Solving Problem 3 in the Eighteenth International Linguistics Olympiad (2021)

Jinyuan Wu

July 15, 2022

1 Data

(1) Bibani navasi yena minasina tetala tau.
'One man will catch these four fish.'

(2) Lekota dimdim mtona.
'This white man arrived.'

(3) Bikota gwadi magudiwena.
'That child will arrive.'

(4) Legisi waga makesiwena namwaya minana.
'This old woman saw those canoes.'

(5) Amtona tau lekalimati nayu bunukwa?
'Which man killed two pigs?'

(6) Leyamatasi teyu tauwau nunumwaya.
'The old women looked after two men.'

(7) Bigisi kwetala vivila minawena nakabitam.
'That clever woman will see something.'

(8) Navila ka'ukwa lekotasi?
'How many dogs arrived?

(9) Amakena waga legisesi gweguyau?
'Which canoe did the chiefs see?'

(10) Legisi dakuna makwena gwadi magudiwena gudimanabweta.
'That beautiful child saw this stone.'

(11) Kwevila lekamkwamsi dimdim mtosiwena?
'How many things did those white men eat?'

(12) Lekalimati natala bunukwa nagasisi guyau tokabitam.
'The clever chief killed one wild pig.'

(13) Navila vivila biyamatasi tau mtona?
'How many women will look after this man?'

(14) Navila vivila biyamata tomwaya mtona?

(15) Bikamkwamsi kweyu vivila minasina.

(16) Amagudina gwadi lekota?

(17) Tevila tauwau bigisesi gugwadi gudigasisi?

(18) Legisesi ketala waga vivila minasiwena.

2 Basic partition of the clause

2.1 Finding the verb

From (2), (3), and (8), which all have *arrive* as the main verb, it seems that

- *le-* is the past tense prefix;
- *bi-* is the future tense prefix;
- *-si* is the interrogative suffix.
- *-kota-* means 'arrive'.

It is possible that the apparent verb stem carries argument indexation, but this is not what can be found out here.

(2), (3) are declarative and are verb-initial, while the interrogative (8) is verb-final. It is not uncommon for the main verb to move in an interrogative clause, but it is also not uncommon to have *wh*-movement, so what happens in the interrogative clause is yet to be decided.

The initial position in a clause may be occupied by syntactic objects other than the main verb, like the topic, adjuncts, etc., but in the translations of data given in § 1, it seems such constructions are absent, and so we make the working hypothesis that the first words in the declarative sentences in § 1 are always the verbs. Thus,

- From (1), *-bani-* is 'catch'.
- From (4), (7) and (10), *-gisi-* is 'see'.
- From (5) and (12), *-kalimati-* is 'kill'.
- From (6), (13), (14), *-yamata-* is 'look after'.

The examples (6) and (14), however, does not support the previous claim that *-si* is the interrogative marker: (6) has *-si* but is not interrogative, while (14) does not has *-si* but is interrogative. Now we need to reconsider the syntactic context licensing *-si*. For those with *-si* on the main verb,

- (6), where both A and O are plural.
- (8), where S is plural.
- The potential example (9), where the verb is *legisesi*, which seems to be *le-gise-si* modified by a phonological rule. The A argument is plural, while O is singular.
- The potential example (11), where there is a word *lekamkwamsi* ending with *-si*.
- (13), where A is plural while O is singular.

So it can be generalized that *-si* appears if and only if S or A is plural. If this is the correct generalization, then Kilivila has the accusative property that S and A are treated equally in argument indexation, and *-si* means the subject – S or A – is plural.

Now we need to check whether clauses without *-si* on the verb always have singular subjects. Scanning (1) to (13), we find it seems to be the case. Then, from (11), *-kamkwam-* is 'eat'.

To summarize:

Box 2.1: Verb morphology and lexicon

Kilivila is an accusative language, and henceforth *subject* is S or A. Tense and number agreement with the subject are marked on the verb. The verb template is

tense	verb stem	subject number

The tense prefix can be the past tense *le-* or the future tense *bi-*. The subject number agreement suffix is *-si* when the subject is plural, and zero when the subject is singular. Here is a list of verb stems attested:

- *-kota-* arrive
- *-bani-* catch
- *-gisi-* see
- *-kalimati* kill
- *-yamata-* look after
- *-kamkwam-* eat

There is a remaining question not answered by Box 2.1: the verb in (9), which may involve phonological rules.

2.2 Constituent order

Now clause-level constituents of intransitive clauses (2) and (3) can already be found out:

(19) Constituent order of intransitive declarative clauses

a. Annotation of (2)

Le-kota [dimdim mtona]$_{\text{subject}}$
PST-arrive this white man

'This white man arrived'.

b. Annotation of (3)

Bi-kota [gwadi magudiwena]$_{\text{subject}}$
FUT-arrive that child

'That child will arrive'.

As is noted before, Kilivila is verb-initial, so VS constituent order is expected.

We already know there is no interrogative marker on the verb, but it is still possible that the interrogative clause type has to be marked by something else, possibly *wh*-movement. This seems to be the case since in (5), (8), (9), (11) and (13), the verb – which can be identified by comparing each word with the verb template in Box 2.1 – is not at the initial position. Then what precedes the verb is likely to be the *wh*-phrase, and hence the post-verbal area is filled by the remaining argument, so the clause-level constituents in these five clauses can be annotated:

(20) Constituent order of interrogative clauses

a. Annotation of (5)

[Amtona tau]$_{\text{subject, }wh}$ le-kalimati [nayu bunukwa]$_{\text{object}}$
which man PST-kill two pigs

'Which man killed two pigs?'

b. Annotation of (8)

[Navila ka'ukwa]$_{\text{subject, }wh}$ le-kota-si
how many dogs PST-arrive-PL

'How many dogs arrived?'

c. Annotation of (9) (*-gesi-* may have undergone phonological processes)

[Amakena waga]$_{\text{object, }wh}$ le-gise-si [gweguyau]$_{\text{subject}}$
which canoe PST-see-PL the chiefs

'What canoe did the chiefs see?'

d. Annotation of (11)

[Kwevila] le-kamkwam-si [dimdim mtosiwena]$_{\text{subject, }wh}$
how.many.things PST-eat-PL those white men

'How many things did thos white men eat?'

e. Annotation of (13)

[Navila vivila]$_{\text{subject, }wh}$ bi-yamata-si [tau mtona]$_{\text{object}}$
how many women FUT-look.after-PL this man

'How many women will look after this man?'

Comparing (20c) and (4), we find *waga* seems to mean 'canoe', and then the linear order of the word *waga* in (4) means in declarative clauses, the constituent order is VOS. *waga*

Therefore, the constituent order information is summarized as the follows:

Box 2.2: Clause constituent order

he clausal constituent order of declarative clauses in Kilivila is VOS. As for interrogative clauses, the *wh*-phrase is fronted.

2.3 Partition of declarative clauses

Now all interrogative clauses and intransitive declarative clauses have been divided into clausal constituents in (19) and (20). We need to complete the task for transitive declarative clauses.

Comparing (2) and (11), *dimdim* is likely to mean 'white man', and hence we find in Kilivila NPs, demonstratives follow head nouns. Comparing (7) and (13), it seems *vivila* means 'woman', and since we have the N Dem constituent order, in (7) the NP corresponding to 'that clever woman' is *vivila minawena nakabitam*, and therefore we have *dimdim* N Dem *vivila*

(21) Annotation of (7)

Bi-gisi [kwetala]$_{\text{object}}$ [vivila minawena nakabitam]$_{\text{subject}}$
FUT-see something that clever woman

'That clever woman will see something'.

This also means we have N Dem Adj constituent order, since *minawena*, by comparison with (20d), is obviously a demonstrative, and thus *nakabitam* has to be the adjective. N Dem Adj

From (9) and (12), we find *guyau* means 'chief', and the prefix *gwe-* may have plural markers or something that we are unable to decide now. Then, since the subject in (12) contains two concepts commonly marked by lexical words, while the object contains three, and there are exactly five words beside the main verb, the following glossing is highly plausible: *guyau* *gwe-*

(22) Annotation of (12)

Le-kalimati [natala bunukwa nagasisi]$_{\text{object}}$ [guyau tokabitam]$_{\text{subject}}$
PST-kill one wild pig the clever chief

'The clever chief killed one wild pig.'

And from the above annotation, we have N Adj constituent order in NPs. N Adj

Comparing (22) and (1), we find *tetala* and *natala* have similar forms, and a reasonable guess is these two are alternants of the numeral 'one' in different syntactic environments. Since Kilivila is VOS, *tetala tau* in (1) is likely to be a constituent filling the subject slot, and hence clausal constituents of (1) can be recognized: *tetala*, *natala*

(23) Annotation of (1)

Bi-bani [navasi yena minasina]$_{\text{object}}$ [tetala tau]$_{\text{subject}}$
FUT-catch these four fish one man

'One man will catch these four fish.'

Also, it seems numerals precede the head noun. Num N

Partition of (4) can be seen by the N Dem fact and comparison with (11): *makesiwena* seems to be a demonstrative, and thus by the N Dem constituent order, we have

(24) Annotation of (4)

Le-gisi	[waga makesiwena]$_{\text{object}}$	[namwaya minana]$_{\text{subject}}$
PST-see	those canoes	this old woman

'This old woman saw those canoes.'

Then it is likely that *namwaya* means 'old woman' as a whole. In a largely same line of reasoning, (10) is analyzed as the following: *namwaya*

(25) Annotation of (10)

Le-gisi	[dakuna makwena]$_{\text{object}}$	[gwadi magudiwena gudimanabweta]$_{\text{subject}}$
PST-see	this stone	that beatiful child

'That beautiful child saw this stone.'

By counting lexemes and comparison with (24), (6) can be annotated as

(26) Annotation of (6)

Le-yamata-si	[teyu tauwau]$_{\text{object}}$	[nunumwaya]$_{\text{subject}}$
PST-look.after-PL	two men	the old women

'The old women looked after two men.'

Now all examples with known translations are divided into clause-level constituents.

3 The noun phrase

The above discussion hints a high complicated inner structure of NPs. In this section, we first summarize plausible variatns of the same word and then search for the details of the syntactic environment of each variant.

3.1 Some quick observations

3.1.1 Variants of numerals

Here is a list of numerals attested:

- 'one': *tetala* in *tetala tau* 'one man' in (23), *natala* in *natala bunukwa nagasisi* 'one wild pig' (22).
- 'two': *nayu* in *nayu bunukwa* 'two pigs' in (20a), *teyu* in *teyu tauwau* 'two men' (26).
- There is 'four' hidden in (23), but currently we are unable to find it, because the relative order of numerals and demonstratives are not clear yet. By observing the form resemblance with demonstratives attested before, it seems *minasina* is the demonstrative, so one word in *navasi yena* has to be 'four'.

By observing the alternation in the form of the four numerals attested, it seems we have *-tela* 'one' and *-yu* 'two', while *te-* and *na-* are prefixes marking certain grammatical categories. Thus *navasi* is more likely to be the numeral 'four', while the remaining word *yena* is 'fish'. Thus we have the constituent order Num N Dem, and hence by comparison with (21), the structure of a declarative NP is Num N Dem Adj. *navasi*, *yena* Num N Dem Adj

3.1.2 Variants of nouns

Here is a list of nouns attested:

- *yena* 'fish' in (23).
- *dimdim* 'white man' in (19a), (20d).
- *tau* 'man' in (23), (20a). A variant *tauwau* is found in (26).
- *waga* 'canoe' in (24), (20c).
- *gwadi* 'child' in (19b), (25).
- *namwaya* 'old woman' in (26). A variant *nunumwaya* is attested in (26).
- *vivila* 'woman' in (21), (20e).
- *ka'ukwa* 'dog' in (20b).
- *dakuna* 'stone' in (25).
- *guyau* 'chief' in (22). A variant *gweguyau* is found in (20c).

Also, according to the Num N Dem Adj constituent order, from the 'one wild pig' NP in (22) we have the noun *bunukwa* 'pig' and the adjective *nagasisi* 'wild'.

3.1.3 Variants of adjectives

- *gudimanabweta*: 'beatiful' in (25).
- *tokabitam*: 'clever' in (22). The variant *nakabitam* is found in (21).
- *nagasisi*: 'wild' in (22).

3.2 Demonstratives

There are two types of demonstratives attested: proximate 'this' and distal 'that'. Comparing (19a) and (20e), we find Kilivila lacks case marking on at least the demonstrative, and the grammatical categories marked on the demonstrative are likely to be all from the head noun. Attested examples are summarized here, where the source, the form of demonstrative, and glossing of the head noun are listed:

	proximate, 'this' and 'these'			distal, 'that' and 'those'		
singular	(19a)	*mtona*	white man	(19b)	*magudiwena*	child
	(24)	*minana*	old woman	(21)	*minawena*	woman
	(25)	*makwena*	stone	(25)	*magudiwena*	child
	(20e)	*mtona*	man			
plural	(23)	*minasina*	fish	(24)	*makesiwena*	canoe
				(20d)	*mtosiwena*	white men

Here some examples without translation also provide useful information about how to find morphemes in the demonstratives:

- *minasina* in (15), which is attached to *vivila* 'woman';
- *minasiwena* in (18), which is also attached to *vivila*.

By comparing the above attested demonstratives, the following facts can be found:

- All of them end in *-na* and start with *m-*.

- The morpheme *-si-* appears, if and only if the NP is plural.
- It seems the distal feature is marked by *-we-*. The only exception to this is *makwena* in (25), but the *-we-* sequence in this word possibly origins from another morpheme.

So now *mtona* in (19a) and *mtosiwena* in (20d) can be labeled as

(27) Demonstratives pertaining to *dimdim* 'white man'

a. mto-na
???-DEM

b. mto-si-we-na
???-PL-DIST-DEM

Now the problem is what is marked by *mto-* and whether *mto-* can be decomposed into smaller units. Comparison between (19a), (20d), and (20e) reveals the stability of *mto-* in different argument slots and with different clearly masculine nouns, so it is highly likely that *mto-* contains a classifier corresponding to roughly the masculine noun class.

Similarly, *minana* in (24) and *minawena* in (21) can be glossed as

(28) Demonstratives pertaining to 'woman' and 'old woman'

a. mina-na
???-DEM

b. mina-we-na
???-DIST-DEM

And *mina-* seems to be – or contains – the classifier of the feminine noun class. The demonstrative *minasiwena* is also attested (in (18)), and comparison between it and (27b) means the order of the distal marker and the plural marker is always plural marker > distal marker. Interestingly, *mina-si-na* – CLS.female-PL-DEM – also works with 'fish', so the latter seems to be feminine in Kilivila.

Applying the above scheme to *magudiwena* and *makesiwena*, we find *magudi-* is or contains the classifier for the noun class containing 'child', while *make-* is or contains the classifier for the noun class containing 'canoe'. Then we may further decompose the two into

(29) a. ma-gudi-
DEM-CLS.child-

b. ma-ke-
DEM-CLS.canoe-

Thus a demonstrative prefix *ma-* is distinguished. Thus *makwena*, the apparent counterexample to the generalization that *-we-* is the distal marker, can be parsed as *ma-kwe-na*, where *-kwe-* is the classifier corresponding to the stone-like noun class.

Now compare this with (27) and (28). The question: why do we see alternation in the demonstrative prefix? What changes *ma-* into *mi-* and *m-*? Since *t* is a stop and the place of articulation *n* is about the teeth, the alternation of the initial *ma-* may simply be a result of phonological assimilation.

Box 3.1: Demonstratives and noun classes

The morphological structure of Kilivila demonstratives can be summarized as follows:

ma-	classifier	plural marker	distal marker	*-na*

When the demonstrative is a distal one, the distal marker position is filled by *-we-*, otherwise it is zero. When the head noun is plural, the plural marker position is filled by *-si-*, otherwise it is zero. The morpheme in the classifier slot corresponds to one of the following noun classes, before some of which the *ma-* prefix is altered, possibly for phonological reasons:

- *-na-*: the feminine class, members of which include *vivila* 'woman', *namwaya* 'old woman', and 'fish'. The *ma-* prefix becomes *mi-* with this classifier.
- *-to-*: the masculine class, members of which include *tau* 'man', *dimdim* 'white man'. The *ma-* prefix becomes *m-* with this classifier.
- *-gudi-*: the "child" class, members of which include *gwadi* 'child'.
- *-ke-*: the "canoe" class, memebers of which include *waga* 'canoe'.
- *-kwe-*: the "stone" class, memebers of which include *dakuna* 'stone'.

3.3 Interrogative pronouns

In the same line of reasoning in § 3.2, we list all interrogative pronouns attested here:

number interrogative: how many			interrogative determiner: which		
(20b)	*navila*	dog	(20a)	*amtona*	man
(20d)	*kwevila*	– (class: things)	(20c)	*amakena*	canoe
(20e)	*navila*	woman			

The three number interrogatives are distinguished by similar forms and by the interrogative pronoun > head noun constituent order.

The inner construction of interrogative determiners can be easily identified: we just need to add *a-* to the corresponding singular proximate demonstratives. They are glossed as

(30) *Which* pronouns

a. a-m-to-na
WH-DEM-CLS.masculine-DEM

b. a-ma-ke-na
WH-DEM-CLS.canoe-DEM

The numer interrogatives have a different ending: all of them end in *-vila*. Once this is observed, their inner structure is obvious: they are constructed by put a classifier and the *-vila* ending together. Besides, now we see the role of *-kwe-*: it is the class of "things" or "objects", like stones. The *-na-* class includes both women and dogs.

Here is a summary of interrogative pronouns:

Box 3.2: Interrogative pronouns

The interrogative determiner ('which') is made by adding *a-* to the singular proximate demonstrative pertaining to the head noun (see Box 3.1). The interrogative pronoun about number ('how many') is made by putting the *-vila* suffix to the classifier of the head noun.

3.4 The structure of adjectives

Now it can be seen that classifiers appear frequently in NP dependents. By comparing Box 3.1 and § 3.1.3, the structure of adjectives is similar to *how*-interrogatives: the first morpheme is the classifier, and the second morpheme is the adjective stem. The adjectives *tokabitam* 'clever' in (22) and *tokabitam* 'clever' in (21) can be glossed as

(31) Glossing of adjectives

a. to-katitam
CLS.masculine-clever
'clever (masculine)'

b. na-katitam
CLS.feminine-clever
'clever (feminine)'

The fact the adjective starts with *to-* in (22) means the noun *guyau* 'chief' is masculine, which is expected. The adjective *gudimanabweta* 'beatiful' in (25) has the same structure:

(32) gudi-manabweta
CLS.child-beatiful
'clever (child)'

Here *gudi-* is the classifier attested for children.

3.5 Noun morphology

§ 3.1.2 shows little alternation in noun morphology. There seems to be several nouns that have (irregular) plural forms: *tauwau* 'men', *nunumwaya* 'old women', *gweguyau* 'chiefs'.

3.6 The noun phrase

4 Result: grammar sketch

Here is a Basic Linguistic Theory (BLT)-based, bottom-up grammar sketch of Kilivila.

5 Result: lexicon