

SOLUTIONS

***The Twelfth
Annual***

**North American
Computational
Linguistics
Olympiad**

2018

www.nacloweb.org

Open Round

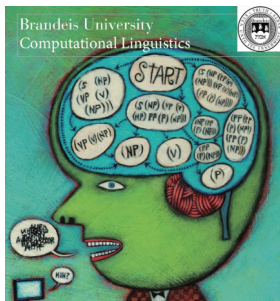
January 25, 2018



Carnegie Mellon University
Language Technologies Institute



**MIDDLE
TENNESSEE**
STATE UNIVERSITY



Yale University

Serious language puzzles that are surprisingly fun!

-Will Shortz, Crossword editor of The New York Times and Puzzlemaster for NPR

(A) At Ease in Gilbertese (1/1)

The order of the words/phrases in a Gilbertese sentence conforms to the following pattern:

Person marker	Verb	Prepositional phrase	Pronoun/noun subject	Temporal phrase
---------------	------	----------------------	----------------------	-----------------

The “person marker” at the beginning of the sentence indicates whether the subject is 1st person singular (I), 2nd person singular (Ko), 3rd person singular (E) or 3rd person plural (A).

A1.

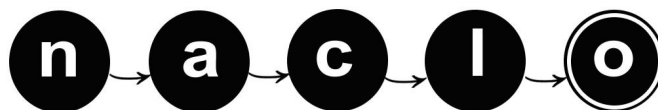
a. e nakonako nakon te auti Meeri n te bong aei
Mary is walking to the house today.

b. i tekateka inanon te kawai ŋai
I am sitting (on) in the road.

c. a nakonako nakon te kawai taian ataei ningaabong
The children will be walking to the road tomorrow.

A2.

a. a takaakaro taian aiine ningaabong
b. ko tekateka irarikin te titooa ŋkoe n te bong aei



(B) The Pame Game (1/1)

Northern Pame:

- **base:** $\text{tenhiu}\eta = 8$;
- **order:** multiplier_base_addend, i.e. the numbers have the following form:
 $\alpha \text{ tenhiu}\eta = 8 * \alpha + \beta$, where $1 \leq \alpha, \beta \leq 7$;
- numbers between 1 and 3 have two different names which are used as multipliers and addends, respectively.

	1	2	3	4	5	6	7
α	kara	kanuje	karnu?	giriuj			
β	sante	nuji	rnu?	giriuj	gitʃʹaj	tiria	teriuhiŋ

Central Pame:

Since the two languages are cognate, it can be expected that at least some words with identical meanings are going to turn out to be more or less similar.

- **base:** 20, subbase 10 (**seskaʔai**);
- **order:** multiplier_base_addend, i.e. the numbers have the following form:
 $\text{seskaʔai} \beta = 10 + \beta$, $\beta \geq 1$;
- α **lien** $\beta = \alpha * 20 + \beta$, $\alpha \geq 1, \beta \geq 1$.

1	nda	5	kik'ai	9	nda ntsaw? nda
2	nui	6	tilija	10	seskaʔai
3	ranhũ?	7	tiliŋũhũŋ	20	nda lien
4	kijui	8	nda ntsaw?	$\beta * 20$	β lien

Answers:

B1.

a.

nuji = 2

karnu? tenhiuŋ teriuhiŋ = $3 * 8 + 7 = 31$

b.

1 = nda

3 = ranhũ?

9 = nda ntsaw? nda

56 = nui lien seskaʔai tilija

60 = ranhũ? lien



(C) All Roads Lead to Tirana (1/1)

C1.

1. G.
2. H.
3. J.
4. A.
5. D.
6. C.
7. F.
8. I.
9. M.
10. B.
11. N.
12. K.
13. E.
14. L.

C2. Playground

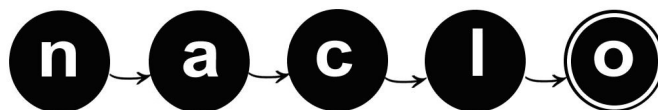
C3.

Drita: 114
Kustrim: 120
Lindita: 117
Sokol: 111
Valon: 116



(D) Intergalactic Grammars (1/1)

- D1.**
- A turtle chased a turtle.
 - Bir kaplumbağa bir kaplumbağa kovaladı.
- D2.**
- nahita sokatra saka. A cat saw a turtle.
 - nanenjika saka sokatra. A turtle chased a cat.
 - nanaikitra saka sokatra. A turtle bit a cat.
 - nihinana saka sokatra. A turtle ate a cat.
- D3.**
- The pet bites. chop Saj.*
 - The U.S.S. Enterprise will battle the Klingons. tlhInganpu' ghobrupqa' 'ejDo' 'entepray'.*
 - The leaders know that the Klingons know that the commander learned that a spy saw that the Klingons have a pet. Saj lughaj tlhInganpu' 'e' leghpu' ghoqwl' 'e' ghojpu' la' 'e' SovTah tlhInganpu' 'e' SovTah Devwl'pu'.*



(E) Parties in Palauan (1/1)

All of the cardinal numbers in Palauan, when used as adjectives, have three possible forms (actually there are more than three) depending on the class of nouns which they qualify. For units of time, the number is preceded by *e-*; for people, the number is preceded by *te-*, and for animals, fruits and some inanimate objects, the number is preceded by *kl-*. Numbers above ten consist of *tęruich me a* (ten) followed by the appropriate number (in the “ones” column), which is marked to indicate the class of noun.

Ordinal numbers are preceded by the morpheme *onge-*, which attaches to the “ten” unit for numbers above ten. The words *sils* and *ureor* both mean “day”, but the former refers to the unit of time while the latter refers to a day of the week.

Answers:

E1.

- a. six teachers
- b. nineteen months
- c. December
- d. Friday

E2.

- a. eai ęl sils
- b. tęruich me a tetiu ęl chad
- c. teuid ęl sensei
- d. ongelolem ęl buil
- e. ongeai ęl buil

E3.

- a. ede
- b. tede
- c. klde



(F) A Match Made in Vietnam (1/1)

F1.

#	Vietnamese	Literal meaning	English answer letter	English
1.	băng	ice	I	ice
2.	bó	cluster	D	cluster
3.	bó hoa	flower cluster	A	bouquet
4.	cánh hoa	flower wing	L	petal
5.	đá	rock	P	rock
6.	đá lửa	fire rock	G	flint
7.	đá phấn	powder rock	B	chalk
8.	đường	road	O	road
9.	đường vòng	circle road	E	detour
10.	hoa	flower	H	flower
11.	lửa	fire	F	fire
12.	mở	to open	S	to open
13.	mở đường	to open a road	T	to pave the way
14.	mở mắt	to open eyes	R	to make aware
15.	núi	mountain	K	mountain
16.	núi băng	ice mountain	J	iceberg
17.	núi lửa	fire mountain	U	volcano
18.	nước đá	rock water	I	ice
19.	nước mắt	eye water	Q	tear
20.	phấn	powder	N	powder
21.	phấn hoa	flower powder	M	pollen
22.	vòng	circle	C	circle
23.	vòng hoa	flower circle	V	wreath

F2. eye(s) (“eye” and “eyes” are both acceptable answers).



(G) By the Same Token (1/2)

G1. “Types” refers to the number of distinct letters found in the word. “Tokens” refers to the total number of letters in the word (i.e., its length. Here are some possible entries for the chart:

Word	Types	Tokens
rock	4	4
paper	4	5
scissors	5	8
shoot	4	5
shaping	7	7
add	2	3
senselessness	4	13
metal	5	5
assesses	3	8
hemidemisemiquaver	11	18
squeegee	5	8
a	1	1
dog	3	3
speaking	8	8
noon	2	4
singing	4	7
unpredictably	13	13

All of the numbers must be the exact values shown, but there are many options for the words; e.g., *a* could be replaced with *l* and *unpredictably* could be replaced with *troublemaking*. (It is very difficult to think of words that fit the 13-13 requirement of being 13 letters long with no repeated letters, but careful observers might notice the hint of *unpredictably* being heavily emphasized in the intro text). Fun fact: The longest English words with no repeated letters are *uncopyrightable* and *dermatoglyphics* (15 letters each), according to <https://www.morewords.com/unique-letters/>.

For half credit, solvers could have use nonsense letter strings that fulfill the given criteria; e.g., *abcdbbb* could have been used instead of *singing*.



(G) By the Same Token (2/2)

G2. Here is what the table looks like with the extra letters removed:

Shilha word	Types in correct Shilha word	Tokens in correct Shilha word	English translation
AFULUS	5	6	male chicken
IFULUSN	6	7	male chickens
TAFULUST	6	8	female chicken
TIFULUSIN	7	9	female chickens
ATBIR	5	5	male pigeon
ITBIRN	5	6	male pigeons
TATBIRT	5	7	female pigeon
TITBIRIN	5	8	female pigeons
AFUNAS	5	6	ox (male bovine)
IFUNASN	6	7	oxen (male bovines)
TAFUNAST	6	8	cow (female bovine)
TIFUNASIN	7	9	cows (female bovines)
ASHLHIY	6	7	male speaker of Shilha
ISHLHIYN	6	8	male speakers of Shilha
TASHLHIYT	7	9	female speaker of Shilha
TISHLHIYIN	7	10	female speakers of Shilha
ADRAR	3	5	mountain (masculine)
IDRARN	5	6	mountains (masculine)
TADRART	4	7	mountain (feminine)
TIDRARIN	6	8	mountains (feminine)

Thus, the answers to the specific questions asked are:

- (l) TATBIRT
- (m) ATBIR
- (n) TAFUNAST
- (o) ADRAR
- (p) TIDRARIN



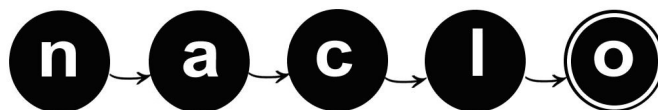
(H) N'ko, M'kay (1/3)

H1. Match up the names 1-12 with their N'Ko equivalents A-L.

1	2	3	4	5	6	7	8	9	10	11	12
E	H	K	D	C	L	J	B	F	G	A	I

H2. Write in N'Ko script the following names:

	Name in transcription:	
a.	Kodo Wari	ƵƵƵƵ ƵƵƵƵ
b.	Liberia	ƵƵƵƵƵƵƵƵ
c.	Mamoun	ƵƵƵƵ
d.	Firiton	ƵƵƵƵƵƵ ƵƵƵƵƵƵ
e.	Netaro	ƵƵƵƵƵ
f.	M'bour	ƵƵƵƵ
g.	N'djala	ƵƵƵƵ
h.	Gberia Fotombou	ƵƵƵƵƵƵƵƵ ƵƵƵƵƵƵƵ



(H) N'ko, M'kay (2/3)

H3. Write the following names in transliteration:

	Name in N'ko	
a.	ⵓⵙⵓⵏⵏ	Djula (Djoula, Dyula, Dyoula)
b.	ⵎⵏⵍⵉ	Mali
c.	ⵍⵉⵏⵏⵉⵎⵓⵏⵏ	Sénégal(i) (Senegal)
d.	ⵃⵏⵔⵓⵏⵏⵓ	Bandoundou (Bandundu)
e.	ⵎⵏⵏⵏⵏⵏ	Maninka
f.	ⵍⵓⵏ ⵍⵓⵏ	Sara Lon (Sierra Leone)
g.	ⵃⵎⵃⵔⵉ	Bambaré
h.	ⵎⵏⵎⵏⵏⵏ	Kambia (Gambia)
i.	ⵃⵓⵔⵓⵎⵏⵏⵏⵏⵏⵏⵏⵏ	Bour(ou)kina Faso (Burkina Faso)
j.	ⵏⵓⵔⵏⵏⵏ	Ir(e)land or Irland

Key to translation:

A	B	D	DJ	E	É	F	G	GB	I	K	L	M	N	O	OU	ɔ	P	R	S	T	W	Y	N'
ⵏ	ⵃ	ⵔ	ⵓ	ⵎ	ⵎ̣	ⵑ	ⵓ̣	ⵓ̣	ⵏ	ⵏ̣	ⵏ̣	ⵎ̣	ⵏ	ⵓ	ⵓ̣	ⵓ̣	ⵑ	ⵑ	ⵑ	ⵑ	ⵑ	ⵑ	ⵑ

Three more letters, ⵏ̣, ⵏ̣, and ⵏ̣, do not appear in this problem, as well as dotted variants of DJ (z) and K (kh), used (like G) for foreign words.

Note (of course) that the writing direction is right-to-left. Mostly the transliteration is simply letter-for-letter, but notice the digraphs DJ, GB, and OU, and that E and É are different. GE, GÉ and GI are written GUE, GUÉ and GUI, and the U is not transcribed (though in fact none of the into-Roman transcriptions require GU).¹ Where a vowel is repeated in a CVCV pattern only the second vowel is written, so Soromaya for example is written SROMYA. This rule applies pairwise, so N'sérégbédé is written N'SRÉGBDÉ (remember GB is a single letter). A dot under a vowel makes it nasalised, transcribed as N or M depending on the following letter. The 'repeated vowel' rule does not apply to nasalised vowels. The sequence transliterated as 'ia' requires a 'y' to be inserted, as in LIBERI(Y)A. And finally, note that the syllabic nasal, written as N' or M' depending on the following sound, has the same symbol ɲ and only occurs word-initially.

1. Ironically, the name of the country written in English as 'Guinea' would actually be DJINÉ in this transcription. No dot.



(H) N'ko, M'kay (3/3)

Some place names have slightly different spellings conventionally (Conakry, Nzerekore, Gueckedou, etc) but have been rendered here in a manner to keep the rules consistent.

Source: <https://catalogingafricana.files.wordpress.com/2012/01/guineepapada1.jpg>

