

Language Puzzles

Fall 2019

Hi, everyone! My name is Evan Hochstein, and I'm very excited to share languages and language puzzles with all of you! I first heard about all this through the North American Computational Linguistics Olympiad (NACLO), which is a long and intimidating name for a very fun competition structured all around puzzles. I competed in NACLO all 4 years of high school, and I hope you will try it out.

In the way of introduction, I'm from Austin, Texas (yeehaw), and my language of choice in middle and high school was Spanish. In college, I have chosen to start again from scratch, and I am learning classical hieroglyphic Egyptian. And some people think learning *Latin* is useless... Feel free to pick my brain about any language, however — I love all of them.

Now some information about the class.

Topics

- Week 1: English-based, number system, phonology, sound change, writing system, Rosetta Stone, matching
- Week 2: Rosetta Stone, matching, phonology/morphophonology, scripts, directional systems
- Week 3: no translation, hard matching, possession, hard number system, cultural difference, morphology

Note: All of our puzzles come from NACLO (nacloweb.org) except for one or two. For each puzzle, I might walk us through it, or have you work in small groups, or see what you can do on your own. Also, I probably won't get to all of the topics, but they're all in your packets for you to explore.

Suggested Puzzles

At the back of your packets, there will be a list of puzzles that are similar to the ones in your packets so you can practice the skills we've built up in class and see the same concepts applied in different languages and styles. These puzzles are not required, but I think you will enjoy them. Even if a puzzle is difficult and you don't figure out "the trick" or you miss something, taking the time to really work at these puzzles will help you make the most gains, and if you actually make a breakthrough and solve a difficult puzzle after a long time, it's that much more rewarding (plus you'll be on the lookout for that trick if it ever appears again).

If you ever have any questions, feel free to contact me at evan.hochstein@yale.edu.

YOUR NAME:

REGISTRATION #

H. Waorani Numbers (I/I)

[10 points]

Waorani is an indigenous language from Ecuador. The following arithmetic expressions involve the numbers from 1 to 10 in Waorani. The \times symbol is used to denote multiplication. The square of a number n is marked as n^2 . Each underlined sequence represents one number from 1 to 10.

- (a) měña mēña mēña mēña + mēña go mēña = ãëmãëmpoke go aroke \times 2
- (b) aroke² + mēña² = ãëmãëmpoke
- (c) ãëmãëmpoke go aroke² = mēña go mēña \times ãëmãëmpoke mēña go mēña
- (d) mēña \times ãëmãëmpoke = tipãëmpoke

H-I Can you write these numbers in Waorani?

3 =

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

5 =

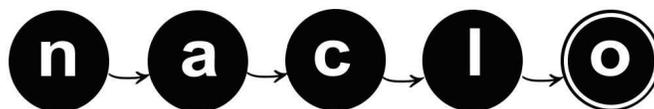
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

7 =

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

10 =

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



What's the Time in Tallinn? (NACLO 2014)

Tallinn is the capital of Estonia, where about 1 million people speak Estonian, a non-Indo-European language closely related to Finnish.

The following expressions show how to tell the time in Estonian:



Kell on üks.



Kell on kaks.



Veerand kaks.



Pool neli.



Kolmveerand üksteist.



Viis minutit üks läbi.

Here are some numbers in Estonian:

6 kuus

7 seitse

8 kaheksa

10 kümme

A1. Translate the following times into Estonian:

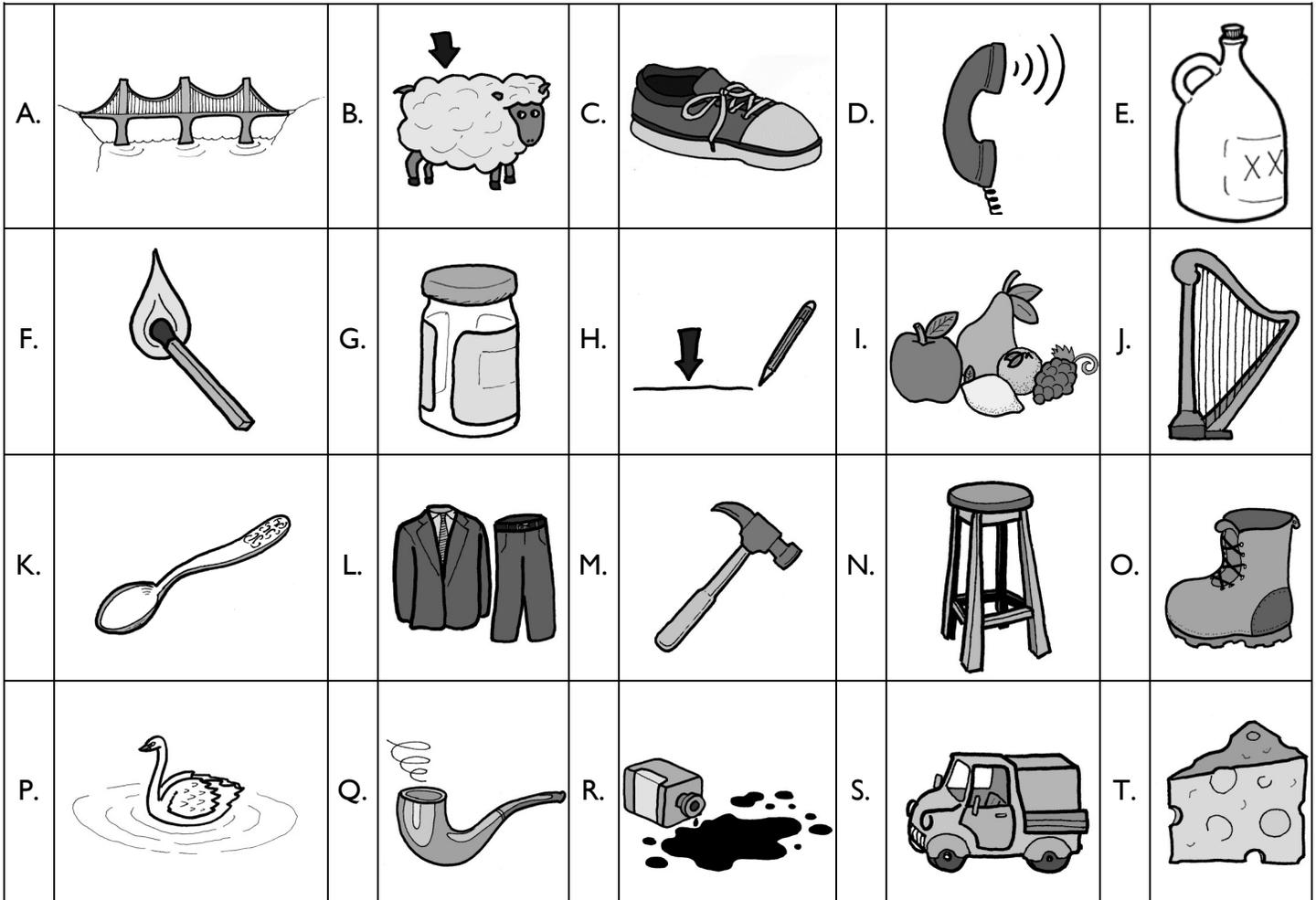
- a. 8:45
- b. 4:15
- c. 11:30
- d. 7:05
- e. 12:30

A2. What do the following Estonian time expressions mean?

- a. Kaksikümmend viis minutit üheksa läbi.
- b. Veerand neli.
- c. Pool kolm.
- d. Kolmveerand kaksteist.
- e. Kolmkümmend viis minutit kuus läbi.

(B) Māori Loanwords (1/2) [15 points]

The Māori language, or “te reo Māori”, is the language of the Māori, the indigenous people of New Zealand. It is one of the official languages of New Zealand, along with English and New Zealand Sign Language, and over several centuries it has borrowed many English words. These words are often adapted to better fit the sounds of the Māori language.



B1. Below are 20 Māori words that have been adapted from English words. Note that Māori uses a line over vowels to mark them as long. Can you match each word below to the picture that illustrates it?

1.	hāma	6.	māti	11.	raina	16.	tīhi
2.	hāpa	7.	paipa	12.	taraka	17.	tūru
3.	hū	8.	piriti	13.	terewhono	18.	wāna
4.	hūtu	9.	pūnu	14.	tiā	19.	whurutu
5.	iniki	10.	pūtu	15.	tiaka	20.	wūru



YOUR NAME:

REGISTRATION #

(B) Māori Loanwords (2/2)

B2. Many English loanwords in Māori deal with introduced Western professions and titles. To what English words do the following Māori words correspond?

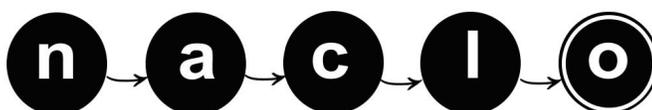
hekeretari	
pirinihehe	
pirihimana	
tiati	

B3. What countries are these?

Iharaira	
Kiupa	
Peina	
Tiamani	
Tiapana	

B4. For each of these English words, predict what the Māori form would be:

beef	
bull	
cart	
clock	
lease	
meat	
seal	
street	
time	
watch	



Old English (NACLO 2015)

English has changed a lot since the period in which Germanic languages were brought to the British Isles over 1500 years ago. In this puzzle, you will look at some ways in which Old English (the name we give to the varieties of English that existed from the 5th century to the 12th century) differs from Modern English.

Note: The letter þ is pronounced like the 'th' in *thin*; the letter æ is pronounced like the 'a' in *cat*.

Old English	Modern English
wit lufodon þæt mægden	we two loved the girl
þæt cild unc lufode	the child loved us two
ge lufodon þone cyning	you all loved the king
se cyning inc lufode	the king loved you two
þæt mægden we lufodon	we all loved the girl
we inc lufodon	we all loved you two
wit eow lufodon	we two loved you all
se æþeling unc lufode	the prince loved us two

I1. Translate the following from Old English into Modern English.

- se cyning eow lufode
- ge lufodon þæt mægden
- wit inc lufodon

I2. Translate the following from Modern English into Old English.

- The prince loved the child.
- The child loved the prince.
- We all loved the child.
- The child loved you two.

(D) Let's Roll! (1/1) [10 points]

Tavla is a variant of backgammon played in Turkey. In the game, two six-sided dice are thrown, resulting in two random numbers between 1 and 6. Each possible outcome has a name, and these names consist of a curious mixture of Persian and Turkish. Some letters used to write these languages are not used to write English, such as *ü*, *ş*, and *ı*, but you do not need to know how these are pronounced to solve this problem. The table on the left lists some of the possible outcomes, while the table on the right lists the names of those outcomes, but the outcomes and their names are not given in the same order. Answer these questions in the Answer Sheets.

D1. Match up which name goes with which outcome. Write your answers in the Answer Sheets (that is, next to each outcome in the Answer Sheet, write the letter corresponding to that outcome's name). Note that the order of the dice does not matter for the naming scheme: For example, 2-6 and 6-2 have the same name.

	<i>Outcomes</i>		<i>Names</i>
a.	1-3	(A)	pencü yek
b.	4-6	(B)	pencü se
c.	1-5	(C)	şeşi yek
d.	3-5	(D)	şes cıhar
e.	1-4	(E)	se yek
f.	1-6	(F)	şesi dü
g.	2-6	(G)	cıharı yek

a.

b.

c.

d.

e.

f.

g.

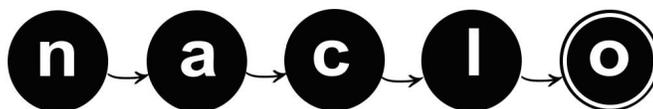
For Questions **D2** through **D4**, your answer should be an element from the following list of possible outcomes: 1-1, 1-2, 1-3, 1-4, 1-5, 1-6, 2-2, 2-3, 2-4, 2-5, 2-6, 3-3, 3-4, 3-5, 3-6, 4-4, 4-5, 4-6, 5-5, 5-6, 6-6

D2. What outcome has the name "pencü dü"?

D3. What outcome has the name "cıhari se"?

D4. What outcome has the name "düşeş"?

NOTE: There is significant variation in the exact names used across different Tavla players from different regions. Therefore, some Tavla players use different sets of names than the ones given in this problem.



(5 points)

(G) Rewrite me badd

You speak a little differently than your parents do. They probably say that you're speaking "bad English". Every generation of parents says this, but this is just how language works. In fact, this is where languages come from: enough generations of young people speaking "bad Latin" and eventually you have Spanish, French, and Italian!

Huishu is a language in the Tangkhulic family and is spoken in the easternmost part of India. Over time, enough changes occurred in this one village that the villagers now speak a different language than any of their neighboring villages. So where they used to say "-lo" ("buy"), they now say "-lu", and where they used to say "-muk" ("cattle") they now say "-mu?". (That symbol at the end represents the sound in the middle of "Uh-oh!", and the dashes in front just mean that these have to occur as parts of larger words.)

Linguists model historical sound changes as "string-rewrite rules". These are very much like a "find-and-replace" procedure in a word processor: look for one character or pattern and replace it with another one. As the old language changed into modern Huishu, the following string-rewrite rules applied:

K-Insertion: When you find an [u] at the end of the word, add a [k] after it.

Vowel-Raising: When you find an [o] at the end of the word, replace it with [u].

K-Deletion: When you find a [k] at the end of a word, replace it with [?].

These changes didn't all just happen at once, though. They happened one after another – although not necessarily in the order above! – and we can see in which order they happened by comparing the old forms to the new forms. Only one order will work; if these changes had happened in any other order, we would have different modern words.

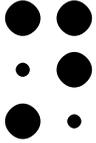
Here are a few such pairs (the old form is at the top; the new one is at the bottom). From these, can you determine the order in which the above changes must have occurred? Write the names of the rules in the blanks on the left. The blanks in between each pair are for your benefit: if you write how each word changed as each rule applied, you should be able to work out their ordering in time.

Proto-Tangkhulic form:	-ru ("bone")	-khuk ("knee")	-ko ("nine")
Rule 1: _____			
Intermediate form 1: Rule 2: _____	_____	_____	_____
Intermediate form 2: Rule 3: _____	_____	_____	_____
Huishu form:	-ruk	-khu?	-ku

(20 points)

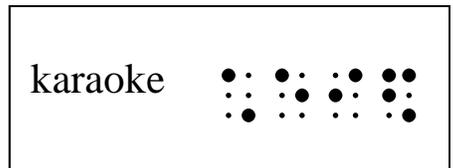
(A) Tenji karaoke (I/I)

Braille is a tactile writing system, based on a series of raised dots, that is widely used by the blind. It was invented in 1821 by Louis Braille to write French, but has since been adapted to many other languages.



English, which uses the Roman alphabet just as French does, required very little adaptation, but languages that do not use the Roman alphabet, such as Japanese, Korean, or Chinese, are often organized in a very different manner!

To the right is a Japanese word written in the *tenji* ("dot characters") writing system. The large dots represent the raised bumps; the tiny dots represent empty positions.



A1. The following *tenji* words represent *atari*, *haiku*, *katana*, *kimono*, *koi*, and *sake*. Which is which? You don't need to know either Japanese or Braille to figure it out; you'll find that the system is highly logical.

a. _____	b. _____
c. _____	d. _____
e. _____	f. _____

A2. What are the following words?

g. _____	h. _____
----------	----------

A3. Write the following words in *tenji* characters. Please write very legibly!!

i. samurai	j. miso
------------	---------



(A) At Ease in Gilbertese (1/1) [10 points]

The Gilbertese language is spoken by most of the over 100,000 people who live in the Pacific island nation of Kiribati. Here are some sentences in Gilbertese with their English translations:

Gilbertese	English
ko nakonako ɲkoe	You are walking.
e nakonako te aiine	A woman is walking.
i takaakaro ɲai	I am playing.
e nakonako nakon te titooa Meeri	Mary is walking to the store.
a tekateka irarikin taian auti taian aiine	The women are sitting next to the houses.
a tebotebo taian nakekei n te bong aei	The people are bathing today.
i tebotebo inanon te auti ɲai	I am bathing in the house.
a takaakaro inanon te titooa taian ataei	The children are playing in the store.
ko tekateka ɲkoe ningaabong	You will sit tomorrow.
a takaakaro irarikin te kawai taian ataei n te bong aei	The children are playing next to the road today.

A1. Below are some Gilbertese sentences with their words in scrambled order:

- bong n Meeri auti e nakon te te nakonako aei
- inanon ɲai kawai i te tekateka
- ataei a nakon taian ningaabong kawai nakonako te

For each of these three sentences:

- Rearrange the words in the sentence into the correct Gilbertese word order.
- Translate the rearranged sentence into English.

a.i

a.ii

b.i

b.ii

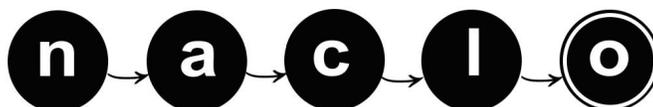
c.i

c.ii

A2. Translate into Gilbertese:

- The women will play tomorrow.

- You are sitting next to the store today.



YOUR NAME:

REGISTRATION #

(M) Come to Istanbul (I/I) [10 points]

Turkish is spoken by about 63 million people, of whom most live in Turkey but about 100,000 live in the UK. It is a non-Indo-European language, so it is unrelated to English but related to languages of Central Asia such as Azeri and Uzbek.

Turkish words are built up by adding one or more endings to a root word; the vowels in most word endings vary depending on the vowels in the root word ("vowel harmony"), as you will see in the following examples. Here are some sentences in Turkish, with their English translations. Note:

- 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 letter "ğ" is usually silent (like the "gh" in "although").
- Square brackets [] enclose English words that are not directly translated.

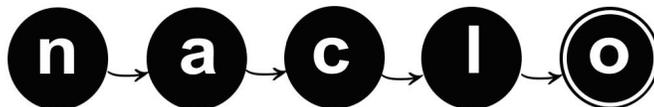
Arkadaşlarım şehirde mutlu	My friends [are] happy in [the] city.
Baban İstanbul'u seviyor mu?	Does your father like Istanbul?
Fakirler Van'dan İstanbul'a gelmek istiyor	Poor [people] want to come from [the city of] Van to Istanbul.
İstanbul en büyük şehir	Istanbul [is the] biggest city.
Eve geliyorlar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

M1. How would you translate the following into English?

A.	Baban mutlu mu?	
B.	"Şehrimize gel" diyoruz.	
C.	Arkadaşım doktor olmak istiyor.	
D.	Fakir evimi seviyorlar mı?	
E.	İstanbul'dan mı geliyorsun?	

M2. The following examples introduce a new pattern. What do you think these examples mean?

A.	Geldiğimde "merhaba" diyorlar.	
B.	Baban geldiğimizden mutlu mu?	
C.	Fakir olduğunu diyorlar.	
D.	Aldığın ev büyük mü?	
E.	En mutlu olduğum şehir, Van.	
F.	Fakir olduğumuz halde mutluyuz.	



To Be Determined (NACLO 2017)

Bulgarian is a Slavic language and, along with the closely related Macedonian language, has several characteristics that set it apart from all other Slavic languages. Old Bulgarian was a highly synthetic language; synthetic languages express meaning by building complex words from smaller pieces. Modern day Bulgarian is a typical analytic language, one which expresses meaning more with word order, like English. Bulgarian currently uses the Cyrillic script, created after the older Glagolitic alphabets. In this problem, a Roman transliteration is used. The letter ‘ă’ represents the vowel in the English word *but*.

Here are some sentences given in Bulgarian and the translated equivalents (in no particular order) in English.

Bulgarian	English
1. Veshterăt nahrani maymunata.	(A) Your son watched you.
2. Kamilata vărvoja.	(B) The girl hugged the cat.
3. Momicheto pregărna kotkata.	(C) You dressed yourself.
4. Veshtitsata prokle kotkata.	(D) The cat scratched you.
5. Kotkata prokle tvoja sin.	(E) You fed the son.
6. Ti nahrani sina.	(F) The witch cursed the cat.
7. Kotkata te odraska.	(G) The camel walked.
8. Ti skochi.	(H) The cat cursed your son.
9. Tvoiat sin te gleda.	(I) The wizard fed the monkey.
10. Veshterăt pregărna edna kamila.	(J) The son dressed your baby.
11. Ti se obleche.	(K) You jumped.
12. Sinăt obleche tvoeto bebe.	(L) The wizard hugged a camel.

N1. Match the Bulgarian sentences with their corresponding English translations.

N2. Translate the following sentences into English:

- Maymunata gleda tvoyata veshtitsa.
- Tvoyata kamila obleche edno momiche.
- Veshterăt se prokle.
- Ti pregărna bebeto.
- Ti vărvoja.
- Ti prokle edin veshter.

N3. Translate the following sentences into Bulgarian:

- The witch dressed you.
- The baby watched the girl.
- The monkey jumped.
- You hugged a son.
- Your son dressed a baby.

Tangkhol Tangle (NACLO 2010)

Tangkhol is a language spoken in the northernmost district of the Indian state of Manipur. Like Manipuri (or Meithei) and many other languages of Northeast India, Tangkhol is related to Tibetan and Burmese rather than to Hindi, Bengali, Marathi, Gujarati, or other well-known languages of India.

Tangkhol words can be very long and quite complicated in their structure. Sometimes single words may have to be translated with whole sentences in English. Also, pronouns (words like he, she, it, and they) can be left out if their meanings can still be filled in from context. Following are a list of sentences from Tangkhol and their English translations (in alphabetical order). In the English translations, pronouns are enclosed in parenthesis when they are left out of the Tangkhol sentences. Tangkhol, unlike Modern English (but like Old English), distinguishes three different grammatical numbers: singular (referring to one person or thing), dual (referring to two persons or things), and plural (referring to three or more persons or things). The abbreviations sg., dl., and pl. indicate “singular,” “dual” and “plural,” respectively.

G1. Match the Tangkhol sentences with their English translations by writing the number of the English translation by the corresponding Tangkhol sentence (8 points)

Tangkhol sentences

- a) a masikserra
- b) āni masakngarokei
- c) āthum masakngarokngāilā
- d) ini thāingarokei
- e) na thāilā
- f) ithum thāingāihāirara
- g) rāserhāira
- h) āni rāra
- i) nathum rāserhāiralā

English translations

- 1) Do they (pl.) want to pinch one another?
- 2) Do you (sg.) see it?
- 3) Have you (pl.) all come?
- 4) He/she will pinch all (of them).
- 5) (They) all have come.
- 6) They (dl.) pinched one another.
- 7) They (dl.) will come.
- 8) We (pl.) will have wanted to see (it).
- 9) We (dl.) saw one another.

G2. Translate the following sentences into English. Always start with the leftmost box. Please follow the style of the English translations given in G1 as closely as possible. (6 points)

- a) nathum masakserngāira
- b) āthum thāiei
- c) i thāiserhāiralā

G3. Translate the following sentences into Tangkhol (6 points).

- 1) Do you (dl.) want to come?
- 2) You (sg.) have seen (it) all.
- 3) We (pl.) will want to see one another.

(N) What happened at the chess tournament? (1/1)

Hungarian is a Finno-Ugric language spoken in Hungary by about 10 million speakers and about 2.5 million speakers in the surrounding countries, as well as the diaspora. Hungarian is often called a nonconfigurational language, which means that a) the words are unambiguously marked for their role in the sentence and b) the word order is not rigid but often determined by the conversational context the sentences appear in.

Match the Hungarian sentences with their English translations.

- | | |
|--|---|
| N1. Valaki megvert valakit. | (A) No one beat everyone (at e.g. chess). |
| N2. Kit vert meg valaki? | (B) Who wasn't beaten by anyone? |
| N3. Senki nem verte meg a Petyát. | (C) No one got beaten. |
| N4. Valakit senki nem vert meg. | (D) Someone beat Martin. |
| N5. Senki nem vert meg mindenkit. | (E) I didn't beat anyone. |
| N6. Senkit nem vert meg a Petya. | (F) No one beat Peter. |
| N7. Ki nem vert meg senkit? | (G) Who got beaten by someone? |
| N8. Valaki senkit nem vert meg. | (H) Someone beat someone. |
| N9. Mindenki megvert valakit. | (I) Everyone beat someone. |
| N10. Valaki megverte a Marcit. | (J) Who didn't beat anyone? |
| N11. Senkit nem vert meg senki. | (K) There's someone who didn't beat anyone. |
| N12. Kit nem vert meg senki? | (L) Peter beat no one. |
| N13. Nem vertem meg senkit. | (M) There is someone who didn't get beaten. |

1. 2. 3. 4. 5. 6. 7. 8.
9. 10. 11. 12. 13.

N14. Explain your answers.



It's True: The Truth About Chalcatongo Mixtec (NACLO 2018)

Chalcatongo Mixtec is a language spoken by just under 6,000 people in Oaxaca State of South-Central Mexico. It is famous among linguists for its many unusual characteristics, but it is an endangered language, at risk of extinction.

J1. Here are some sentences in Chalcatongo Mixtec. Their English translations have been given in scrambled order on the right. Match the Chalcatongo Mixtec sentences to their English translations.

Note: Chalcatongo Mixtec (when written in the Roman alphabet) has some letter symbols and accents that we do not use in English—these are not relevant for solving this problem. Whenever a word or phrase is bold and underlined (**like this**), it means that the word or phrase is emphatic, but otherwise it is not emphatic. Finally, here is a hint to help you get started: Chalcatongo sentence (1) corresponds to English choice (a).

Chalcatongo sentences

1. Nduča kaa ñí?ní.
2. Maria kúu † † xasúčí.
3. Ñí?ní nduča.
4. Juan kaa lúlí.
5. Ndežu kaa ža?u.
6. Súčí Maria.
7. Juan kúu xažiiri.
8. Pedro kúu xalúlíri.
9. Kaa kwíí.

English translations

- a. **The water** is hot.
- b. Pedro is my child.
- c. Juan is my husband.
- d. **It** is green.
- e. Maria is a young person.
- f. The water is hot.
- g. Maria is young.
- h. **The food** is expensive.
- i. **Juan** is small / short.

Here are some words in Chalcatongo Mixtec. The English translation for each word is given on the right.

ndáa <i>true</i>	bíší <i>sweet</i>	bèè <i>heavy</i>	tûû <i>black</i>	nde?é <i>brave</i>
kúnú <i>deep</i>	ñí?ní <i>hot</i>	kwa?á <i>red</i>	súkú <i>tall</i>	

J2. Translate the following into Chalcatongo Mixtec.

- a. depth
- b. heat
- c. Maria is brave.
- d. **Pedro** is tall.
- e. Pedro is a tall person.
- f. The fruit is red.
- g. My fruit is the green one.
- h. **It** is true.
- i. It is true.
- j. It is the truth.

J3. Explain your answers by describing how Chalcatongo Mixtec words and sentences are structured.

Suggested problems, Week 1 (all from NACLO, www.nacloweb.org/practice.php):

- Similar to Gelda's House of Gelbelgarg (2010):
We are all molistic, in a way (2007)
Predicate etiquette (2019)
- Similar to Rewrite Me Badd (2007):
Sound judgments (2016)
- Similar to Māori Loanwords (2014):
Traevölörs freisbuk (2012)
- Similar to At East in Gilbertese (2018) and Come to Istanbul (2014):
Learn Yolmo with pleasure (2012)
The silence of the goats (2012)
Waanyi (2012)
A case of Pali (2013)
Beja (2013)
To play or not to play (2014)
Learning Yidiny (2014)
Don't sell the house (2016)
A little Tshiluba (2017)
To know or not to know literary Tamil (2018)
You Speak Ndebele, don't you (2019)
Hate it or Lopit (2019) — difficult
- Similar to To Be Determined (2017), Tangkhul (2010), What Happened at the Chess Tournament? (2016), and It's True: The Truth About Chalcatongo Mixtec (2018):
A Menya Puzzle (2018) — difficult
I know what I saw (2018)