UKLO Round 2 March 2010

Data booklet

1. (5 points) Ardhay Uzzlepay

Minangkabau is spoken by about 7 million people around the West Sumatran city of Padang in Indonesia. Its speakers generally also speak Indonesian but Minangkabau is a distinct language. Minangkabau has a number of 'play languages' that people use for fun, like Pig Latin in English. Ordinary language words are changed into play language by following just a few rules. One of these 'play languages' is called **Sorba**.

Here are some examples of standard Minangkabau words and their Sorba play language equivalents:

Standard Minangkabau	Sorba	English Translation
raso	sora	'taste, feeling'
rokok	koro	'cigarette'
rayo	yora	'celebrate'
susu	sursu	'milk'
baso	sorba	'language'
lamo	morla	'long time'
mati	tirma	'dead'
bulan	larbu	'month'
minum	nurmi	'drink'
lilin	lirli	'wax, candle'
mintak	tarmin	'request'
cubadak	darcuba	'jackfruit'
mangecek	cermange	'talk'
bakilek	lerbaki	'lightning'
sawah	warsa	'rice field'
pitih	tirpi	'money'
manangih	ngirmana	'cry'
urang	raru	'person'
apa	para	'father'
iko	kori	'this'
gata-gata	targa-targa	'flirtatious'
maha-maha	harma-harma	'expensive'
campua	purcam	'mix'

1.1. (1 point) Using the rules that explain the words in the Table above, write the Sorba equivalents of the following standard Minangkabau words in the answer sheet.

	Standard Minangkabau	English Translation
a	*rancak*	'nice'
b	*jadi*	'happen'
c	*makan*	'eat'
d	*marokok*	'smoking'
e	*ampek*	'hundred'
f	*limpik-limpik*	'stuck together'
g	*dapua*	'kitchen'

1.2. (1 point) If you know a Sorba word, can you work backwards to a single standard Minangkabau word? Demonstrate with the Sorba word *lore* which means 'good'. (Use the answer sheet.)

1.3. (1 point) Another 'play language' is called **Solabar**. The rules for converting a standard Minangkabau word to Solabar can be worked out from the following examples:

Standard Minangkabau	Solabar	English Translation
baso	*solabar*	'language'
campua	*pulacar*	'mix'
makan	*kalamar*	'eat'

What is the Solabar equivalent of the Sorba word *tirpi* 'money'?

1.4. (2 points) In writing Minangkabau does the sequence 'ng' represent one sound or two sounds? Provide evidence that supports your answer. (Write your answer in the answer sheet.)

2. (10 points) Be Cree-ative

Cree is the most widely spoken of the Canadian aboriginal languages, with about 117,000 people speaking one of its many varieties. Here are six words in Plains Cree (Nehiyawewin), a dialect spoken across much of the Western Canadian prairie and in parts of Minnesota, written using the Roman alphabet:

tehtapiwin	'chair'	*mistikwan*	'head'
iskwahtem	'door'	*tipahikan*	'hour
sakahikan	'nail'	*astotin*	`hat'

2.1. (1 point) Below are six words, meaning 'little hat', 'little nail', 'little door', 'little head', 'minute', and 'little chair'. Which means which? (Use the answer sheet.)

1. cipahikanis	**2.** miscikwanis	**3.** cehcapiwinis
4. sakahikanis	**5.** ascocinis	**6.** iskwahcemis

Pronunciation note: When writing Cree in the Roman alphabet, the letter <c> represents the [ts] sound.

Although Cree can be written in the Roman alphabet, it is more frequently written in a writing system known as "**Syllabics**". This writing system has been adopted by speakers of other Canadian aboriginal languages as well; Inuktitut Syllabics are in wide use, and speakers of Ojibwe (Anishinaabemowin), Blackfoot, and Carrier (Dakelh) have also written their languages in Syllabics.

2.2. (5 points) The twelve words provided above in the Roman alphabet are given below (in random order) in Syllabics. Write their Roman alphabet equivalents in the answer sheet.

a. ∩<ⁱ'Δbᴐ   b. ◁ᴖJᴦσᴖ   c. ᐟbⁱ'Δbᴐ

d. ᒥᴖ∩bᴐ   e. ᴦ<ⁱ'Δbσᴖ   f. Δᴖbᐧⁱ'�ᒥᴖ

g. ᐟbⁱ'Δbσᴖ   h. Uⁱ'ᑕΛΔᴐ   i. ᒥᴖᴦbᐧσᴖ

j. Δᴖbᐧⁱ'Uᶜ   k. ◁ᴖᴐ∩ᴖᴐ   l. ᒉⁱ'ᒪΛΔᐧσᴖ

2.3. (4 points) Explain your answer (on the answer sheet).

3

3. (10 points) F u cn rd ths

Abbreviations are hard. We are used to thinking of standard abbreviations like km, UK, Mr or Ave. But in fact people make up new abbreviations all the time, if they are under time pressure (e.g. instant messaging) or if they have severe space limitations (e.g. classified ads in a printed newspaper).

One place where you find lots of abbreviations is the notes taken by the overworked people who staff call centres. They have to record what was discussed, but they don't have the time to type everything out. So you often get things that look like this, from the logs of a call centre run by a major telecommunications company:

cust rcvd lttr cncrng local srvc

which of course is supposed to mean:

customer received letter concerning local service

Let's say you are designing a computer program to try to flesh out abbreviated messages automatically. You can't just have a fixed list of abbreviations: the set is pretty open ended. But what the computer can do is to look at all the messages it's got, and hope that someone somewhere has spelled out the complete words. So if for example it is looking at *rcvd lttr*, and somewhere else in the database someone has done it the favour of reporting on a different call using the fully spelled phrase *received letter*, then it has a chance of guessing the expansion of *rcvd lttr*. That is, *rcvd* is a plausible abbreviation of *received*, *lttr* is a plausible abbreviation of *letter*, and the two occur together in the right order.

Of course, you know English, so you could have figured this out. But the computer really doesn't. To the computer the problem looks as follows: It has a bunch of words and phrases, some of which are abbreviated, written in a bunch of otherwise meaningless symbols (remember the computer doesn't know English and to it, the strings are ultimately just a bunch of characters anyway).

Here are some messages (A-R), transposed into symbols so that you see what the computer "sees", and below them a list of full phrases from elsewhere in the collection (1-15). There are two warnings to bear in mind:
    (1)   When you are under time pressure, you make mistakes. There are actually three typos in the abbreviations—typos in that all the letters are there, but are out of the expected order, and therefore are not strictly speaking reasonable abbreviations for the words.
    (2)   There are three phrases in the second list that are not found in the abbreviations, and (of course), any of these phrases may be represented by more than one of the abbreviations.

A. ⊤⊖⊙ ⊕∩⊔

B. ⊤⊙ ⊤±∘∘∩×

C. ⊤⊖⊘●⊕ ±×○⊖×

D. ∓⊘⊖　∓±∘∘∩✕

E. ∓⊙⊖　±✕○✳⊖∩✕

F. ∓∧•　⊙⊓⊖∩⊘

G. ∓⊖⊘⊕　∓∘±✳•⊖

H. ∓⊙⊖⊘•　∓∘∘✕

I. ∓•⊕　✕⊖∓∨∨∓⊘✕

J. ∓⊙⊖⊘⊕　⊙⊓⊖⊘

K. ∓⊖⊘　⊙∨✕∩⊕⊖⊘∧∧✕

L. ∓⊖⊙⊘　‡∨⊘✕

M. ∓⊖⊘•　‡∨⊘⊖

N. ∓⊖　∓∘∪

O. ∓⊖⊕　∓±∘∘∩✕

P. ∓⊖•⊕　∓∘∘∨∪

Q. ∓⊙⊖⊘　∓±⊕∩

R. ∓⊙⊖⊘∧　∓±∘∘

(1)   customer advised

(2)   customer call

(3)   customer called

(4)   customer calling

(5)   customer care

(6)   customer claims

(7)   customer disconnected

(8)   customer likes

(9)   customer needs

(10)   customer request

(11)   customer says

(12)   customer understood

(13)   customer upset

(14)   customer wanted

(15)   customer wants

3.1. (7 points) In the table on the answer sheet, match each of the encoded abbreviations from the previous page to one of the phrases above.

3.2. (3 points) What message is abbreviated in the symbols below?

4. (15 points) Tale of Kieu

Vietnamese is now written with a writing system consisting of Roman letters, just like English, but with lots of special markers or "diacritics" to show many distinctions between the sounds of many different vowels and six different tones. That writing system, called **Quốc Ngữ**, was developed by European missionaries, which is why Vietnamese is written with the same letters as European languages.

However, writing first came to Vietnam from China. At first, Vietnamese scholars actually wrote in Chinese, and because of the status of Chinese – as a language of government, literature, and culture – many Chinese words were borrowed into Vietnamese. When Vietnamese scholars started writing their own language, it was easy to see how to write these borrowed words: just use the same character that was used when writing Chinese. The following table gives a number of such characters used to write borrowings from Chinese into Vietnamese, their pronunciations in Vietnamese, and their approximate translation in English:

天 **thiên** sky; heaven; god

上 **thuong** top; highest; go up

工 **gông** labor; work; laborer

南 **nam** south

病 **bênh** illness; sickness

沖 **trong** pour; infuse; wash out

年 **nên** year; person's age

糸 **mich** silk

人 **nhân** man; human; mankind

木 **môc** tree; lumber; wood; wooden

見 **kiên** see, observe, perceive

告 **cáo** tell; announce; inform; accuse

弄 **lòng** do; play or fiddle with; alley

豆 **đâu** peas; beans

沐 **múc** bathe; cleanse; wash

心 **tâm** heart; mind; intelligence; soul

皮 **bì, bê'** skin; hide; fur

.

In order to write native Vietnamese words, however, these writers had to invent new characters. They did this by using a strategy that was already used, within the Chinese writing system, for creating new characters out of existing characters. In the Chinese writing system, new characters can be made by combining two or more simpler characters. These components provide hints regarding either the meaning of a character or its pronunciation. Components may be stacked on top of one another, place beside one another, or even placed so one surrounds another. While most of the characters given above are simple characters (with only one component) a few are complex characters (with more than one component). The resulting writing system is called **Chữ Nôm**. It was through the spread of this Chữ Nôm that Vietnamese literature finally came into its own. In fact, the Vietnamese national epic, the Tale of Kieu, was composed in Chữ Nôm.

Here is a translation of the first six lines of the Tale of Kieu in English (i-vi). Beneath it, but out of order, are the same lines in Vietnamese, both in Chữ Nôm (a-f) and in Quốc Ngữ.(1-6).

4.1. (11 points) Show on the answer sheet which of the Vietnamese lines corresponds to each line of English; one correspondence is already entered to get you started.

4.2. (4 points) Explain your answer.

English translations

A hundred years—in this life span on earth   (i)
    talent and destiny are apt to feud.   (ii)
You must go through a play of ebb and flow   (iii)
    and watch such things that make you sick at heart.   (iv)
Is it so strange that losses balance gains?   (v)
    Blue Heaven's habit is to strike a rose from spite.   (vi)

Quốc Ngữ

1. Trải qua một cuộc bể dâu,

2. Trăm năm, trong cõi người ta,

3. Trời xanh quen thói má hồng đánh ghen.

4. Những điều trông thấy mà đau đớn lòng;

5. Chữ tài, chữ mệnh, khéo là ghét nhau.

6. Lạ gì bỉ sắc, tư phong,

Chữ Nôm

a. 孨才孨命窖羅恄饒

b. 邏之彼蔷私豐

c. 歪青慣退 媽紅打悭

d. 仍調齟𥾒罵忉疸悳

e. 庪戈没局波橵

f. 聂䢔𩵜堁貁嗟

5. (20 points) Possessed in Vanuatu

Vanuatu is a South Pacific country with 74 populated islands and more than 100 languages belonging to the Oceanic language family made up of languages spoken from Papua New Guinea to Hawaii to Easter Island. In Vanuatu, speakers of these languages have developed interesting ways of saying that something belongs to someone. You are invited to examine some examples from the language spoken on the island of Tanna. Take a look at the examples of how possession is expressed in this language (given on the next page) and then answer the questions which follow.

Pronunciation notes: [ə] represents a sound like the last sound of 'the' in 'the book'. [ŋ] represents a velar nasal sound as in 'singing'.

1	ralah neŋow	their canoe
2	rahan nasumien	his garden
3	raham nima	your house
4	nepikə kahaw	rat's tail
5	nəməm nəkawə	your (sg) kava (a drink)
6	netetamlaw	your child (speaking to mother and father of child)
7	niŋlaw nahwel	their laplap pudding (a food) (for both of them)
8	nenien raha enteni	Tanna's speech (*enteni* 'earth' = Tanna)
9	ratah naŋhatien	our language (=yours (sg) and mine)
10	narmen	his image
11	rahak nien	my coconut (that I'm selling)
12	rahak sot	my shirt
13	narfu tem	man's belly
14	neiwok mil	my two female cousins
15	pukah asoli	big pig
16	niŋək nien	my coconut (for eating)
17	nelkak	my leg
18	piam	your (sg) same-sex sibling
19	nisiməteliŋəm	your (sg) ear-wax
20	narunien raha tjotam	Jotham's knowledge
21	niŋlah kuri	their dog (for eating)
22	niŋən nawanien	his food

23	nepiken	his tail
24	ratalaw jow	their turtle *(belonging to both of them)*
25	rahak jerehi	my lobster
26	nisin	his excrement
27	nentowi jow	turtle's neck
28	nerow raha jow	the turtle's spear
29	nelka pukah	pig's leg
30	nakale naw mil	the two edges of the knife OR the two knives' edges
31	nisi kunget	louse excrement
32	nəmtalaw nəkawə, ian mwamnəm	As for the kava (drink) belonging to you two, go and drink it!
33	ratamlaw kuri ije?	Where is your dog (belonging to the two of you)?
34	niŋək kuri u, ojakawan	My dog here, I'm going to eat (it).
35	rahak nima takaku	My house is small

5.1. (5 points) Translate each of these five expressions into the Tanna language.
   a. rat's ear
   b. my two dogs (that I own)
   c. their bellies
   d. a brother of those two (men)
   e. our child (= yours (sg) and mine)

5.2. (5 points) Translate these five expressions into the Tanna language.
   a. Jawkelpi's house
   b. the pig's canoe
   c. my picture of you (=the one that I own that is an image of you)
   d. The house belonging to you two is big
   e. Where is my lobster (that I am going to eat)?

5.3. (5 points) There are several ways of saying 'their' in Tanna. On the answer sheet, list those found in the Tanna examples and explain any differences in the meaning they express.