## <1> Miwoc in California (1/3)                    12 points

Miwoc is a member of the Utian family of indigenous languages spoken in central and northern California. The language is seriously endangered: currently there are only about 3,500 Miwoks alive, and it is not known how many of them – if any – speak this language.

The data below lists 50 simple verb forms with their translations. The transcription has been simplified slightly. The č and š symbols represent "ch" and "sh" respectively, while x represents a glottal stop, the sound between the two syllables in "*uh-oh*". Miwoc distinguishes between singular and plural "you", the plural glossed here as "you all".

| | | | |
|---|---|---|---|
| *yilim* | I bite | *yilaputon* | They bit you all |
| *yiliš* | You bite | *kači* | He speaks |
| *yili* | He bites | *kačiton* | He speaks to you all |
| *yilimaš* | We bite | *kačimuš* | I speak to you |
| *yilitoš* | You all bite | *kačimu* | You speak to me |
| *yilip* | They bite | *kačitoknixmaš* | We speak to you all |
| *yilak* | I bit | *lotaknixmaš* | We caught you |
| *yilaš* | You bit | *lotakmutoš* | I caught you all |
| *yila* | He bit | *lotaktoknixmaš* | We caught you all |
| *yilakmaš* | We bit | *lotakmuč* | You all caught me |
| *yilaktoš* | You all bit | *lotaput* | They caught me |
| *yilap* | They bit | *lotakmušmeč* | You all caught us |
| *yilakmutoš* | I bit you all | *lotipum* | They catch us |
| *yilak* | I bit him | *xinam* | He came to us |
| *yilimaš* | We bite them | *xinapum* | They came to us |
| *yilakmu* | You bit me | *xinakmuš* | I came to you |
| *yilimuč* | You all bite me | *xinakmušme* | You came to us |
| *yilimušmeč* | You all bite us | *xiniš* | You come to him |
| *yilit* | He bites me | *xiniputon* | They come to you all |
| *yilat* | He bit me | *šiyičit* | He watches me |
| *yilim* | He bites us | *šiyičimušme* | You watch us |
| *yilin* | He bites you | *šiyičimutoš* | I watch you all |
| *yilaton* | He bit you all | *šiyičinixmaš* | We watch you |
| *yiliput* | They bite me | *šiyičapun* | They watched you |
| *yilipun* | They bite you | *šiyičakmuš* | I watched you |

# <1> Miwoc in California (2/3)

In general, the endings of each of these Miwoc words indicate information about who or what is doing something to whom or what, i.e., a combination of *subject* and *object* features — to use linguist speak. Before attempting the tasks assigned below, it may help you to fill out the following grid. Some cells are filled in as a guide, e.g., *nixmaš* on the end of a verb indicates that 'we' do something to 'you', while *put* indicates that 'they' do something to 'me', and *š* means that 'you' do something. HINT: In some cells you will have alternating endings, while the same ending may appear in more than one cell.

| | | A | | B | | C | |
|---|---|---|---|---|---|---|---|
| | **Subject →** | I | we | you | you all | he | they |
| | **Object ↓** | | | | | | |
| 1 | me | | | | | | *put* |
| | us | | | | | | |
| 2 | you | | *(k)nixmaš* | | | | |
| | you all | | | | | | |
| 3 | him | | | | | | |
| | them | | | | | | |
| 4 | (none) | | | *š* | | | |

## Task 1:

Write the five verb stems (i.e., minimal part of word expressing verb meaning without the Tense or Subject and Object endings) corresponding to *bite*, *speak (to)*, *catch*, *come (to)* and *watch*. (Write the Miwoc verb stem below the corresponding English verb.)

| bite | speak (to) | catch | come (to) | watch |
|---|---|---|---|---|
| | | | | |

## Task 2:

a. How is tense marked on Miwoc verbs? Answer by specifying the sound which marks *present* and *past* respectively. (Answer by completing each of the sentences below.)

Present tense is marked by _____

Past tense is marked by _____

b. Indicate where in the verb these sounds occur by marking the spot with 'X':

___ stem ___ subject-object ending ___

# <1> Miwoc in California (3/3)

**Task 3:**

Translate the following into Miwoc:

| | | |
|---|---|---|
| a. | *We spoke to you all.* | |
| b. | *He came to them.* | |
| c. | *You bit us.* | |
| d. | *They speak to us.* | |

**Task 4:**

Translate the following into English:

| | | |
|---|---|---|
| a. | *šiyičaton* | |
| b. | *lotimu* | |
| c. | *yilanimaš* | |
| d. | *kačan* | |

## <2> Come to Istanbul (1/2)      20 points

Turkish is spoken by about 63 million people, who mostly live in Turkey, but many Turkish speakers also live elsewhere. It is not an Indo-European language like English. It is related to languages of Central Asia such as Azeri and Uzbek.

Turkish is sometimes described as 'agglutinative' because its words are built up by adding one or more endings to a basic or 'root' word. As you will see, the vowels in most endings vary depending on the vowels in the root word — a phenomenon called 'vowel harmony'.

NOTES:
- The Turkish letters "ş", "ç" and "ı" are pronounced like English "sh", "ch" and the "a" in "above".
- The letters i and ı represent different vowels.
- The letters u and ü represent different vowels, rather like the difference between the vowel sound in "foot" as opposed to "food" or "boot".
- The letter "ğ" is usually silent (like the "gh" in "although").

Here are some Turkish expressions, with their English translations. (Square brackets [ ] enclose English words that are not directly translated.)

| | Turkish | English |
|---|---|---|
| 1. | Arkadaşlarım şehirde mutlu | My friends [are] happy in [the] city. |
| 2. | Baban İstanbul'u seviyor mu? | Does your father like Istanbul? |
| 3. | Fakirler Van'dan İstanbul'a gelmek istiyor | Poor [people] want to come from [the city of] Van to Istanbul. |
| 4. | İstanbul en büyük şehir | Istanbul [is the] biggest city. |
| 5. | Eve geliyorlar | They come home. |
| 6. | Babam "Merhaba! Gel, arkadaşımız ol", diyor. | My father says "Hello! Come [and] be our friend". |
| 7. | Evimizde büyük pencereler var | There are big windows in our house. |
| 8. | Pencereden atlıyoruz | We jump from [the] window. |
| 9. | Ev almak mı istiyorsun? | Do you want to buy [a] house? |

### Task 1:

An English word like *rebuilding* can be analysed as having a maximum of 3 meaningful parts, e.g., *re - build -ing.* We can roughly indicate the meaning of each part by using other words with same or similar meanings, e.g., *again - construct - action (of).*

Now break up these Turkish words into their maximum number of meaningful parts, using spaces and hyphens; indicate the meaning associated with each by writing a one word (if possible) indication of the meaning of each of these parts in English in the third column.

| Turkish word | Meaningful parts of word | Meaning of each part |
|---|---|---|
| *arkadaşlarım* | | |
| *evimde* | | |
| *pencerelerden* | | |

# <2> Come to Istanbul (1/2)

**Task 2:** Translate the following Turkish sentences into English?

| | |
|---|---|
| a. | Baban mutlu mu? |
| b. | "**Ş**ehirimize gel" diyoruz. |
| c. | Arkada**ş**ım doktor olmak istiyor. |
| d. | Fakir evimi seviyorlar mı? |
| e. | **İ**stanbul'dan mı geliyorsun? |

**Task 3:** The following examples introduce a new pattern. What do these examples mean?

| | |
|---|---|
| a. | Geldi**ğ**imde "merhaba" diyorlar. |
| b. | Baban geldi**ğ**imizden mutlu mu? |
| c. | Fakir oldu**ğ**unu diyorlar. |
| d. | Aldı**ğ**ın ev büyük mü? |
| e. | En mutlu oldu**ğ**um **ş**ehir, Van. |
| f. | Fakir oldu**ğ**umuz halde mutluyuz. |

## &lt;3&gt; Hidden in Hungary (1/2)                    20 points

The grid below represents a field divided into a 7 x 7 grid, aligned north-south and east-west. There are rocks, represented by **x,** in some of the squares of the grid.

There are four Hungarians standing in the field, but *not* on squares with rocks. Each one is facing in one of the four cardinal directions. Each person (1-4) makes three statements describing the positions of the rocks; for instance, the first person's first statement translates as "*To the east (behind me) there's one stone.*" The second person's first statement translates as "*To the south (to the left) there are no stones.*" (All compass points are to be understood as meaning "*due east/west/north/south".*)

**Task 1:** Find each person's place in the field and the direction (N, S, E or W) they're facing.

1. (a) Keletre (mögöttem) egy kő van.
   (b) Délre két kő van.
   (c) Jobbra nincs kő.

2. (a) Délre (balra) nincs kő.
   (b) Északra egy kő van.
   (c) Mögöttem két kő van.

3. (a) Északra (előttem) nincs kő.
   (b) Nyugatra egy kő van.
   (c) Jobbra két kő van.

4. (a) Nyugatra (jobbra) két kő van.
   (b) Északra egy kő van.
   (c) Balra nincs kő.

## <3> Hidden in Hungary (2/2)

| | | | X | | | |
|---|---|---|---|---|---|---|
| | | X | X | X | | |
| | | | | | | |
| X | | | X | | | X |
| | | X | | X | | |
| | | | | | | |
| | | | X | | | |

Complete this task by writing the number corresponding to each person (1-4) in the appropriate cell of the table above, and complete these sentences with the appropriate compass term:

Person 1 is facing _____

Person 2 is facing _____

Person 3 is facing _____

Person 4 is facing _____

**Task 2:**

Give the Hungarian equivalents of the English compass points:

north          _____

south          _____

east           _____

west           _____

## <4> Lontara (1/2)          15 points

Buginese is a language of the southern region of Sulawesi, the third largest island of Indonesia. It is written using a script called *Lontara.* The script's name comes from the Malay word for palm, *lontar*. The long, thin leaves of the palm were once used to create scroll-like manuscripts.

Here is a passage from the epic Indonesian creation myth *Sureq Galigo,* written in Buginese using the Lontara script. Text is thus usually written as one long line.

# <4> Lontara (2/2)

**Sureq Galigo: English translation**

*There is no one to call the gods Lord, or to offer praise to the underworld. Why Lord don't you have one of your children descend, and incarnate him on the earth; do not leave the world empty and the earth uninhabited. You are not a god, Lord, if there are no humans under the heavens, above the underworld, to call the gods Lord.*

**Sureq Galigo: Buginese written in the Latin alphabet**

The Buginese text of Sureq Galigo has been chopped up into several pieces and scrambled up. These lines are the same pieces of the same creation myth in Buginese, but written in the Latin alphabet instead of the Lontara script.

> *Notes about spelling: **e** is a vowel (/ə/)* pronounced like 'e' in *betray.* **é** is a vowel (/e/) pronounced like 'e' in *bed.* **q** is the glottal stop (/ʔ/) which is the sound in the middle of *uh-oh.*

**Task 1:**

Unscramble the pieces: figure out the correct order for the fragments A through to J, and write them in order here, one letter per blank. (NOTE: no capital letters are used here, reflecting their absence in the Lontara writing conventions.)

| A | ajaq naonro lobbang linoé | 1 | _____ |
| B | lé namasuaq mua na sia | 2 | _____ |
| C | makkatajangeng ri atawareng. | 3 | _____ |
| D | mappaleq wali ri pérétiwi. | 4 | _____ |
| E | mattampa puang lé ri batara, | 5 | _____ |
| F | mattampa puang lé ri batara. | 6 | _____ |
| G | ri awa langiq, lé ri ménéqna pérétiwié, | 7 | _____ |
| H | tabareq-bareq ri atawareng. | 8 | _____ |
| I | tammaga puang muloq séwa rijajiammu, | 9 | _____ |
| J | teddéwata iq, puang, rékkua masuaq tau | 10 | _____ |

**Task 2:**

Translate these English words into Buginese, using the **Latin script** only. You don't need to write them using the Lontara script.

lord            _____

underworld      _____

earth           _____

**Task 3:**

Returning to the original text in the Lontara script, circle the written form of *puang* each time it occurs.

## <5> Navajo's Great Chain of Being (1/3)          14 points

 Navajo (sometimes also "Navaho") is an American Indian language from the South-West United States. Famously, Navajo Code-Talkers (soldiers in US Army units fighting in the Pacific during the Second World War) were used to send secret army messages; this was the only "code" that the Japanese never managed to break (Japanese code breakers managed to break every artificial code invented by the US military).

Navajo (in common with all American Indian languages) is very different from any European or other well-known languages such as Arabic, Chinese or Japanese.

Here are some of Navajo's rules you'll be working with in this problem:

i. Subject (the person, animal, or thing doing the action) & object (the person, animal or thing having the action done to them) can go in either order in the sentence; compare this with English: "Man bites dog" means something completely different from "Dog bites man".

ii. However, there is an exception to the rule above: sometimes you're not allowed to vary the order of subject and object depending on the category of the nouns (i.e., words that indicate people, animals, places, or things).

**TASK 1:**
Study these Navajo sentences:

| Navajo Sentence | English Translation | Alternative Navajo Sentence |
|---|---|---|
| a   Diné ashkii yiztał. | The man kicked the boy. | Ashkii diné biztał. |
| b   Ashkii diné yiztał. | The boy kicked the man. | Dine ashkii biztał. |
| c   Ashkii łééchąąʔí yiztał. | The boy kicked the dog. | * Łééchąąʔí ashkii biztał. |
| d   *Dóola diné yizgoh. | The bull gored the man. | Diné dóola bizgoh. |
| e   Dóola shash yizgoh. | The bull gored the bear. | Shash doola bizgoh. |

**Important:**
*denotes ungrammatical sentence – a sentence that would be "wrong" in spoken Navajo.

ʔ and ł represent particular sounds in Navajo – you don't need to think about these for solving this problem.

a. Although the word order can vary in Navajo, one particular type of word always appears in the same location in Navajo.

i. What sort of word is this? _____

ii. Where does this word-type appear in the Navajo sentence? _____

b. This particular word-type is subject to a small change, depending on "who" or "what" is doing the action of the sentence to "whom" or "what". Explain in your own words what this change is and how it works.

# <5> Navajo's Great Chain of Being (2/3)

c. You'll have noticed that Navajo's tendency to categorize nouns has had an effect on the sentences a-e above. The nouns are categorized according to a particular hierarchy. Based on these sentences (a to e), explain the hierarchy at play here.

_____

_____

_____

**Task 2:**
Building on what you have learnt in Task 1 and by looking at the data below, work out the rules about noun hierarchy in Navajo in more detail.

| | Navajo Sentence | English Translation | Alternative Navajo Sentence |
|---|---|---|---|
| f | Shash mósí yishxash. | The bear bit the cat. | *Mósí shash bishxash. |
| g | *Mósí shash yishxash. | The cat bit the bear. | Shash mósí bishxash. |
| h | Mósí tązhii yinoołchééł. | The cat is chasing the turkey. | Tązhii mósí binoołchééł. |
| i | Mósí naʔazízí yinoołchééł. | The cat is chasing the gopher (small rodent). | *Naʔazízí mósí binoołchééł. |
| j | Naʔazízí wóláchíí yinoołchééł. | The gopher is chasing the ant. | *Wóláchíí naʔazízí binoołchééł. |
| k | *Dibé awééchíʔí yiztał. | The sheep kicked the baby. | Awééchíʔí dibé biztał. |
| l | Dibé tązhii yiztał. | The sheep kicked the turkey. | Tązhii dibé biztał. |
| m | Naʔazísí wóláchíí yiisxí. | The gopher killed the ant. | * Wóláchíí naʔazísí biisxí. |
| n | *Awééchíʔí diné yiztał. | The baby kicked the man. | Diné awééchíʔí biztał. |
| o | Shash awééchíʔí yinoołchééł. | The bear is chasing the baby. | Awééchíʔí shash binoołchééł. |
| p | Tsísʔná naʔashjéʔii yishish. | The bee stung the spider. | Naʔashjéʔii tsísʔná bishish. |
| q | Naʔashjéʔii wóláchíí yiisxí. | The spider killed the ant. | Wóláchíí naʔashjéʔii biisxí. |
| r | *Tsísʔná naʔastsʔǫǫsí yishish. | The bee stung the mouse. | Naʔastsʔǫǫsí t sísʔná bishish. |

The category at the upper end of the hierarchy is in the leftmost position in a precedence chain where categories are separated by the symbol ">" (A > B > C ...) where '>' means 'precedes'. List all the words (in English) found in sentences a-r that denote a member of each of these categories. NOTE: There is not enough information in the Navajo sentences above (a-r) to confidently classify *one* of the nouns used.

| A | > | B | > | C | > | D | > | E |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |

# <5> Navajo's Great Chain of Being (3/3)

**Task 3:**

a. Which noun were you unable to place in one of the five categories above? _____

b. What additional information would a linguist need in order to confidently classify this word? Answer this question by giving an English sentence that the linguist might ask a Navajo speaker to translate into Navajo which would allow the linguist to decide the category that the word identified in your answer to a. belongs to.

_____

c. Explain how the translation of your English sentence will give you the information you need.

_____

_____

_____

# <6> Combinatorial Categorial Grammar (1/4)     19 points

**PART I: ENGLISH**

One way for computers to understand language is by forming a structure that represents the relationships between words. One such approach is known as Combinatorial Categorial Grammar (CCG). Computer scientists and linguists can use CCG to parse sentences (that is, try to figure out their structures) and then extract meaning from the structures.

As the name suggests, Combinatorial Categorial Grammar parses sentences by combining categories.  Each word in a sentence is assigned a particular category:

| | |
|---|---|
| I | NP |
| books | NP |
| sleep | S \ NP |
| enjoy | (S \ NP) / NP |

These categories are then combined in systematic ways.  We will not explain how, but we will give you two successful parses…

| I | sleep |
|---|---|
| *NP* | *S \ NP* |

$$\text{S}$$

| I | enjoy | books |
|---|---|---|
| *NP* | *(S \ NP) / NP* | *NP* |

$$S \setminus NP$$
$$S$$

… and four unsuccessful parses…

| enjoy | books |
|---|---|
| *(S \ NP) / NP* | *NP* |

$$S \setminus NP$$

| I | sleep | books |
|---|---|---|
| *NP* | *S \ NP* | *NP* |

$$S$$

| I | enjoy |
|---|---|
| *NP* | *(S \ NP) / NP* |

| Books | I | sleep |
|---|---|---|
| *NP* | *NP* | *S \ NP* |

$$S$$

If a parse is successful, the sentence is declared "grammatical"; if not, the sentence is declared "ungrammatical".

# <6> Combinatorial Categorial Grammar (2/4)

**TASK 1**: Using the above examples as evidence, figure out how CCG parses sentences, and describe it briefly here:

**TASK 2:**

In the sentence "I enjoy long books", what categories could we assign to "long"?

**TASK 3:**

Not every grammatical sentence of English will be declared "grammatical" by the process above. Using only the words "I", "books", "sleep", and "enjoy", form a grammatically correct English sentence that will fail to parse given the categories above.

**PART 2: TOK PISIN**

Tok Pisin (also referred to as New Guinea Pidgin or Melanesian Pidgin) is a creole language spoken in the northern mainland of Papua New Guinea and surrounding islands. It is an official language and the most widely used language in the country, spoken by over 5 million people.

Many Tok Pisin words come originally from English – its name comes from "talk" and "pidgin"[1] — but Tok Pisin isn't just English. It has a distinct grammar and uses these words in different (but systematic!) ways.

---

[1] A pidgin language is one developed by two or more groups of people who do not share a common language. Tok Pisin started out as a pidgin but has since developed into a creole, a complex language in its own right.

# <6> Combinatorial Categorial Grammar (3/4)

**TASK 4**:

Below are sentences in Tok Pisin with a scrambled list of English translations.  Match each sentence to its English equivalent. (Answer by writing letter (A-H) to left of Tok Pisin sentence.)

|  |  |  |  |
|---|---|---|---|
| ____ | *1. Brata bilong em i stap rit.* | A. | He has read a book. |
| ____ | *2. Ol i stap dringim wara.* | B. | My sister boils the water. |
| ____ | *3. Ol i ken ritim buk bilong mi.* | C. | They can read my book. |
| ____ | *4. Em i ritim buk pinis.* | D. | His sister can write. |
| ____ | *5. Em i laik rit.* | E. | His brother is reading. |
| ____ | *6. Susa bilong em i ken rait.* | F. | The water has boiled. |
| ____ | *7. Susa bilong mi i boilim wara.* | G. | He wants to read. |
| ____ | *8. Wara i boil pinis.* | H. | They are drinking water. |

**TASK 5**:

Translate the following Tok Pisin sentence into English:

*"Brata bilong mi i stap ritim buk bilong susa bilong mi."*

**TASK 6:**

Translate the following English sentence into Tok Pisin:

"Their sister wants to write a book."

# <7> Combinatorial Categorial Grammar (4/4)

**TASK 7:**

Describing these words in terms of their CCG categories highlights that these aren't English words combined according to English rules, but are Tok Pisin words combined according to Tok Pisin rules. One new category label, VP, is introduced here.

Match each Tok Pisin word below to its CCG category by writing a letter (A - G) corresponding to the category.  Some categories will be used more than once. The categories of two words are already identified for you, by way of example.

| | | | | | |
|---|---|---|---|---|---|
| ____ | 1. *bilong* | G | 10. *laik* | A. | NP |
| ____ | 2. *brata* | ____ | 11. *mi* | B. | (NP \ NP) / NP |
| D | 3. *boil* | ____ | 12. *ol* | C. | (S \ NP) / VP |
| ____ | 4. *boilim* | ____ | 13. *pinis* | D. | VP |
| ____ | 5. *buk* | ____ | 14. *stap* | E. | VP / NP |
| ____ | 6. *dringim* | ____ | 15. *raitim* | F. | VP \ VP |
| ____ | 7. *em* | ____ | 16. *rit* | G. | VP / VP |
| ____ | 8. *i* | ____ | 17. *ritim* | | |
| ____ | 9. *ken* | ____ | 18.  *susa* | | |

# Acknowledgements

# This OzCLO quiz was devised in cooperation with:

AILO (All Ireland Linguistics Olympiad)
NACLO (North American Computational Linguistics Olympiad)
UKLO (UK Linguistics Olympiad)

# Problem Credits