



OzCLO

Australian Computational and
Linguistics Olympiad
National Round 2012

Marking Guide

A. Gone fishing (1/3)

[20 points]

Waanyi is an Australian language traditionally spoken south of the Gulf of Carpentaria in country straddling the border between the state of Queensland and the Northern Territory. Few fluent speakers remain and our knowledge of this language now relies mainly on recordings made between the 1960's and 2008. Study this story and see how much Waanyi you can learn so you are able to answer the questions that follow.

1	Karrinja nyulu kirriya barrawunu.	<i>The woman is standing in the house.</i>
2	Jungku nyulu burrurri kundana.	<i>The man is sitting under a tree.</i>
3	Jungku bula nawunu rajini.	<i>They are here in the camp.</i>
4	Dabarraba nyulu waliji, nangkani burrurrii.	<i>This man is cooking meat.</i>
5	Balikajba nyulu, walijiyanyi, nana kirriya.	<i>She is hungry for meat, that woman.</i>
6	Nayi burrurri, lalujbu nyulu.	<i>This man, he gets up.</i>
7	Kanungku barri nyulu jilaba kirriyawurru.	<i>He then goes up to the woman.</i>
8	Wijbi barri nyulu kirriya walijiyanyi jangkaranyiyanyi, karrinjawurru.	<i>Then he gives some cooked meat to the woman who's standing.</i>
9	Nanangkani kirriyaa, nanganja barri nyulu manii nana waliji burrurrinanja.	<i>That woman, she then takes that meat with her hand from the man.</i>
10	Jarrba barri nyulu, balikajini, nanangkani kirriyaa, nana waliji, karrinjana nanawunu barrawunu.	<i>Then that woman hungrily eats that meat, standing there in the house.</i>
11	Jawikajba barri nyulu burrurri: Ninji, wanyi ninji jarrba?	<i>She then asks the man. What are you eating?</i>
12	Budangku ngawu jarrba jalanya.	<i>I'm not eating now.</i>
13	Jilakanyi ngawu kakuwanyi nanganjaanyi. Karubu-yanyba ngawu.	<i>I'll go and catch some fish. I'm going fishing.</i>
14	Wunjuku ninji jilaba?	<i>Where are you going?</i>
15	Kularra ngawu jilaba, nanangkurru manangkawurru.	<i>I'm going south, to that river.</i>
16	Ngabungabu, malijibi nyulu kirriyaa, banjana nyulu jilaba.	<i>Late afternoon, the woman followed him, she went after.</i>
17	Najba barri nyulu, burrurri, jungkuwurru, karubu-yanykurru.	<i>Then she saw the man sitting fishing.</i>
18	Manangkana nyulu jungku, nana burrurri.	<i>That man was sitting by the river.</i>
19	Najba nyulu kirriya, kanungkuwurru.	<i>He saw the woman approaching.</i>
20	Kawa! Jilanji nangkurru.	<i>Come! Walk over here!</i>
21	Jawikajba nyulu burrurri kanungkunu.	<i>She asked the man as she approached.</i>
22	Kaku ninji nanganja?	<i>Have you caught any fish?</i>
	(Continued on next page)	

23	Budangku ngawu kakuwanyi.	<i>I've got no fish.</i>
24	Budangku nayi kakuwanyi.	<i>There's no fish here.</i>
25	Ngamuyu-kiya ninji nanganja kaku nawunu. Kaja.	<i>I thought you would have caught fish here. Lots.</i>
26	Yanyba nyulu nangangi.	<i>He said to her:</i>
27	Najba ngawu kaku nawunu wanamini, bilikijawurru, bungkuna.	<i>I saw fish swimming here in the water yesterday.</i>
28	Budangku yalu balikajba walijiyanyi jalanya.	<i>They are not hungry for meat right now.</i>
29	Ngadijbi yaluwangka bulinjana.	<i>They are hiding in the water-grass.</i>
30	Rajiwurru barri bula kannga, budangku kakuwanyi.	<i>They both returned home, without any fish.</i>
31	Balikajini bula kannga rajiwurru, kirriya, bururri.	<i>They both return home hungry – the woman (and) the man.</i>

A-1 From the Waanyi text you will see that there are several words that translate as *here* in English. Complete the sentences below writing the appropriate 'here' word in the space indicated, then translate your completed sentence into English. [6] 0.5 / correct answer in both columns (Disregard tense of translation – can be present or past) (Accept other translations if they express same meaning.)

	Waanyi	English
a.	<i>Jungku bula nawunu.</i>	<i>They are (both) (sitting) here.</i>
b.	<i>Jilaba ngawu nangkurru .</i>	<i>I come/came/walk/walked (to) here.</i>
c.	<i>Budangku nayi kundaanyi.</i>	<i>There's no trees here. / Here are no trees / No trees are here</i>

A-2 Translate these Waanyi sentences into English: [3] 1 point each (Disregard tense of translation of 1 and 2.) Need to get correct meaning to get marks, with all elements of meaning included in sentence. Again, students might use different words in their English translations.

1. Jungku ngawu rajini.

I'm (sitting/staying) at home. OR I'm (sitting/staying) in camp. OR I was.../ sat/stayed...

2. Jawikajba barri bula nayi bururri.

They (both) then asked/questioned this man. OR They are both now asking/questioning
....

3. Budangku ngawu balikajba jalanya.

I'm not hungry now/today/at present.

A-3 Translate these English sentences into Waanyi: **[8]** (2 marks / sentence) must have correct subject pronoun (underlined) in second position to get full mark. All constituent orders are possible except that underlined must be in second position. Constituents between square brackets.

4. The man and the woman are sitting here.

[Jungku] bula [nawunu] [bururri, kirriya].

5. That woman eats fish.

[Jarrba] nyulu [nanangkani kirriya] [kaku]. (OR [kirriya nanangkani])

6. This man cooks that meat standing near a tree.

[Dabarraba] nyulu [nana waliji] [nangkani bururrii] [kundana karrnjana]. OR [bururrii nangkani] or [karrnjana kundana]

7. She gives the man fish.

[Wijbi] nyulu [bururri] [kakuwanyi]. (OR [kakuu] in place of [kakuwanyi])

A-4 Explain your answers to A-3 here. **[3]** Explanations will vary (use discretion).

to get full marks need to include something about

second position of subject pronoun or auxiliary – must follow first word OR first word + barri – otherwise free word order

use of endings on words (or different forms of words) to indicate different roles of the word/its referent (e.g., as agent or location etc)

Need to express the subject pronoun as well as the noun(s) with the same referent

If they explain that transitive subject (however expressed) is marked on nouns by lengthening final vowel – give 1 point.

They may also include:

something about vowel harmony on locative endings, e.g., *kunda-na*, *karrinja-na* vs *barrawu-nu* (line 1 & 10), *raji-ni* (line 3): vowel of suffix = final vowel of word/stem.

Here is Richard Hudson's explanation which may help:

This is an exercise in exploring the structure of a language with very different grammatical rules from English. The real challenges lie in the syntax, as I explain below.

- Finding the lexical correspondences (i.e. the dictionary words or vocabulary) is relatively easy, especially when you recognise that some words contain a long suffix (e.g. *kirriyawurru* in 7 contains *kirriya* in 1). Fortunately the morphology is very straightforward,

so endings are easy to recognise; but pay attention to small differences such as that between *kirriya* in 1 and *kirriyaa* in 9.

- The first syntactic problem concerns words like *nyulu* and *bula*, which don't seem to correspond systematically to anything in the English. Almost every sentence contains one of these words (or some other bisyllabic non-vocabulary word: *yalu*, *nayi*, *ninji*, *ngawu*), so you might suspect (rightly) that they're like auxiliary verbs; so let's call them auxiliaries. But what distinction do these verbs indicate? You might consider alternatives such as tense or positive/negative, but it turns out in fact that it's the 'person' and 'number' of the sentence's 'actor' (more on this below): *ngawu* = 'I', *ninji* = 'you', *bula* = 'the two of them', *yalu* = 'they', *nyula* = 'he/she/it'. (*Nayi* in 24 is a puzzle; maybe it just means 'there isn't'.)
- The second syntactic problem is word order. One regularity is that the auxiliary is very often the second word in the sentence, especially if you ignore on the one hand introductory material separated from the rest of the sentence by a comma (e.g. 6), and on the other the word *barri*, which means 'then'. With these two exclusions, there are no exceptions: the auxiliary always follows just one word. Moreover, that one word is very often the main verb of the sentence – but not always. For instance, 31 starts *rajiwurru barri bula*, 'they both returned then', but 32 moves the verb *rajiwurru* after the auxiliary: *Balikajini bula kannga rajiwurru*, 'hungry they both return'. Apart from these two rules, the order of elements in a sentence is free.
- The third problem is that nouns have more than one form which varies from sentence to sentence; for instance, 'woman' is *kirriya* in 1 and 5, but *kirriyaa* in 9 and 10. In some examples the extra material seems to express the same meaning as an English preposition; for example, *kirriyawurru* is translated as 'to the woman' (combined with *kanungku*, translated as 'approaching' in 19 and 21). But this doesn't help with the variation between *kirriya* and *kirriyaa*, or the similar alternation between *burrurri* and *burrurrii* for 'man' in 2 and 4. You may think it's like the 'case' changes found in languages like German (or in English pronouns, such as *he* and *him*), but that contrasts grammatical subjects and objects (e.g. *he came*, and *she saw him*). This pattern is called a 'nominative-accusative' case system, where 'nominative' and 'accusative' are the traditional names for the cases used as subjects and objects. That's not what's happening in Waanyi, where the short form is used both as subject in 1 or 2 (X is standing/sitting ...) and also as object in 17 and 19 (... saw X). Waanyi follows a completely different system, called a 'nominative-ergative' case system, where a special form (in Waanyi, the longer one) is used for the 'actor' in a two-part action, where the actor does something to someone or something else – in other words, for the subject of a verb that also has an object. For instance, 'woman' is normally *kirriya*, but in 9 it's *kirriyaa* because it means 'that woman takes that meat ...' and likewise in 10 meaning 'that woman eats that meat'.

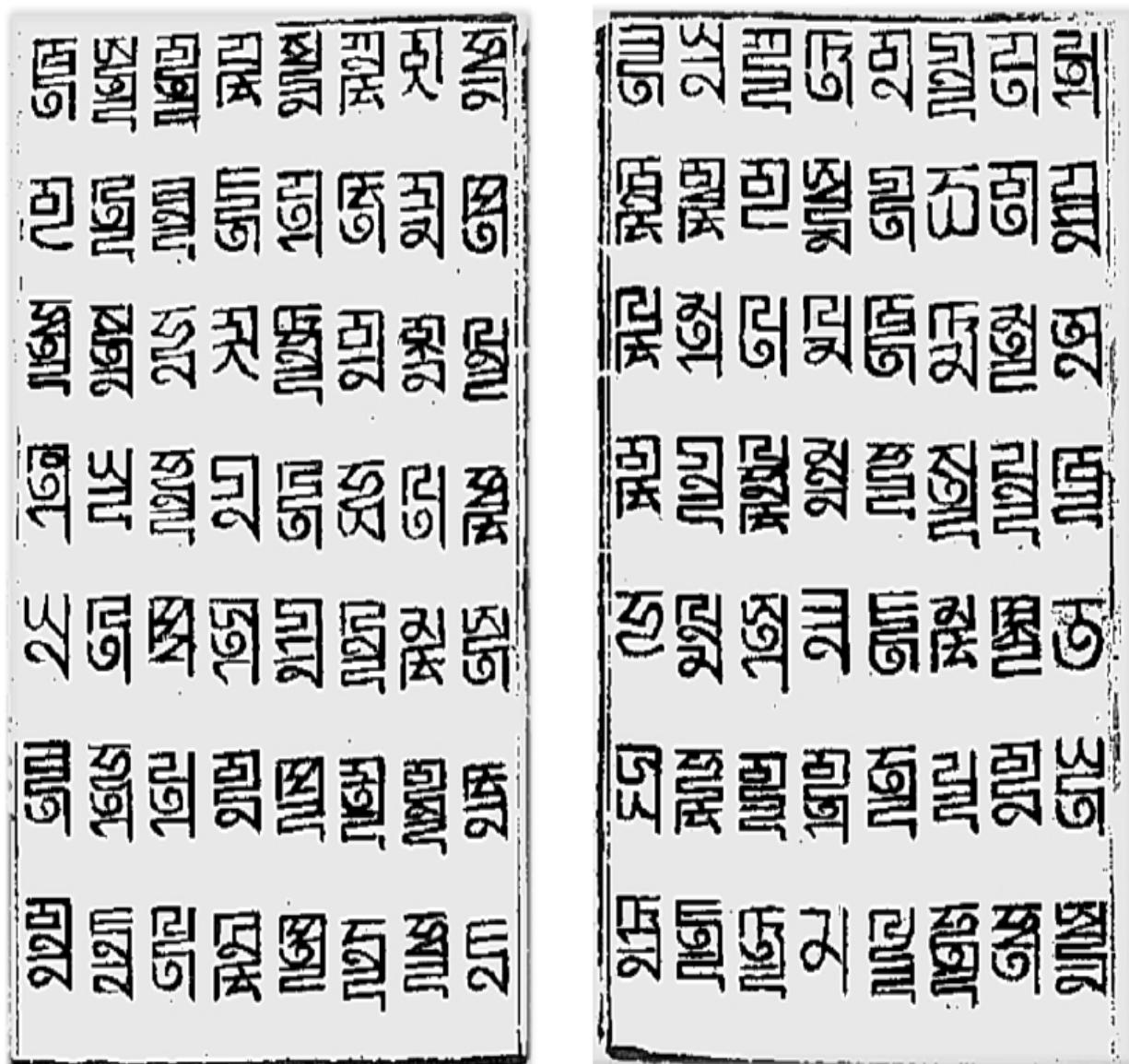
B. 100 Surnames (1/4)

[30 points]

When the Mongol Emperor Kublai Khan initiated the Yuan dynasty (1271–1368 A.D.) in China, he commissioned Lama 'Gro-mgon Chos-rgyal 'Phags-paa to create a unified script to write all the major languages under his rule. Although the resulting system (now called 'Phags-pa) never caught on beyond official use, some classic Chinese texts survive in a 'Phags-pa version.

The Bǎijiāxìng (Hundred Surnames) is a Song Dynasty (960–1279) poem listing over 400 classical Chinese family names. Although originally written in Chinese characters, during the Yuan dynasty this poem was written in 'Phags-pa characters as well.

Figure 1: Two consecutive pages of the Bǎijiāxìng Měnggǔwén (The Hundred Surnames in Mongol Script), from a 1340 A.D. manuscript.



In Figure 2 are twenty lines (9-28) from the Yuan-era Bǎijiāxìng, with some names missing. The two pages in Figure 1 correspond to a portion of the poem below. Your task is to work out which

portion of this poem the pages in Figure 1 represent, and use this to work out what the missing names must be.

Figure 2: Yuan-era Bāijiāxìng (fragment)

	a	b	c	d	e	f	g	h
9	Fi	Lem	Drxim	Sia,	Lue	Ho	Yi	Thang
10	Dxing	'In	Lo	Pi,	Haǔ	'U	'An	Srang
11	Yaǔ	Yiu	Sri	Fu,	Bue	Pen	Dzi	Khang
12	U	Yiu	Ngiǔan	Pu,	Ku	Mung	Bing	Hōang
13	_____	Fu	Sring	Tai,	Dam	Sung	_____	Bang
14	Xiung	Ki	_____	Khiu,	_____	Triu	Tung	Liang
15	Du	_____	Lam	_____	Zi	_____	_____	Giang
16	Kia	Lu	Lxiǔ	Ngue,	Kiang	Dung	_____	Kuaǔ
17	_____	Sring	Lim	Xiaǔ,	Trung	Ziu	Khiǔ	Laǔ
18	Kaǔ	_____	Tshai	Den,	Fan	Hu	_____	Fau
19	Ngiu	Wan	Tri	Ko,	_____	Koan	Lu	Maǔ
20	Kiing	_____	_____	Wu,	Kan	Xiai	'Ing	Tsung
21	Ting	Siuān	Pue	Dxing,	'Iu	Sren	Hang	Hung
22	Paǔ	Triu	_____	Sri,	Tshue	Kii	Nriū	Kiung
23	Dring	Xii	Xiing	_____	Bue	Liu	Ngiung	'Ung
24	Siuān	Yang	_____	Xiue,	Trin	Khieu	Kia	Fung
25	Nyue	Yi	Driu	Kin,	Ki	Ping	Mue	Ziung
26	Tsing	Dōan	Fuǔ	Wu,	'U	Tsiaǔ	Pa	Kiung
27	Wu	Ngue	Sran	Ku,	Trhia	Hiǔ	Fu	Bung
28	Dziǔan	Trhi	Pan	Ngiāng,	Tshiǔ	Driung	Yi	Kiung

This transcription represents Yuan dynasty pronunciation rather than modern pronunciation.

- *r* indicates that the previous sound is pronounced with an r-like curve of the tongue, and *h* indicates that the previous sound is pronounced with an extra puff of breath.
- *ny* is pronounced as in *canyon*, *ng* as in *sing*.
- The apostrophe sign (') represents a glottal stop – the sound in the middle of “uh-oh”.
- *x* indicates something like a whispered *y* or a *hy* sound.
- A *^* over a vowel means that it is a “glide” – a short vowel-like sound transitioning into or out of the syllable’s main vowel.

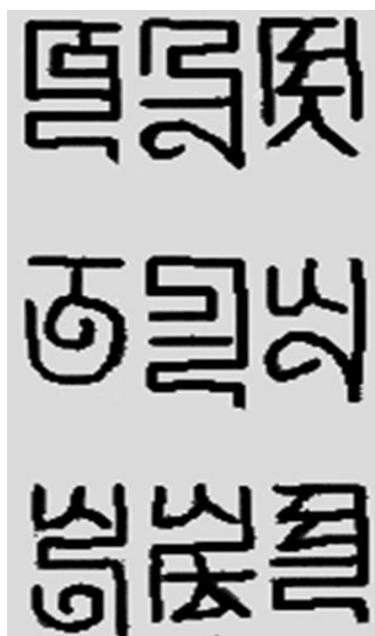
B-1 Which cell in Figure 2 corresponds to the topmost, rightmost name in Figure 1? (Answer by giving cell coordinates and the name written in that cell.) [2]

CELL 11E NAME Bue

B-2 Eighteen names are missing from Figure 2. Write them in the table below. [18] 1 x each correct answer (remove 0.5 for missing accent OR for one wrong/missing letter)

13a.	Kǐi	15f.	Kǐue	19e.	Tsam
13g.	Maǔ	15g.	Ma	20b.	Fang
14c.	Srǐu	16g.	Yan	20c.	Giǔ
14e.	Xǐang	17a.	Mue	22c.	Tso
15b.	Ngiǔan	18b.	Xǐa	23d.	Hǔa
15d.	Min	18g.	Ling	24c.	‘U

B-3 Here is a partial 3x3 excerpt from one page of a 1418 A.D. manuscript of the Bǎijiāxìng Měnggǔwén. Six of the names have been left out. Draw them in the spaces provided. [6] 1 point each missing name glyph.



B-4 Briefly explain how the *Phags-pa* writing system works. [4] (Use discretion with this one)

Need to have at least 4 of these elements in answer: [See explanation below]

1. Glyphs consist of subparts/subglyphs
2. Sub-glyphs represent parts of names
3. Direction of writing
4. (Four) different classes of glyphs
5. Same sound may be represented by different glyph depending on its position in word
6. No shape for /a/; missing V symbol indicates that vowel is /a/
7. When [ü] and [i] co-occur their order is reversed.

FYI I've attached the explanation written by a creator of this problem, highlighting parts corresponding to above.

To begin, we can notice that, given the complexity of the glyphs and the frequent recurrence of sub-glyph shapes within them, that the **glyphs likely consist of multiple sub-parts**.

We can then notice that the diversity of shapes at the top of the glyphs is much greater than the shapes at the bottom: there are a few dozen possible "tops" and only a handle of possible "bottoms". This, when compared to the names in the table, shows that the top shapes of glyphs represent the possible beginnings of names, of which there are many, and the bottoms represent the ends, of which there are few. This imbalance helps to confirm our earlier insight: that **sub-glyphs represent sub-parts of names in some systematic manner**. It also gives us a hypothesis about the overall writing direction of the text: that it might be top-to-bottom overall as well.

At this point, there are several ways to try to fix what part of the poem these two pages represent. One way is by rhyme scheme: given that the poem is made of 8-name rhyming couplets, there should be a pattern of repeating "bottom" shapes every eight characters. Given the 8x7 shape of the pages, if the text is written horizontally (in either direction), this pattern should occur in vertical lines, and if the text is written vertically (in either direction), this pattern should occur in diagonal lines. It is easy to see that there are no such vertical-line patterns (especially not at the ends of lines where we would hope them to be). There is, however, an every-eight-characters diagonal rhyme pattern running through the text: from 3rd row/1st column ("3a") upward and rightward to (1c), wrapping around to 7e and then up to 1k, wrapping around again to 7m, then up to 4p.

Going by the bottom shape, the rhyme scheme of the segment pictured here looks to be something like AAAAABBBBAAAAAA. There is only one place in the poem this could be: the BBBB glyphs have to represent Kűau, Lau, Faű, and Maű and the rest represent names ending in -ng. This result, however, is still compatible with two writing directions (upward then left-to-right, or downward then right-to-left). The former of these is unlikely given the downward writing direction within glyphs, but it is still at least possible. Looking at the "top" shapes decides it for us: we have an ABCDEFAGCHGIDJ pattern, and this only fits one way with the pattern of name beginnings in the text (**the downward then right-to-left direction**).

This is one way of determining the position and direction of the manuscript relative to the poem; other ways are equally valid and can receive equal points.

Now that we know which glyphs represent which, we can begin to determine what exactly the sub-glyph shapes represent. If you've come this far, this process will be mostly straightforward. There are four "classes" of glyphs, which I'll call A, B, C, and D:

- A: initial sounds (representing word-initial b, p, dz, kh, y, tr, tsh, etc.)
- B: "on-glides" (i, ü, etc.), written after initials

-- C. vowels (i, u, e, etc.), written after on-glides (if any)

-- D. codas (m, ng, ī, etc.), written last

Depending on where in a word a sound occurs, a sound like [ng] or [ū] might have to be in a different class and thus have a different shape.

One complication to watch out for is that there is no shape for [a]. If there is no other class C (vowel) glyph in the word, that word's vowel is [a]. You can tell the difference between (say) Húa and Haū, even though [a] is not written, by the shape [ū] gets: the class B or class D shape.

The other complication to note is that when class B [ū] and [ī] co-occur, they are written in ‘Phags-pa in an order opposite from what we would expect.

C. Roasted Red Potato Pancake (1/2)

[15 points]

English has the wonderful feature that it lets you stick two nouns together to make a **compound noun**, whose meaning derives in some idiosyncratic way from the meanings of its parts:

- water *fountain*: a fountain that *supplies* water
- water *ballet*: a ballet that *takes place in* water
- water *meter*: a device (called *meter*) that *measures* water
- water *barometer*: a barometer that *uses* water instead of mercury (to measure air pressure)
- water *biscuit*: a biscuit that is *made with* water
- water *glass*: a glass that is *meant to hold* water

Even more fun is that one of the two nouns in the compound noun could itself be a compound noun, as in the case of *china tea pot*. But what are we talking about? It depends. You make a [[*china* [*tea pot*]] out of fine porcelain called *china* because of its country of origin, whereas you use, maybe exclusively, a [[*china tea*] *pot*] to brew tea that either comes from China or which is of a style that is grown in China.

C-1 The paragraph above used **[square brackets]** to distinguish two possible meanings of *china tea pot*, one of them being the conventional (probably most obvious) meaning.

Add brackets to each compound below to indicate whether the most likely meaning corresponds to [[X Y] Z] or [X [Y Z]]. **[7]** (**1 point each correct**)

- a. [[ice cream] soda]
- b. [[science fiction] writer]
- c. [[customer service] representative]
- d. [state [chess tournament]]
- e. [[Mars Rover] landing]
- f. [plastic [water cooler]]
- g. [[typeface design] report]

C-2 Choose the most likely bracketing for the 4-word compound noun *country song platinum album*. Indicate your choice by ticking the box to the right of the chosen compound. **[1]**

- a. [country [song [platinum album]]]
- b. [country [[song platinum] album]]
- c. [[country song] [platinum album]]
- d. [[country [song platinum]] album]
- e. [[[country song] platinum] album]

C-3 Give a plausible definition of [[*space mission*] [[*control freak*] *show*]]. (If you must use compound nouns in your definition, define them too.) **[3]**

A drama about control freaks (i.e., freaks about control), performed during a space mission (i.e., a mission to space). (or other plausible definition – might depend on how ‘show’ is interpreted)

C-4 Show the most likely bracketing for the 8-noun sequence below. As in the examples above, your bracketing must have the form [X Y], where each of X and Y is either a single-word noun or a compound noun (which must also be written as a bracketing [X Y] and so on.) [2]
for expected answer, 1 for other plausible answer.

Expected answer is:

[[[[family [board game]] [togetherness effect]] [government study]] author]. [2]

Although the following might also be defensible:

[[[[family [[board game] togetherness]] effect] [government study]] author] [1]

Or even perhaps:

[[[[[family [board game]] togetherness] effect] [government study]] author] [1]

C-5 Paraphrase below the super8 compound in **C-4** to make clear the meaning expressed by your bracketing analysis above (in **C-4**): [2] full marks if definition matches bracketing analysis in **C-4**.

[[[[family [board game]] [togetherness effect]] [government study]] author].

= The author of a government study (study commissioned by the government/ done for the govt) on/of the effect on togetherness of board games (games played on boards) played by families/ or suitable for families or whose target audience is families.

Although the following might also be defensible:

[[[[family [[board game] togetherness]] effect] [government study]] author]

= author of a government study on the effect on families of board game togetherness

[[[[[family [board game]] togetherness] effect] [government study]] author]

= The author of a government study on the effect of family board game togetherness (togetherness associated with board games played/owned by families).

Markers: please use your own intuitions and judgments here!

D. A fox among the h (1/2)

[10 points]

Dr. Dumutche is compiling an online biology reference, and he is currently working on the information retrieval system, so that people can type in questions like “What do whales eat?” or “How much does a bee weigh?” and get relevant answers.

Part of this task involves a process called *stemming* – taking text and figuring out what the “stem” of each word is. (The “stem” is the form of the word without any prefixes or suffixes, so *dance* is the stem of *dancing*, *happy* is the stem of *unhappiness*, etc.). The system needs this so that it can determine that a request about “whales” needs data from the article WHALE and one about “fungi” needs data from the article FUNGUS.

So, Dr. Dumutche writes a series of rules for determining the singular form of plural nouns. He writes a rule, “Remove final S” to handle *whales*→*whale*. He writes another rule, “Replace final I with US” to handle *fungi*→*fungus* and a rule “Remove final E” to handle *algae*→*alga*, plus some other rules to handle other types of plural words and their singular counterparts.

He ends up with the following seven rules:

- | | |
|--------------------------------------|----------------------------------|
| 1 Remove final S | 5 Remove final EN |
| 2 Replace final ICE with OUSE | 6 Replace final A with UM |
| 3 Replace final IES with Y | 7 Replace final I with US |
| 4 Remove final E | |

When he applies his little program to a series of real words, however, it doesn’t always work.

Here is the output of his program:

Input	Intended Output	Actual Output
cats	cat	cat
dogs	dog	dog
walruses	walrus	walrus
foxes	fox	fox
oxen	ox	ox
bacteria	bacterium	bacterium
fungi	fungus	fungus
horses	horse	hors
chimpanzees	chimpanzee	chimpanze
algae	alga	algum
guppies	guppy	guppi
hens	hen	h
mice	mouse	mous

D-I What singular form would Dumutche’s program produce for the following words: [2.5] (0.5 point each correct)

Input	Actual Output
bees	be
kiwis	kiwus
flies	fli
fleas	fleum
geese	gees

D-2 What went wrong with Dr. Dumutche's program? [2] Need to mention reapplication of rules with an example of the type below to get full marks. The observation in curly brackets {but it doesn't apply a same rule twice} is not required to get full marks.

Dumutche's program doesn't stop applying rules when one of them succeeds – it keeps looking for applicable rules and then applies them to the output of the previous rules {but it doesn't apply a same rule twice}. (

This gives the right output for "walruses" and "foxes" – it removes the "s", then continues on and removes the "e" – but goes very wrong with "horses", "hens", etc.

An answer such as "D's program can't distinguish correct from incorrect words" only gets 0.5.

D-3 What can you determine about the order in which Dr. Dumutche's program applied the rules? [4] 0.5 for each of these – only give 0.25 if no example given.

We can determine that:

"Remove S" must come before "Remove E"; otherwise, we would get WALRUSE, FOXE, MOU, etc. instead.

"Remove S" must come before "Remove EN"; if it came after, we would get HEN instead.

"Remove S" must come before "Replace IES with Y"; if it came after, we would get GUPPY instead.

"Remove S" must come before "Replace I with US"; if it came after, we would get FUNGU instead.

"Replace I with US" must come before "Remove E"; if it came after, we would get GUPPUS instead.

"Remove E" must come before "Replace A with UM"; if it came after, we would get ALGA instead.

"Remove E" must come before "Replace OUSE with ICE"; if it came after, we would get MIC instead.

PLUS [0.5] for observation that:

Rules do not apply twice – that is, Dumutche's program probably goes through each rule in the list exactly once and only goes through the list once. Otherwise, we would get things like CHIMPANZ or WALRU, etc.

(if this has already been stated in answer to D-2, but not repeated here, then still give additional 0.5 here)

D-4 Could putting the rules in a different order cause the program to work? What order would produce the intended results? Or why isn't there one? (You can refer to rules by their

corresponding number or name.) **[1.5 point]**

Need to justify answer with ONE relevant example along the lines of answer below to get full mark. Only give 0.5 for an answer like: "There is no order of rules that will make Dumutche's program work" if there is no justification – i.e. no answer to "Why isn't there one?"

There is no one order of rules that will make Dumutche's program work, for several reasons:

1. The feeding of "Remove E" by "Remove S" is necessary to get "walrus" and "fox" correct, but it's exactly this interaction that produces "hors" and "chimpanzee".
2. No order will correctly produce "mouse". Consider the two rules (A) "Remove E" and (B) "Replace ICE with OUSE". If A comes before B, we get "mic"; if B comes before A, we get "mous" (or even "mou").

E. Ik heb voorspeld (1/2)

[15 points]

Here are some examples of regular Dutch verbs in their infinitive or plain form and their past participle form; for instance, *slibben* means **to silt up**, and its past participle *geslibd* means **silted up**, as in “It has silted up”. The English meaning is given for information only: it has no bearing on the solution.

Table I: Some Dutch verb forms

Verb	Translation	Past participle
slibben	<i>to silt up</i>	geslibd
klagen	<i>to complain</i>	geklaagd
branden	<i>to burn</i>	gebrand
weren	<i>to resist</i>	geweerd
tochten	<i>to make a draft (wind)</i>	getocht
tellen	<i>to count</i>	geteld
raken	<i>to hit (target)</i>	geraakt
lijmen	<i>to glue</i>	gelijmd
kunnen	<i>can, to be able</i>	gekund
vertellen	<i>to tell</i>	verteld
telen	<i>to tell</i>	geteeld
verhoren	<i>to cultivate</i>	verhoord
trouwen	<i>to interrogate</i>	getrouwed
schaven	<i>to marry</i>	geschaafd
razen	<i>to shave (woodwork)</i>	geraasd
prijzen	<i>to storm</i>	geprijsd
lappen	<i>to put a price on</i>	gelapt
smaken	<i>to clean</i>	gesmaakt
praten	<i>to taste</i>	gepraat
fietsen	<i>to cycle</i>	gefietst
boffen	<i>to be lucky</i>	geboft

E-1: Write the past participle form for each of the following verbs under its English translation.

[10] 1 point each correct answer.

1 gedeeld	2 gehoord	3 getapt	4 verhuisd	5 geland
6 geklopt	7 gemokt	8 gerookt	9 gerot	10 getobd

E-2: In E-1 you were asked to predict (or derive) the past participle from the plain form. But doing it the other way round, i.e. deriving the plain form from its past participle, is not always possible. Give one reason why. Illustrate your answer with one of the examples given in Table I or in the Table in E-1. [5] (3 points for the explanation; 2 points for a relevant example, i.e. one of those in Table below)

Past participles ending in -d or -t may or may not have had a d or t in the plain form (like *branden* and *tochten*). (part in parenthesis not essential to get full mark of [3] for this part of Q)

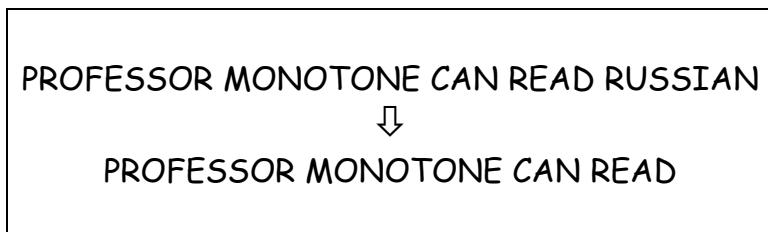
So any one of the following examples would demonstrate this: (following linguistic convention, we mark with an asterisk hypothetical unattested forms).

Past participle	Actual plain form	Might have been
geslibd	slibben	*slibden
geklaagd	klagen	*klaagden
geweerd	weren	*weerden
geteld	tellen	*telden
geraakt	raken	*raakten
gelijmd	lijmen	*lijmden
gekund	kunnen	*kunden
verteld	vertellen	*vertelden
geteeld	telen	*teelden
verhoord	verhoren	*verhoorden
getrouw'd	trouwen	*trouwden
geschaaf'd	schaven	*schaafden
geraasd	razen	*raasden
geprijsd	prijzen	*prijsden
gelapt	lappen	*lapten
gesmaakt	smaken	*smaakten
gepraat	praten	*praatten
gefietst	fietsen	*fietsten
geboft	boffen	*boften

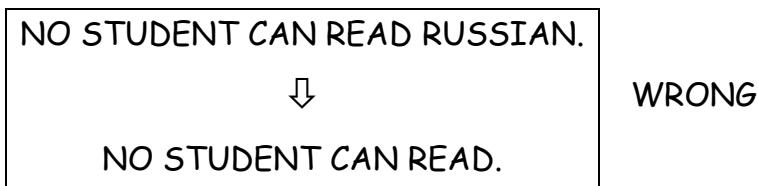
F. The Little Engine that Can... Read (1/2)

[10 points]

Professor Monotone's "Astounding Linguistic Knowledge Engine for Making Inferences" (ALKEMI), when given a list of true statements, can deduce further true statements from it. For example, if it knows that "Professor Monotone can read Russian", it can deduce that "Professor Monotone can read". We represent this as:



This means that whenever the first statement is true, the second has to be true, too; there's no way for the first to be true while the second is false. We call this a *legitimate inference*. The Professor's machine can go through statements and, by making particular sorts of changes, generate further statements that follow from them. However, it's not as easy as replacing "can read Russian" with "can read" anywhere you find it. For example, funny things happen when the statement contains one of a set of words called "quantifiers", including *every*, *some*, *no*, *a*, *few*, *many*, *three*, and so on.



The inference is not legitimate; even if no student reads Russian, it's entirely possible that they read Japanese, English or Spanish.

Each quantifier allows a different pattern of legitimate inferences, so the professor's machine keeps a special table of patterns and uses it to derive new statements from given ones. We've reproduced it on the next page. It may look mysterious, but given the information in this table and a list of inferences produced by the machine, you can work out what each part means and how the machine works.

Figure 1: Inference patterns used by Monotone's Machine

	Quantifier	Side	Direction
A	Every	Left	Downward
B	Every	Right	Upward
C	No	Left	Upward
D	No	Right	Downward
E	Some	Left	Upward
F	Some	Right	Upward

Unfortunately, however, there is one error in Figure 1 that is causing the professor's machine to draw some illegitimate inferences!

Figure 2. Some inferences declared legitimate by Monotone's Machine:

Every teacher can read English. ↓ Every English teacher can read English.	No student can read Russian. ↓ No student can read English and Russian.
Some English students can read English. ↓ Some English students can read.	Every teacher can read English. ↓ Every Russian teacher can read English.
No English student can read Russian. ↓ No student can read Russian.	Some Russian students can read English. ↓ Some students can read English.
Every teacher can read English and Russian. ↓ Every teacher can read Russian.	No English student can read. ↓ No English student can read English.

F-1 Which row (A-F) in Figure 1 contains a mistake and caused the machine to draw one or more illegitimate inferences? (Answer by writing appropriate letter in box below.) [2]

C (The illegitimate inference is —No English student can read Russian => No student can read Russian. We certainly hope that Russian students, for example, can read Russian!)

F-2 The list of inferences isn't complete. Monotone's Machine could draw additional inferences as well. Using only words that appear in Figure 2, generate another legitimate inference that the machine could have drawn from "Every teacher can read English". [2]

Every teacher can read.

F-3 Monotone's Machine doesn't yet understand every quantifier. Help it learn the quantifiers *at least three*, *at most three*, and *not all* by putting "Upward" or "Downward" in the appropriate cells: [6] (1 point each)

	Quantifier	Side	Direction
G	At least three	Left	upward
H	At least three	Right	upward
I	At most three	Left	downward
J	At most three	Right	downward
K	Not all	Left	upward
L	Not all	Right	downward

See more explanation on next page:

The terms “left” and “right” determine which of the two “arguments” of the quantifier (the first one or the second one, respectively) the rule applies to.

No English student reads English



The terms “upward” or “downward” indicate whether a *more general* or *more specific* replacement is warranted. (So “English teacher” is more specific than “teacher”, “reads English” is more specific than “reads”, “reads English and Russian” is more specific than “reads English”, etc.)

So a “left upward” quantifier allows the replacement of the “left” argument of the quantifier with a more general phrase, and a “right downward” quantifier allows the replacement of the “right” argument with a more specific phrase.