OL LO MALTA

Linguistics Olympiad

The Problem Solvers' Challenge

PRACTICE SESSION 2

Give yourselves around ½ hour per puzzle

SOLUTIONS

Answer as many of the questions as you can: read the instructions carefully.

Use pen and paper to note down your observations as you work towards.

Puzzles

1. Free the Friulian three   Dick Hudson & Harold Somers
2. Georgia's always on my mind   Harold Somers
3. Transitions   Harold Somers
4. Basquing in the sun   Aleka Blackwell
5. Making a mark in Choctaw   Babette Newsome

Puzzle 1: Free the Friulian three   (15 points)

Friulian is sometimes referred to as a dialect of Italian, but it is sufficiently different to be classed as a separate language, spoken by around 600,000 people in Northeast Italy.

Study the following Friulian phrases, then fill in the blanks in the table below.

my house	*la mê cjase*
the houses	*lis cjases*
the bicycle	*la biciclete*
your bicycles	*lis tôs bicicletes*
our tickets	*i nestris tagongs*
my job	*il gno lavôr*
my sister	*mê sûr*
our brother	*nestri fradi*
your cousin	*to cusin*
my sisters	*lis mês sûrs*
our Irish grandmother	*la nestre ave irlandese*
my cousins	*i miei cusins*
our Irish cousin	*i nestri cusin irlandês*

(a) our house	**la nester cjase**	(f) your sister	**tô sûr**
(b) the ticket	**il tagong**	(g) our cousins	**i nestris cusins**
(c) my bicycle	**la mê biciclete**	(h) our jobs	**i nestris lavôrs**
(d) your brother	**to fradi**	(i) the ticket	**lis nestres aves**
(e) my Irish sister	**la mê sûr irlandese**	(j) my brother	**gno fradi**

Explanation

Like in most Romance languages, Friulian nouns belong to one of two genders, masculine and feminine. The masculine nouns in the data are *tagong*, *lavôr*, *fradi* and *cusin*; the feminine nouns are *cjase*, *biciclete*, *sûr* and *ave*. As far as we can see, the plural of all nouns is formed by adding *–s*. All adjectives, articles and possessives agree in both gender and number. So these words have four forms as follows (only the forms given in the data are shown):

Word	masc sing	fem sing	masc plur	fem plur
the	*il*	*la*	*i*	*lis*
my	*gno*	*mê*	*miei*	*mês*
your	*to*	*?*		*tôs*
our	*nestri*	*nestre*	*nestris*	*?*
Irish	*irlandês*	*irlandese*		

For the answers, it was mostly a question of looking up the appropriate form in the above table. However (f) and (i) required you to "construct" the feminine singular for 'your' and the feminine plural for 'our' respectively. It is reasonable to assume that the plural is the same as the singular, but with an –s added, as is the case with nouns and some of the possessive pronouns. Of course that's not a hard and fast rule (cf. 'the' and 'my'), but if 'your' and 'our' were not regular, would it be fair to ask you to guess?

But there's something else going on: why do most but not all of the phrases with possessives also have an article. Let's look at the examples we have:

	la mê cjase
	lis tôs bicicletes
	i nestris tagongs
	il gno lavôr
mê sûr	*lis mês sûrs*
nestri fradi	*la nestre ave irlandese*
to cusin	*i miei cusins*
	i nestri cusin irlandês

The rule – and don't ask why this is – is that in general the article is included, EXCEPT for a family member in the singular and not otherwise modified. So 'our aunt' would be *nestre ave*, but 'our Irish aunt', with the additional modifier 'Irish' reinstates the article *la*.

Puzzle 2: Georgia's always on my mind (20 points)

Tbilisi is the capital of Georgia, a country in the Caucasus region of Eurasia.

On the next page is a list of (some of) the stations on its subway (underground) system, written in Georgian, together with the equivalent English names of the stations, though not in the same order. While most of the names are simply transliterated, a few of them are translated and so do not match exactly. Note that **J** is pronounced like the 'y' in 'yes', and that sequences DZ, SH, TS are regarded as single letters.

The Tbilisi metro system – don't worry if you can't read the names on this map, they are reproduced below.

1	ვარკეთილი	A	Aragveli
2	სამგორი	B	Avlabari
3	ისანი	C	Didube
4	არაგველი	D	Gotsiridze
5	ავლაბარი	E	Guramishvili
6	სადგურის მოედანი	F	Isani
7	რუსთაველი	G	Marjanishvili
8	მარჯანიშვილი	H	Medical University
9	ნაძალადევი	I	Nadzaladevi
10	დიდუბე	J	Rustaveli
11	გურამიშვილი	K	Samgori
12	სარაჯიშვილი	L	Sarajishvili
13	ტექნიკური უნივერსიტეტი	M	Station Square
14	სამედიცინო უნივერსიტეტი	N	Technical University
15	გოცირიძე	O	Varketili

B1. Match up the station names 1-15 with their equivalents A-O. [15 points]

Add the letter corresponding to the equivalent English station names of each of the Georgian station names 1-15 below.					
1	**O**	2	**K**	3	**F**
4	**A**	5	**B**	6	**M**
7	**J**	8	**G**	9	**I**
10	**C**	11	**E**	12	**L**
13	**N**	14	**H**	15	**D**

B2. How is the station name translated here as 'Station Square' pronounced in Georgian? [3 points] **sadguris moedani**

B3. How is the Georgian word for 'Medical' pronounced? [2 points] **sameditsino**

Explanation

Although this may look complex because of the unfamiliar alphabet, it is actually quite a simple decoding question, once you realise that Georgian is a straightforward alphabet with consonants and vowels much like ours.. There are several "entry points" to the code, any one of which will get you started, after which you just have to patiently fill in the letters.

One entry point would be the fact that there are two stations with "University" in their name, and two stations ending in უნივერსიტეტი (two of only three with two-word names). So you might reasonably try to match up the letters. The Georgian is actually *universiteti*, but the first part of the word will give you 6 or 7 letters. Another "way in" is to look for similar names such as three ending in *–ishvili* matching up with იშვილი. You could also count letters: *Isani* is the only 5-letter name, *Didube* the only 6-letter name. Or you could use the picture, the title of which, you might guess, says *Tblisis Metro*. How do you know which university is which? One is *sameditsino*, the other *texnikali*. To be fair, the *x* (pronounced like Irish/German/Welsh *ch*) doesn't appear anywhere else. But it's not hard to guess is it?

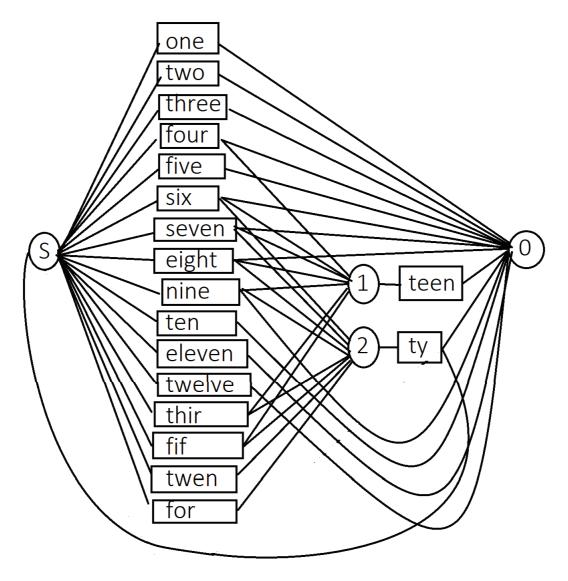Here is the Georgian alphabet: not all the letters appeared in the data:

ა ბ გ დ ე ვ ზ თ ი კ ლ მ ნ ო პ ჟ რ ს ტ უ ფ ქ ღ ყ შ ჩ ც ძ წ ჭ ხ ჯ
a b g d e v z t i k l m n o p zh r s t u p x gh q sh ch ts dz ts'ch kh j

Actually there are five more letters which are not used in modern Georgian.

Puzzle 3: Transitions   (20 points)

The diagram on the next page represents a "transition network". The circles represent "states" while the boxes represent sequences of letters that can be "generated" from any given state, as indicated by the lines (the "transitions"). The aim is to start at "S" and get to the end state "0". For some boxes there is a choice of transition. The lines are directional (it would have been even more messy to add the arrowheads), so note that you can only enter a state from the front (left). For example, you can go from "ty" to "S", but not the other direction. As you can see, by following the paths you can generate the names of some numbers in English.

The above diagram is already quite messy, and it can be represented more neatly by a set of rules as below. Each rule is identified (in square brackets) but this is ONLY for ease of reference in answering the questions. Apart from that each rule consists of a state (the symbol before the ":"), a sequence of letters, and then, after the arrow ("→") a list of states to which you can then move. Starting at position "S", you generate the text indicated, and then continue to any ONE of the rules whose start state is listed after the arrow. State "0" is a special case meaning "finish".

[a]	S: one → 0		[j]	S: ten → 0
[b]	S: two → 0		[k]	S: eleven → 0
[c]	S: three → 0		[l]	S: twelve → 0
[d]	S: four → 0,1		[m]	S: thir → 1,2
[e]	S: five → 0		[n]	S: fif → 1,2
[f]	S: six → 0,1,2		[o]	S: twen → 2
[g]	S: seven → 0,1,2		[p]	S: for → 2
[h]	S: eight → 0,1,2		[q]	1: teen → 0
[i]	S: nine → 0,1,2		[r]	2: ty → S,0

So, for example, starting at S we can generate "fourteen" by taking rule [d] to state 1, then rule [q] to finish. We cannot generate "twelveteen" because rule [l] only allows one way to progress, namely to finish.

C1. Write out the sequence of rules and states followed to generate the following words: For example, for "fourteen" write "d 1 q 0". [12 points]

- (a) sixteen
- (b) ninetythree
- (c) twentyeight
- (d) fifteen

C1. Solution

- (a) **(S) f 1 q 0**
- (b) **(S) i 2 r S c 0**
- (c) **(S) o 2 r S h 0**
- (d) **(S) n 1 q 0**

C2. The network above "overgenerates", that is, it allows us to create sequences which are not valid number names. Indicate whether each of the following words can be generated by the network or not. [4 points]

- (a) oneten
- (b) fiftytwelve
- (c) sixteensix
- (d) twentyseventeen
- (e) fortythirty
- (f) eleventythree
- (g) fivety

C2. Solution – Put a tick (✓) in the box if the network generates this string**.**

(a)	(b)	(c)	(d)	(e)	(f)	(g)
	✓		✓	✓		

C3. The above network wrongly generates a misspelling in the case of "eighteen" and "eighty...". Suggest a simple fix for this (i.e. a change to one of the existing rules and an additional rule). [4 points]

C3. Solution

Change rule h as follows: S: eight → 0
Additional rule S : eigh → 1,2

(Some other fixes are possible, but this is the most elegant and obvious)

Puzzle 4: Basquing in the sun   (25 points)

Basque is a language spoken by around 700,000 people in an area straddling France and Spain, in the western Pyrenees. Linguistically, Basque is unrelated to the other languages of Europe and indeed, is a "language isolate", unrelated to any other known language.

Study the following sentences, shown in Basque together with their translations in English.

Basque	English
Emakumeak gizona ikusi du.	The woman has seen the man.
Zuk umea ikusi duzu kalean.	You have seen the child in the street.
Non ikusi duzu umea?	Where have you seen the child?
Nork ikusi du umea kalean?	Who has seen the child in the street?
Umeak ez du gizona ikusi.	The child has not seen the man.
Emakumeak liburua irakurri du.	The woman has read the book.
Umea etorri da.	The child has come/arrived.
Umea kalean erori da.	The child has fallen in the street.
Igela agertu da.	The frog has appeared.

D1. Translate (a)-(d) into Basque and (e)-(f) into English in the table below: [16 points]

(a) The woman has come.
(b) The man has seen the woman.
(c) The book has not arrived.
(d) The street has appeared in the book.
(e) *Nork ikusi du gizona?*
(f) *Igela kalean agertu da.*

(a) The woman has come.	*Emakumea etorri da*
(b) The man has seen the woman.	*Gizonak emakumea ikusi du*
(c) The book has not arrived	*Liburua ez etorri da*
(d) The street has appeared in the book.	*Kalea liburuan agertu da*
(e) *Nork ikusi du gizona?*	Who has seen the man?
(f) *Igela kalean agertu da*	The frog has appeared in the street

Basque is a free word-order language. This means that all the phrases shown in Set A below translate into English as 'The child has fallen in the street.' (In English different word order would change the meaning, compare 'The man bit the dog' and 'The dog bit the man', but not so in Basque.)

SET
A   *Kalean umea erori da.*
       Umea erori da kalean.
       Kalean erori da umea.
       Erori da umea kalean.
       Erori da kalean umea.

By contrast, all the phrases in Set B below are considered ungrammatical. That is, a native speaker of Basque would consider these structures unacceptable. Following convention, we mark ungrammatical sentences with an asterisk *.

SET
B   * Erori umea da kalean.
       * Erori kalean da umea.
       * Erori umea kalean da.

6

D2. Briefly explain the restriction which applies to the word order options available in Basque. [3 points]

Now consider the following additional data.

(a) *Gizonak umeari liburua eman dio.*
    The man has given the book to the child.
(b) *Irakasleak umeari liburua irakurrarazi dio.*
    The teacher has made the child read the book.

D2. Solution

Erori **can't be separated from** *da*. **More generally, the word order restriction is that a MAIN VERB and its following associated AUXILIARY VERB have to appear adjacent to each other and in that order.**

D3. Translate (a) into English, and (b) into Basque in the table below. [6 points]

(a) *Emakumeak umeari etorrarazi dio*.
The woman has made the child come.
(b) The child has given the frog to the woman.
Umeak emakumeari igela eman dio.

Explanation

Basque nouns carry endings showing their function (what linguists call "case") in the sentence. But unlike more familiar (Indo-European) languages, the cases are not exactly subject/object. Basque is a so-called "ergative" language, which means that the case corresponding to subject in a transitive sentence (where there is also an object) is not used for what would be the subject of an intransitive sentence (with no object): in that case the marking is the same as the object. The best way to understand it is to think of sentences like *John opens the door* and *the door opens*: in familiar languages (that show case marking like German, Latin, Greek and so on) – door would be object in the first, subject in the second. But in an ergative language it is the same case in both sentences, which has some logic when you think that the door plays the same role in both sentences (it gets opened).

Linguists call this role "absolutive", while the "doer" or agent is called "ergative". And quite simply, the nouns in the ergative case add a –*k*. There is also another case-marker in the first set of example sentences,

which we can call locative –*n*. The endings occur on pronouns (*zuk*, *nork*, *non*) as well as nouns.

The verbs in the data mostly end in –*i*, but there is another word that needs explaining: *da*, *du*, *duzu*. This is an auxiliary verb (like *do* in English negatives, questions and past tenses) and takes the form *du* when there is an ergative (ending in –*k*), *da* when there is not, and *duzu* when 'you' is the ergative.

Word order is more or less free, so you could present the answers to D1 in any order, with the exception of *erori* which can't be separated from *da*.

Puzzle 5: Making a mark in Choctaw   (20 points)

Choctaw is a native American language spoken in Oklahoma and Mississippi by around 10,000 people.

Here are some Choctaw sentences with their English translations.

Baliililitok.	I ran.
Baliilitok.	He/she ran.
Baliililih.	I have run / I am running.
Johnat niyah.	John is fat.
Pamat kayyah.	Pam is pregnant.
Baliilih.	He/she has run *or* He/she is running.
Chikayyah.	You are pregnant.
Saniyah.	I am fat.
Hattakat chaahah.	The man is tall.
Hattakat taloowatok.	The man sang.
Hattak chaahah piisalitok.	I saw the tall man.
Hattak chaahah ishpiisatok.	You saw the tall man.

E1. Translate the following into Choctaw: [10 points]

(a) He/she saw John.	**John piisatok.**
(b) I sang.	**Taloowalitok.**
(c) You are singing.	**Ishtaloowah.**
(d) I am pregnant.	**Sakayyah.**
(e) He/she saw the fat man.	**Hattak niyah piisatok.**

Here are a few more sentences or phrases in Choctaw:

Sayyit sabashah.	My leg is cut.
Sapiisatok.	He/she saw me.
Issapiisatok.	You saw me.
amofi	my dog

Ofit sayyi sakopolitok.	The dog bit my leg.
Ofit amofi kopolitok.	The dog bit my dog.
Chimanolilitok.	I told you.
Amanolitok.	He/she told me.
Sashkit hattak piisatok.	My mother saw the man.

E2. Translate into English: [10 points]

(a) *Chiyyit chibashah.*	**Your leg is cut.**
(b) *chimofi*	**your dog**
(c) *Amofi ishpiisatok.*	**You saw my dog.**
(d) *Chishki piisalitok.*	**I saw your mother.**
(e) *Amofit chiyyi chikopolitok.*	**My dog bit your leg.**

Explanation

Word-order: the verb comes last.

Pronouns are unusual as the verb markings are sufficient to convey subject, object etc. Direct object nouns are unmarked, but subject nouns take the suffix *–at* if the noun ends in a consonant or *–t* if the noun ends in a vowel.

Person and number is marked on the verb, almost always as a prefix; only first person singular is marked as a suffix.
- The person /number markers on the verb indicate argument roles: subject, direct object, indirect object / dative case, benefactor etc.
- Any argument may appear OR if it is clear from context it is possible to leave person / number argument markers out altogether.
- The person / number markers that are suffixes come first, then follow the suffixes that indicate tense / aspect etc.
If a verb has multiple arguments that are all marked as prefixes, the Subject prefix comes first, then the Object prefix etc.
- The person / number argument markers that are used as affixes on verb stems are also affixed on adjectives to make a predicate (complement)

The table below breaks down the person / number argument markers used in this problem: hyphens indicate whether a prefix or suffix.

Person:	Class 1 argument (like Subject / Nominative):	Class 2 argument (like Object / Accusative &	Class 3 argument (like Indirect Object / Dative & alienable

		inalienable possession):	possession):
I	*-li*	*sa-*	*am-* *sam-* (when preceded by subject or object prefixes)
you (singular)	*ish-* *-is* (alternative if stem or following prefix begins with a sibilant)	*chi-*	*chim-*
he/she/it	unmarked	unmarked	not used in problem.

A note on tense & aspect marking on the verb:

- the verb takes the suffix *–tok* in what equates to simple past tense (e.g. "I sang") – a completed action that is now in the past;
- the verb takes suffx *–h* to indicate a predicate which implies a tense that can be translated into English as either a present perfect tense (as in "I have sung") or a progressive present tense (as in "I am singing") – an action that is still having an influence on the present / now, and that is taking up some time.
- there is no verb "to be" (the *–h* suffix on the adjective indicates that it is used predicatively as in sentence 4, where the adjective "niya" becomes "niyah").