of Khasi pronouns, where gender marking existed in both second and third person when Roberts wrote his classic description of the language in 1891. The forms given are: second person singular *me* (masculine) and *pha* (feminine)[9]; third person singular *ʔu* (masculine) and *ka* (feminine), and plural *ki*. Nagaraja (2014) gives a general second person singular *phi* which does not distinguish for gender, noting that the *me/pha* pair is archaic and now tabooed. The third person forms have been grammaticalized as definite articles, governed by the gender of the noun. The parallels with Bit are striking, not only in terms of form but also function. Importantly, these similarities suggest that the basic features of gender differentiation in the pronoun system represents an older system, from which the semantic changes discussed above departed. Thus the *ko/ŋa* pair in Bit and Bumang/Khang should be considered shared retentions from this system. The fact of cognancy between the feminine forms, Bit *koː* and Khasi *ka*, seems non-controversial, even though masculine forms remain to be explained. Yet this fact again reinforces the idea that the feminine historically bears a special semantic load within this system.

---

[9] The aspirated /ph/ is the regular correspondence with the Bit unaspirated /p/.

10. Discussion

The first person inclusive-exclusive has received attention, as a proposed characteristic part of the 'typical Mon-Khmer system' of personal pronouns (Deepadung 2013). Bringing Bit data into the analysis, together with forms from other Khmuic languages, highlights some other interesting questions within this realm of grammar. From the Bit perspective, the main similarities between all these pronoun systems lies in the second person /m-/ and /b-/ forms. Data on the semantic distinctions of this pair discussed in the paper is summarized by language in Table 12.

Table 12: Second person /m-/ and /b-/

	Bit, Bumang/Khang	Khmu	Phong	Ksingmul	Mlabri	Palaung
m-	masculine	masculine	intimate	intimate	singular	singular
b-	feminine	feminine	distant	distant	dual	dual/plural

In these six languages, the second person *m-/b-* notes several rather different semantic dynamics, including gender, intimacy and number. There is a range of relationships that could be drawn between the different semantics indicated here, including multiple paths of semantic shift. From the Bit point of view, it is possible to consider the shifts along a semantic horizon as in Diagram 1:

*Diagram 1: Shift in semantics of pronouns*
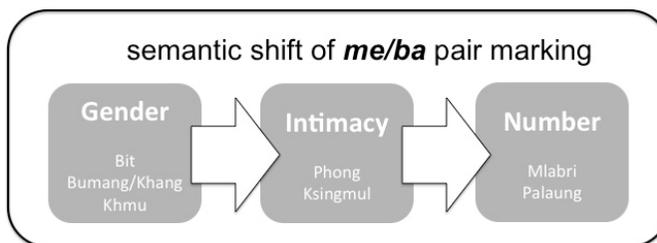
There is patchy, yet enticing evidence for a similar shift in the third person as well, with the feminine singular *ko:* becoming the base for the third person non-singular forms. Because the second person shift is underpinned by a hypothesis that it was motivated by a 'default' value favoring the feminine form within the animate world, it would make sense for the same transition to happen in the third person. The Bit data shows how a gender-marking system, together with morphology, could provide the basis for shifting semantics grounded in the social systems of the speakers. A tendency towards leveling in the sound system could then produce a pronoun paradigm as regular and transparent as Palaung. With several languages in

the interim stages of this horizon, some hints regarding the semantic, morphological, phonological and social dyanmics involved can be gleaned.

This suggests that the Bit system reflects an older system of organizing personal pronouns, which resulted in the Palaung template, through a process of semantic change such as the one discussed above, together with a thorough phonological leveling. Diagram 2 summarizes this proposal:

Diagram 2: Process of transformation in pronoun systems
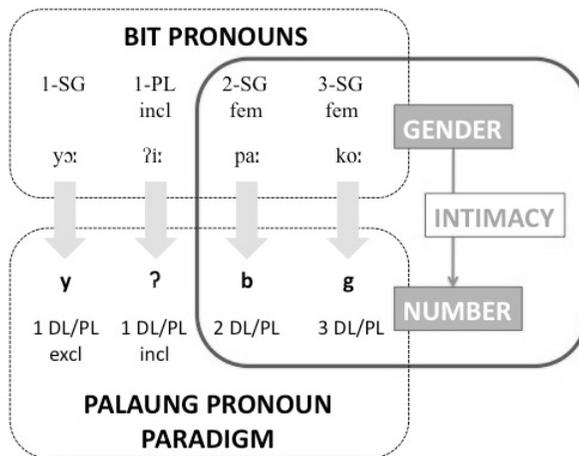
Long ago, Schmidt (1901) asserted that personal pronoun systems provided a potentially fruitful area of comparison among the Austroasiatic languages, with regard to classification. At the same time, he recognized changes in meaning, replacement of forms and other socially-driven factors, stating that these systems were in fact not stable and there is now a significant degree of divergence. Pinnow (1965) followed in this line, believing that broad-based comparison would indeed show that the majority of pronouns can be traced to common source in proto-Austroasiatic. He identified six semantic categories that should be considered in the structure of Austroasiatic pronouns: 1) person, 2) number, 3) exclusive-inclusive (in 1 per dual and plural, 4) classification by class or gender, 5) social status (politeness), and 6) case (Pinnow 1965). In the forgoing analysis we have seen how all of these factors may have worked together to bring about the systems currently observed in the northern Mon-Khmer languages.

Since then, it seems that there have been two issues of concern. First, there is a lexical issue, which seeks to map the contemporary forms to reconstructed proto-forms. Second, there has been discussion about a typically Mon-Khmer or Austroasiatic typology – the most well-known being the inclusive-exclusive distinction that is pervasive in the family. The Bit system, taken together with other evidence for gender distinction in the personal pronoun systems of Mon-Khmer (for example, Bahnaric languages), points to additional areas of questioning. For

example, Daladier (2002) has discussed animate noun classifiers, suggesting that Austroasiatic third person pronouns may be considered as a kind of anaphoric 'classifiers' that mark gender and number. These elements are often grammaticalized as nominal determiners, or act as clitics marking gender. Particularly interesting here is the /k-/ phoneme, which can be observed in an inordinate range of marking functions within the animate world. Moreover, Daladier argues that AA third person pronouns have produced prepositions, conjunctions, relative pronouns and interrogatives, or indefinites. The grammaticalization of the Bit third person pronouns can be considered a part of this phenomenon. Pinnow had already pointed out that the gender distinction could perhaps be better understood typologically as one between animate and inanimate. Here, he gives Khasi as an example of the directions of shift that can happen, as the gender distinction was still made in nouns (at the time of writing), but the original gender distinction in pronouns had become one of politeness. In this sense, one could consider the Bit and Bumang/Khang systems to be older manifestations of a system that has begun to shift towards the system marking social status of Ksingmul and Phong, and eventually Mlabri and Palaung.

The value of personal pronoun systems for classification could use some further discussion, based on new data that is coming to light. The data examined in this short discussion of several northern Mon-Khmer languages is a first step to broaden the discussion, identifying several directions of formal and semantic change within pronoun systems. Central as they are to the language, pronouns in this region are clearly dynamic and subject to many types of change. At the same time they may point to a number of issues pertaining to the classification of animate life, and the relationships between these different categories. There is a need to further examine the evidence from other Austroasiatic languages, and bring this data to a consideration of broader 'universal' trends that have been observed. With the larger discussion of a possible feminine bias in the pronoun systems, we are also encouraged to explore further the anthropological investigation of social structures and the changes they undergo. Investigation of matrilineal social structures in mainland Southeast Asia covers most all of the major linguistic phyla, and remains a fascinating question with implications for how society perceives and organizes itself in relation to the basic categories of animacy and gender that provide the biological foundations of the natural world.

Acknowledgments

References

Badenoch, Nathan. 2015. "Phonological Sketch of the Bit Language of Luang Namtha, Laos", paper presented at Southeast Asian Linguistics Society Annual Meeting, Payap University, Chiang Mai, May 27-29, 2015.

Daladier, Anna. 2002. "Kinship and Spirit Terms Renewed as Classifiers of "Animate" Nouns and their Reduced Combining Forms is Austroasiatic". Proceedings of the Twenty-Eighth Annual Meeting of the Berkeley Linguistics Society: Special Session on Tibeto-Burman and Southeast Asian Linguistics.

Diffloth, Gérard. 1979. "Expressive Phonology and Prosaic Phonology in Mon-Khmer", in T.L. Thongkum, V. Panupong, P. Kullavanijava and M.R. Kalaya Tingsabadh (eds), *Studies in Tai and Mon-Khmer Phonetics and Phonology,* p. 49-59. Bangkok: Chulalongkorn University Press.

Deepadung, Sujaritlak. 2013. "Personal Pronouns in Palaung Dialects". Language and Culture Documentation and Description 22. SIL International Digital Resources.

Janzen, Hermann and Margarete Janzen. 1972. "Grammar Analysis of Pale clauses and phrases", *Journal of the Burma Research Society,* 55.1-2:47-99.

Kosaka, Ryuichi. 1999. "Some features of the Khabet language – a Mon-Khmer language linking Khmuic and Palaungic branches", *Linguistic & Anthropological Study on the Shan Culture Area*, Report of Japanese Society for the Promotion of Science Grant-in-Aid project, Tokyo University of Foreign Studies, Institute for the Study of Languages and Cultures of Asia and Africa.

Lefèvre-Pontalis, Pierre. 1892-6. "Notes sur quelques populations du nord de l'Indo-Chine", JA 8e sér., 19 (1892):237-69.

Mikami, Naomitsu. 2003 *Khang*. Ueda H. ed. Reports on Minority Languages in Mainland Southeast Asia. Report A3-013. Endangered Languages of the Pacific Rim Project, Tokyo University of Foreign Studies.

Nagaraja, K.S. 2015. "Standard Khasi", in Jenny and Sidwell (eds) *Handbook of Austroasiatic Languages Vol. 2*. Leiden: Brill.

Pinnow, H. J. 1965. Personal pronouns in the Austroasiatic languages: A historical study. *Lingua* 14:3–42.

Rischel, Jørgen. 1995. *Minor Mlabri: A Hunter-Gatherer Language of Northern Indochina*. Copenhagen: Museum Tusculanum Press.

Roberts, H. 1891. *A Grammar of the Khasi Language.* London: Kegan Paul, Trench, Trubner & Co.

Schmidt, Pater Wilhelm. 1901. "Die Sprachen der Sakai und Semang auf Malacca und ihr Verhältnis zu den Mon-Khmer-Sprachen", *Bijdragen tot de Taal-, Land-, en Volkenkunde van Nederlandsch-Indië,* 52:399-583 52:399-583.

Suvilai Premsrirat. 2002. *Thesaurus of Khmu Dialects in Southeast Asia.* Mon-Khmer Studies-Mahidol University Special Publication No. 1 Volume 1. Mahidol: Institute of Languang and Culture for Rural Development.

Ueda, Hiromi. (ed) 2003. Reports on Minority Languages in Mainland Southeast Asia. Endangered Languages of the Pacific Rim Publications Series A3-013 Khang. Kyoto: ELPR.

伊藤雄馬　2015　「ムラブリ語の数詞」大西、千田、伊藤（編）『地球研言語記述論集』7，言語記述研究会、

　　　　　-　2013　「研究ノート　ムラブリ語の人称表現における体系性と双数形の無標性」*京都大学言語学研究*、32:77−101.

刀洁　《布芒语研究》 (中国新发现语言研究丛书) 北京：民族出版社 2007.

高永奇　《布兴语研究》(中国新发现语言研究丛书) 北京：民族出版社 2004.