



ROUND ONE

February 2013
2 hours

PROBLEM BOOK

*Answer as many of the questions as you can.
Write your answers in the answer books provided.*

Questions and credits

A. Many Bulgarians	Bozhidar Bozhanov	10 marks
B. A case of Pali	Babette Verhoeven	15 marks
C. Intuitive Inuit	Harold Somers	20 marks
D. Talk like Yoda you must	Harold Somers	15 marks
E. Grammar rules OK	Andrea Schalley/Patrick Littell	20 marks
F. Phoenician fun	Harold Somers	20 marks

Question A: Many Bulgarians (10 marks)

Here are listed some Bulgarian noun phrases in the form quantifier (numeral or “many”) + noun.

Note: Bulgarian is normally written in the Cyrillic alphabet. For this problem, all Bulgarian words have been transliterated. As for pronunciation, *č* is like the “ch” in “chest”, *š* is like the “sh” in “shoe”, *c* sounds like “ts”, and *ē* is a short vowel like the “a” in “patrol”. Other letters are pretty much like in English.

<i>dvama učenici</i>	two students
<i>devet garderoba</i>	nine wardrobes
<i>mnogo učenici</i>	many students
<i>edin sanděk</i>	one chest
<i>tri sanděka</i>	three chests
<i>mnogo sanděci</i>	many chests
<i>devetima balěci</i>	nine morons
<i>mnogo garderobi</i>	many wardrobes
<i>šestima programisti</i>	six programmers
<i>četiri kapaka</i>	four covers
<i>mnogo programisti</i>	many programmers
<i>trima čistači</i>	three cleaners
<i>edin balěk</i>	one moron

In your answer book, write in Bulgarian (using the same transliteration scheme):

- six covers
- many morons
- four cleaners
- many covers
- one student
- two chests

Question B: A case of Pali (15 marks)

Pali is a dead language, like Latin. It was a literary language related to Sanskrit, the ancestor of modern languages spoken in Northern India, such as Hindi; it is also distantly related to English. Pali was first written down around 100 BCE in Sri Lanka by Buddhist monks to preserve the teachings of the Buddha, and is usually written in a special script (which we replace here by our familiar Roman letters, using *ā* and *ī* for long vowels and without capital letters or punctuation). Here are some sentences or phrases in Pali with their English translations:

Pali	English Translation
<i>mahāmatto nisīdati</i>	The minister sits down.
<i>mahāmattam upasamkamanti</i>	They visit the minister.
<i>samano tathāgato hoti</i>	The philosopher is enlightened.
<i>samane attham pucchanti</i>	They ask the philosophers the meaning.
<i>upāsako pucchati</i>	The disciple asks.
<i>loko mahāmattassa</i>	the minister's world

Here are a few more items of vocabulary:
rājo 'king', *devo* 'god', *gāmo* 'village'

B1 Translate the following into English: (4 marks)

- a. *rājo nisīdati*
- b. *rājo gāmassa devo hoti*

B2 Translate the following into Pali: (11 marks)

- a. The minister asks the philosophers.
- b. The philosopher sits down.
- c. They sit down.
- d. The minister asks the kings.
- e. the disciple's village
- f. The meaning of the world is god.

Question C: Intuitive Inuit (20 marks)

The Inuit live in the Arctic regions of Canada and Greenland, and speak a language called Inuktitut. The Inuit used to be known as Eskimos, but this term is now considered insulting. The writing system used for the Inuktitut language is based on the one devised for writing Cree, a Native American language not related to Inuktitut. The writing system is highly regular and systematic, which should make your task all the easier.

The area of Canada where Inuktitut is spoken is called Nunavut and its capital is Iqaluit. Here is how these two words are written:

Nunavut	ᓃᓇᓂᓪ
Iqaluit	ᐃᓪᓃᓪᓂᓪ

English has borrowed some words from the Inuit, such as 'igloo' which in Inuktitut is written as ᐃᓪᓃ and is pronounced *ihlu*: the *hl* is a lateral fricative, like the LL sound in Welsh, a bit like an 's' combined with an 'l', also found in Zulu and some other languages. The *hl* sequence is treated as a single sound.

Question D: Talk like Yoda you must (15 marks)

In the *Star Wars* movies, the character Yoda was heard to speak English with a distinctive word order, apparently a deliberate device to make him seem different and special. Here are some examples of Yoda's English, all genuine examples from the screenplays of the *Star Wars* movies:

Take you to him I will.
 A domain of evil it is.
 Help them you could.
 Always two there are.
 Truly wonderful the mind of a child is.
 Much to learn you have.
 When nine hundred years old you reach, look as good you will not.
 Lost a planet Master Obi-Wan has.
 Begun the Clone Wars have.
 Your father he is.
 Unexpected this is.
 Hiding in the Outer Rim Grievous is.
 Go I will. Good relations with the Wookies I have.
 Chewbacca and Tarfful, miss you I will. Good friends you are.
 For your help much gratitude and respect I have.

However, the script writers are far from consistent. Sometimes Yoda uses ordinary English word order, such as "But now, we must eat" or alternatives which, although slightly unusual, are still acceptable in normal English, such as "A powerful Jedi was he" or "Sick have I become". Your task is to take the following examples (all are genuinely from the *Star Wars* movies) and suggest how they should be spoken in 'pure' Yoda-speak.

- a. But now we must eat.
- b. I cannot teach him.
- c. The boy has no patience.
- d. This one a long time have I watched.
- e. A Jedi's strength flows from the Force.
- f. The fear of loss is a path to the dark side.
- g. Great care we must take.
- h. Disturbing is this move by Chancellor Palpatine.
- i. The capture of General Grievous will end this war.
- j. To a dark place this line of thought will carry us.
- k. A little more knowledge might light our way.
- l. Sick have I become.
- m. Twilight is upon me and soon night must fall.
- n. Into exile I must go.
- o. Your weapons you will not need them.

Question E: Grammar rules OK (20 marks)

One way for computers to understand language or for linguists to describe language is to define the language using a “context-free grammar” (CFG) (also called a “phrase-structure grammar”). A CFG is a set of rules for forming sentences. Only sentences that can be “generated” using such a set of rules are then deemed grammatically correct and “well-formed”. The “language” defined by the CFG is any and all sentences that a given CFG can generate. S is the starting symbol for each sentence.

Here is an example of a simple CFG:

S → N + V
 N → children
 N → squirrels
 V → sing
 V → eat

Each rule says that the element to the left of the arrow can be expanded into (or replaced by) the element(s) to the right of the arrow. By repeatedly replacing symbols, this CFG can expand the symbol S into “squirrels sing”, “children sing”, “squirrels eat”, and “children eat”. It cannot, however, generate “children eat squirrels” or “squirrels eat children” or just “children” – you can see that there is no possible sequence of replacements that turns S into any of these.

The following is another fairly simple CFG. The rules have been numbered for your convenience, but the numbers are not part of the rules.

1. S → NP + VP	2. NP → N	3. NP → D + N
4. NP → NP + CONJ + NP	5. VP → VP + PP	6. VP → VP + CONJ + VP
7. VP → IV	8. VP → IV + PP	9. VP → TV + NP
10. VP → TV + C + S	11. PP → P + NP	12. PP → P
13. IV → runs	14. IV → sits	15. TV → eats
16. TV → catches	17. TV → tells	18. TV → sees
19. TV → chases	20. CONJ → and	21. P → away
22. P → in	23. D → the	24. C → that
25. N → squirrel	26. N → dog	27. N → John
28. N → Mary	29. N → tree	30. N → he

E1. Here is a simple story. Not all the following sentences are, according to the CFG, well formed, meaning they cannot be derived from S by repeated substitution of symbols. In your answer book mark each sentence with a tick or a cross to indicate whether they are well-formed (✓) or not (✗).

(16 marks)

- John sees the dog and Mary sees the dog.
- The dog sees John and Mary.
- The dog sees a squirrel.
- The squirrel sits in the tree.

- e. That squirrel sees the dog.
- f. The squirrel is seen by the dog.
- g. The dog runs.
- h. The squirrel in the tree runs.
- i. The dog chases the squirrel and eats the squirrel.
- j. The dog eats.
- k. John sees that the dog eats the squirrel.
- l. John tells Mary that the dog eats the squirrel.
- m. The dog sees that John sees that he eats the squirrel.
- n. And the dog runs away.
- o. Mary and John chase the dog.
- p. John chases and catches the dog.
- q. John eats dog.

E2. Not all of the sentences that this CFG can generate are actually sentences of English. For example, “The dog and the squirrel sits” can be generated but this isn’t a correct sentence of English. Give three more examples of sentences that can be generated by this CFG but are not correct English sentences. (3 marks)

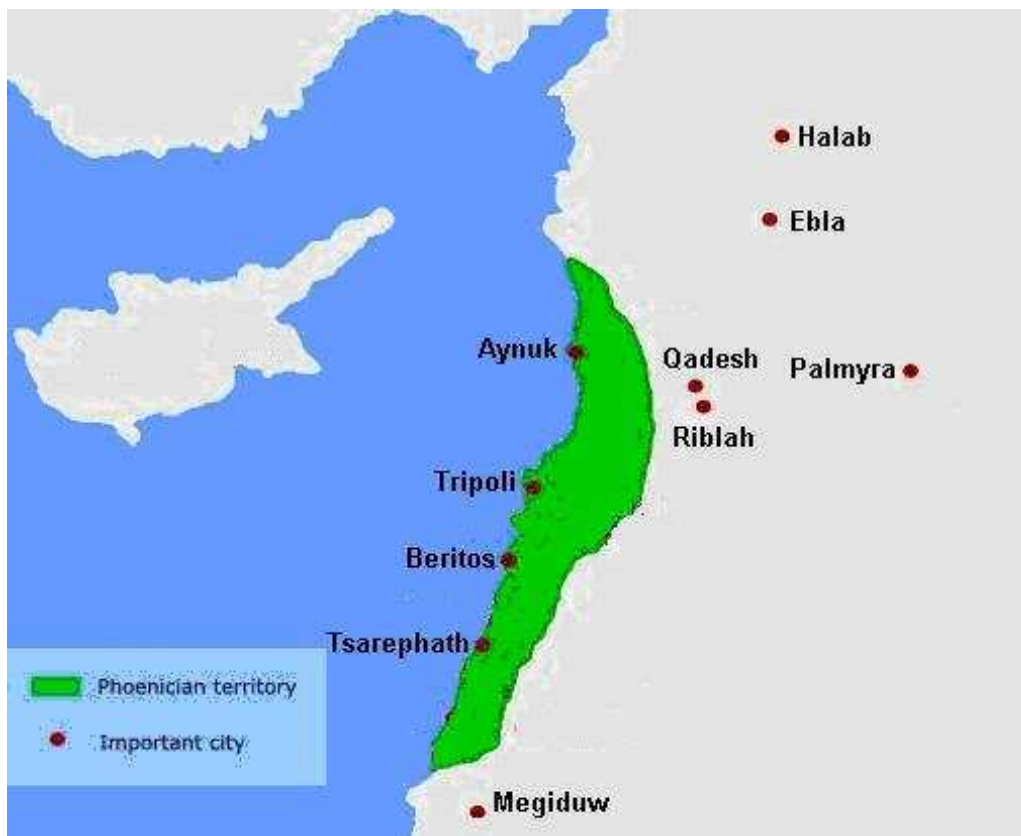
E3. One of the rules in the CFG above is redundant: any sequence of words that can be generated by this rule can already be generated by a combination of other rules. Which is the redundant rule? (1 mark)

Question F: Phoenician fun (20 marks)

The Phoenicians have a special place in the history of linguistics, since their language was one of the first to be written in a phonetic script, unlike other contemporary writing systems which were pictorial. The Phoenician script can be dated at around 1050 BCE, and from it the Arabic, Hebrew and by extension Greek, Roman and Cyrillic scripts evolved.

The Phoenician civilization was centred along the Mediterranean coast in an area known as Cana’an. The map below shows a number of Phoenician cities and some nearby cities that were important trading partners. The spellings reflect their pronunciation in Phoenician. However, two of the cities on the map are shown with their modern names which are very different from what they were called in Phoenician times.

(See next page)



This map is copyright © mapsof.net and is reproduced under the Creative Commons Attribution-ShareAlike 1.0 Licence

F1. Match up the Phoenician names in the list below with the names on the map. Remember, two of the names will not match, so you should have two names left over. (15 marks)

- | | | | |
|---|---------|---|---------|
| A | ⊕ 7 4 ʳ | F | ✕ ʳ ʳ ʳ |
| B | ʳ ʳ ʳ | G | ʳ ʳ ʳ |
| C | ʳ ʳ ʳ | H | ʳ ⊕ ʳ |
| D | ʳ ʳ ʳ | I | ʳ ʳ ʳ |
| E | ʳ ʳ ʳ ⊕ | J | ʳ ʳ ʳ |

F2. Write down how the names of the two left-over cities were pronounced in Phoenician, as far as you can tell. Briefly explain what is missing and why. (1½ marks each + 2 for explanation)