

NACLO thanks the following for their generous contributions:

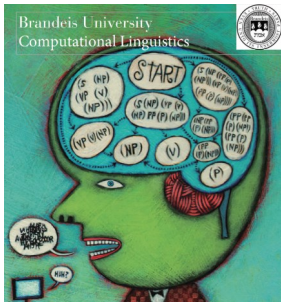


The Fifteenth Annual

North American Computational Linguistics Open Competition

2021

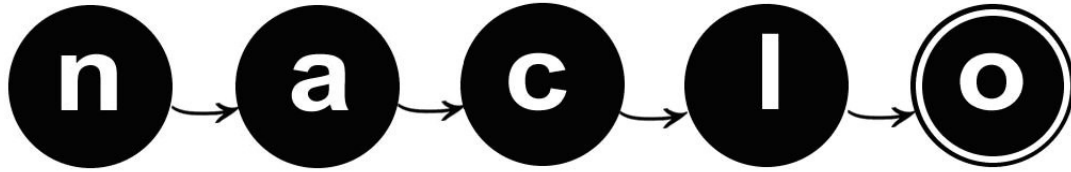
www.nacloweb.org



Invitational Round
March 11, 2021

Serious language puzzles that are surprisingly fun!

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



Welcome to the Invitational Round of the fifteenth annual North American Computational Linguistics Open Competition! In order to be completely fair to all participants across North America, we need you to read, understand, and follow these rules completely.

Rules

1. The contest is four hours long and includes ten problems, labeled J to S.
2. Follow the facilitators' instructions carefully.
3. If you want clarification on any of the problems, talk to your facilitator. The facilitator will consult with the jury and convey their answer.
4. You may not discuss the problems with anyone during or after the contest except as described in item #3.
5. Each problem is worth a specified number of points, with a total of 100 points. **In the Invitational Round, some questions require explanations.**
6. All your answers should be written **clearly** in the Answer Sheets in blue or black ink.
7. Write your name and registration number on each page of the Answer Sheets.
8. You can use the last page of the Answer Sheets if you need extra space to answer a question. Clearly indicate which problem this additional answer applies to.
9. The top students from each country (USA and Canada) will be invited to the next stage.
10. Each problem has been thoroughly checked by linguists and computer scientists as well as students like you for clarity, accuracy, and solvability. Some problems are more difficult than others, but all can be solved using ordinary reasoning and some basic analytic skills. You don't need to know anything about linguistics or about these languages in order to solve them.
11. If we have done our job well, very few contestants will solve all these problems completely in the time allotted. So, don't be discouraged if you don't finish everything.
12. **DO NOT DISCUSS THE PROBLEMS UNTIL THEY HAVE BEEN POSTED ONLINE! THIS MAY BE A COUPLE OF MONTHS AFTER THE END OF THE CONTEST.**

Instructions for At-Home participants

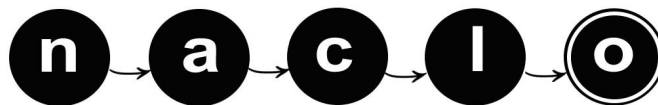
1. Print **one single-sided copy** of the Problems file.
2. Print **two single-sided copies** of the Answer Sheets file.
3. Scan **one copy** of your Answer Sheets to submit for grading. If possible, upload all the pages as one PDF file. Please include all the Answer Sheets pages, even if you left some blank.
4. Be sure your scans are legible before submitting them.
5. If you have technical issues, please ask your facilitator for direction.

...Oh, and have fun!

NACLO 2021 Organizers

Organizing Committee:

Adam Hesterberg — Massachusetts Institute of Technology
Aleka Blackwell — Middle Tennessee State University
Ali Sharman — University of Michigan
Andrés Salanova — University of Ottawa
Andrew Lamont — University of Massachusetts
Andrew Tockman — Massachusetts Institute of Technology
Annie Zhu — Harvard University
Ben LaFond — Harvard University
Brian Xiao — Massachusetts Institute of Technology
Corinne Soucy — Université de Montréal
Daniel Lovsted — McGill University
David Mortensen — Carnegie Mellon University
Dragomir Radev — Yale University
Duligur Ibeling — Stanford University
Ethan Chi — Stanford University
Heidi Lei — Massachusetts Institute of Technology
James Hyett — University of Toronto
James Pustejovsky — Brandeis University
Ji Hun Wang — Stanford University
Ken Jiang — University of Waterloo
Kevin Liang — University of Pennsylvania
Lori Levin — Carnegie Mellon University
Margarita Misirpashayeva — Massachusetts Institute of Technology
Matthew Gardner — Carnegie Mellon University
Mihir Singhal — Massachusetts Institute of Technology
Nathan Kim — Stanford University
Patrick Littell — University of British Columbia
Pranav Krishna — Massachusetts Institute of Technology
Rui Zhang — Yale University
Ryan Chi — Stanford University
Ryan Guan — Stanford University
Shuli Jones — Massachusetts Institute of Technology
Skyelar Raiti — University of Michigan
Sonia Reilly — Massachusetts Institute of Technology
Stella Lau — Massachusetts Institute of Technology
Tom McCoy — Johns Hopkins University



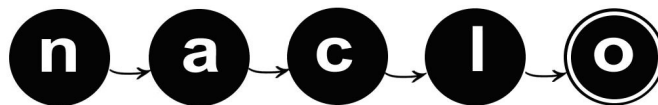
NACLO 2021 Organizers (cont'd)

Organizing Committee (continued):

Tom Roberts — University of California, Santa Cruz
Vanessa Hu — Harvard University
Yilu Zhu — Fordham University

Program Committee:

Adam Hesterberg — Massachusetts Institute of Technology
Alan Chang — University of Chicago
Aleka Blackwell — Middle Tennessee State University
Ali Sharman — University of Michigan
Andrés Salanova — University of Ottawa
Andrew Lamont — University of Massachusetts
Annie Zhu — Harvard University
Babette Verhoeven — Aquinas College
Daniel Harbour — Queen Mary University of London
Daniel Lovsted — McGill University
David Mortensen — Carnegie Mellon University
Dick Hudson — University College London
Dragomir Radev — Yale University
Elisabeth Mayer — Australian National University
Erik Andersen — Brandeis University
Ethan Chi — Stanford University
Gordon Chi — Stanford University
Harold Somers — All Ireland Linguistics Olympiad
Heather Newell — Université du Québec à Montréal
James Hyett — University of Toronto
Jill Vaughan — University of Melbourne
Kevin Liang — University of Pennsylvania
Lori Levin — Carnegie Mellon University
Lynn Clark — University of Canterbury
Margarita Misirpashayeva — Massachusetts Institute of Technology
Mary Laughren — University of Queensland
Oliver Sayeed — University of Pennsylvania
Patrick Littell — University of British Columbia
Simi Hellsten — University of Oxford
Sonia Reilly — Massachusetts Institute of Technology



NACLO 2021 Organizers (cont'd)

Program Committee (continued)

Sophie Ishiwari — University of Pennsylvania

Tom McCoy — Johns Hopkins University

Vlado Keselj — Dalhousie University

Problem Credits:

Problem J: Ethan Chi

Problem K: Harold Somers

Problem L: Evan Hochstein

Problem M: Simi Hellsten

Problem N: Ethan Chi

Problem O: Aleka Blackwell

Problem P: Gordon Chi

Problem Q: Tom McCoy and Ryan Chi

Problem R: Pranav Krishna

Problem S: Ethan Chi and Daniel Lovsted

Booklet Editor:

Pranav Krishna — Massachusetts Institute of Technology

Daniel Lovsted — McGill University

US Team Coaches:

Aleka Blackwell — Middle Tennessee State University

Dragomir Radev — Yale University

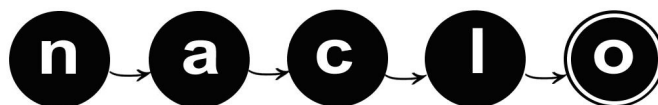
Lori Levin — Carnegie Mellon University

Canadian Coordinator and Team Coach:

Daniel Lovsted — McGill University

We are grateful for the support of many institutional and individual donors who make this contest possible.

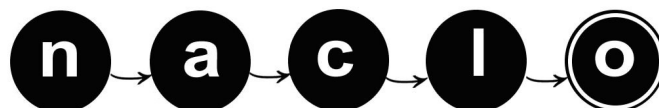
All material in this booklet © 2021, North American Computational Linguistics Open Competition and the authors of the individual problems. Please do not copy or distribute without permission.



(J) A Vintage Sound System (1/2) [10 points]

The Chinese language was first spoken in a small area in Henan, China, around 1000 BCE, during the Zhou Dynasty. However, by the time of the Tang Dynasty (700 CE), many words had changed significantly. For example, the sound *-aj* lost its final *-j*. Below is a table of some words in Chinese, according to a recent historical reconstruction of the language. Each word has two pronunciations: those of the Zhou era (Old Chinese) and those of the Tang era (Middle Chinese). (A pronunciation guide is provided on the next page.) However, some pronunciations in the table are missing.

Old Chinese (Zhou)	Middle Chinese (Tang)	English Translation
(1)	<i>dó</i>	'to come to'
<i>m^ʰə</i>	<i>mō</i>	'soot'
<i>rajs</i>	<i>ljè</i>	'to revile'
<i>p^ʰək</i>	(2)	'north'
<i>p^ʰat</i>	(3)	'to stop in the open'
<i>l^ʰep</i>	<i>dēp</i>	'butterfly'
(4)	<i>bāk</i>	'calm, still'
<i>dzak</i>	<i>dzjĕk</i>	'stone'
<i>braj</i>	<i>bjĕ</i>	'to exhaust'
<i>ŋ^ʰajs</i>	<i>ŋà</i>	'hungry'
<i>p^ʰeks</i>	<i>pè</i>	'favorite'
<i>p^ʰaj</i>	<i>pā</i>	'wave'
<i>dz^ʰə</i>	(5)	'wealth'
<i>t^ʰep</i>	<i>tĕp</i>	'paralyzed'
<i>k^ʰe</i>	<i>kĕ</i>	'chicken'
<i>nə?</i>	<i>ní</i>	'ear'
<i>gres</i>	<i>gjè</i>	'water-chestnut'
<i>prə?</i>	<i>pí</i>	'border town'
<i>gə</i>	<i>gī</i>	'his'
<i>l^ʰaj</i>	(6)	'to flow'
<i>graj?</i>	<i>gjé</i>	'to stand'
<i>tək</i>	<i>tĭk</i>	'to go to'
<i>s^ʰəks</i>	<i>sò</i>	'frontier'
<i>mraj?</i>	<i>mjé</i>	'to share with'
<i>be?</i>	<i>bjé</i>	'female servant'
<i>raj</i>	<i>ljĕ</i>	'to drag into'
(7)	<i>sī</i>	'silk'
<i>l^ʰek</i>	<i>dĕk</i>	(name of an ancient tribe, the Beidi, to the north of China)



(J) A Vintage Sound System (2/2)

J1. Fill in the missing pronunciations from the following choices: *dā, dzō, sǎ, pǒk, l'ǎʔ, b'ak, pǎt*.

J2. Match the Old Chinese words on the left to the Middle Chinese words on the right.

Old Chinese

1. *p'ajʔs*
2. *p'ajʔ*
3. *pajʔ*
4. *pe*
5. *p'ek*
6. *mraj*

Middle Chinese

- A. *pjē* 'humble'
- B. *pà* 'to winnow'
- C. *mjē* 'rice gruel'
- D. *pěk* 'wall (of a house)'
- E. *pjé* 'that'
- F. *pá* 'to limp'

J3. Give the Middle Chinese pronunciations of the following Old Chinese words.

Old Chinese	Middle Chinese	English Translation
<i>kraj</i>	(1)	'bridle'
<i>n'ǎ</i>	(2)	'violent'
<i>rak</i>	(3)	'female servant'
<i>pre</i>	(4)	'pole'
<i>b'ǎʔ</i>	(5)	'double'
<i>m'ajs</i>	(6)	'dust'

J4. Give the Old Chinese pronunciations of the following Middle Chinese words. If there are multiple possibilities, write all of them, separated by commas.

Old Chinese	Middle Chinese	English Translation
(7)	<i>ŋjē</i>	'to make a sacrifice to the deity of the soil'
(8)	<i>tōk</i>	'to obtain'

J5. Explain the sound changes that occurred between Old and Middle Chinese.

Pronunciation Notes:

- The letter $\acute{}$ after a consonant marks pharyngealization of the preceding consonant (constriction of the throat).
- *j* is pronounced like *y* in English *yes*.
- ʔ is the so-called glottal stop (the sound between the two syllables of *uh-oh*).
- *ǎ* is the vowel of English *cut*.
- *ŋ* is the final sound of English *hang*.
- The symbols $\bar{}$, $\acute{}$, $\grave{}$ and \sim mark tones.

Make sure you record your answers in your Answer Sheets!



(K) Putting a Place to a Name (1/2) [5 points]

Tamazight is a family of closely related languages spoken by tens of millions of people across North Africa. Tamazight languages are official languages in Morocco (which is also called *Imeghrib* in Tamazight) and Algeria (in Tamazight, *Dzhayr* or *Lezzayer*).

Tamazight can be written using the Latin alphabet, but it also uses the Tifinagh script, which dates back more than 2000 years, although it has been adapted for modern use. Tifinagh can be written left-to-right, right-to-left, or bottom-to-top, with the orientation of some of the symbols altered accordingly. In this problem, all words are written left-to-right.

On the next page is a list of place names in Tamazight, written in the Tifinagh script, on the left, and the same places named in English on the right, in a scrambled order. Note that the Tamazight names and the English names are not always exactly the same as each other (even after converting from one alphabet to the other). For two of the places, the names are really quite different.

K1. Match up the Tamazight names with their English equivalents.

K2. What do we call the region that in Tamazight is called ⵜⴰⴳⴷⵓⴷⴰ ?



(K) Putting a Place to a Name (2/2)

Tamazight Names	English Names
1. ⵏⵔⵉⵙⵓ	A. Adrar des Ifoghas
2. ⵜⴰⵏⴰⵔⴰⵏⵜ	B. Oran
3. ⵏⴰⵙⴰⵔ	C. Fez
4. ⵏⴰⵙⴰⵏⵓ	D. Toubkal
5. ⵏⴰⵙⴰⵏⵓ	E. Timbuktu
6. ⵏⴰⵙⴰⵏⵓ	F. Béchar
7. ⵏⴰⵙⴰⵏⵓ	G. Safi
8. ⵏⴰⵙⴰⵏⵓ	H. Tangiers
9. ⵏⴰⵙⴰⵏⵓ	I. Oujda
10. ⵏⴰⵙⴰⵏⵓ	J. Nouakchott
11. ⵏⴰⵙⴰⵏⵓ	K. Tamanraset
12. ⵏⴰⵙⴰⵏⵓ	L. Annaba
13. ⵏⴰⵙⴰⵏⵓ	M. Tindouf
14. ⵏⴰⵙⴰⵏⵓ ⵏⴰⵙⴰⵏⵓ	N. Meknès
15. ⵏⴰⵙⴰⵏⵓ	O. Laayoune
16. ⵏⴰⵙⴰⵏⵓ	P. Canaries
17. ⵏⴰⵙⴰⵏⵓ	Q. Agadir
18. ⵏⴰⵙⴰⵏⵓ	R. Tripoli
19. ⵏⴰⵙⴰⵏⵓ	S. Marrakesh
20. ⵏⴰⵙⴰⵏⵓ	T. Sousse
21. ⵏⴰⵙⴰⵏⵓ	U. Rabat
22. ⵏⴰⵙⴰⵏⵓ	V. Constantine
23. ⵏⴰⵙⴰⵏⵓ	W. Gao
24. ⵏⴰⵙⴰⵏⵓ	X. Siwa
25. ⵏⴰⵙⴰⵏⵓ	Y. Algiers
26. ⵏⴰⵙⴰⵏⵓ	Z. Casablanca

Make sure you record your answers in your Answer Sheets!



(L) Is This Problem Intelligible? (1/2) [10 Points]

Hawu¹ (with around 110,000 speakers in 5 dialects) and Dhao² (with around 5000 speakers) are spoken on three islands in the southernmost province of Indonesia. Similarities between the vocabularies of Hawu and Dhao have led some writers to conclude that Dhao should be classified as another dialect of Hawu. For example, here are some related words in Hawu and Dhao (with one cell intentionally left blank):

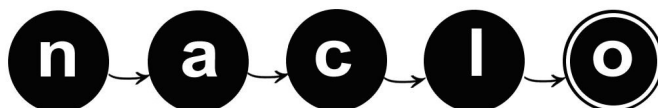
Hawu	Dhao	English
hupa	subha	swear an oath
puru	puru	descend
d'ue	dua	two
b'èhi	bèsi	iron, steel
d'ida	dedha	above
pedutu		follow
ana telora	ana talora	middle child
bèj'i	bhèj'i	sleep
pepuru	papuru	lower (something)
do	dhu	relativizer "which"

Below on the left are twelve sentences, six in Hawu and six in Dhao. The Hawu and Dhao sentences are mixed together. On the right are their English translations in an arbitrary order. Each English translation corresponds to one sentence in Hawu and one sentence in Dhao.

- | | |
|---|--|
| 1. Èi suti. | a. She is walking along the edge of the sea. |
| 2. Pehewina noo ri roo. | b. They keep walking to Seba. |
| 3. Ra kako taruu asa Sèba. | c. They see her head. |
| 4. Ladhe ina na sanède, baku pakèdi. | d. They reminded her. |
| 5. Ta nèru ke noo oro ngidi dahi. | e. If her mother remembers, don't leave. |
| 6. Hutì ne èi. | f. The water spilled. |
| 7. Ra pasanède na. | |
| 8. Na kako madhutu sebhe dhasi. | |
| 9. Kì ta hewina ke ne ina noo, b'ole pekèd'i. | |
| 10. Ta ngède ke ri roo ne kètu noo. | |
| 11. Ta nèru ke roo teruu la Hèb'a. | |
| 12. Ra ladhe kètu na. | |

¹ Until conversions in the 1970s, most Hawu people maintained their traditional religion and ways of life. The Hawu people remember genealogies spanning hundreds of years that preserve Hawu history and structure Hawu society.

² The Dhao people recount that their island was first settled by people from the island of Hawu. The Hawu also tell a version of this history. Traditionally, Dhao women are weavers, while Dhao men are gold- and silver-smiths.



(L) Is This Problem Intelligible? (2/2)

L1. For each sentence, indicate whether it is in Hawu or in Dhao. Then match it to the corresponding English translation.

L2. Translate the following sentences (from Hawu or Dhao as appropriate):

- a. **Ra pasanède ina.**
- b. **Ki ta pedutu ke roo ri ina noo, ta ngède ke noo ri roo.**
- c. **Pehewina roo ri noo.**
- d. **Ladhe na puru, na ladhe sebhe.**
- e. **B'ole bèj'i.**

L3. Translate into Hawu and into Dhao:

- a. Don't walk to the sea.
- b. They keep seeing their mother.
- c. She spilled the water.

L4. In the table of Hawu and Dhao words at the beginning of the problem, the Dhao word related to the Hawu word **pedutu** was left blank. Fill in the blank with the related Dhao word. (Hint: it appears somewhere in this problem!)

L5. Dialects of a language tend to be largely mutually intelligible — that is, speakers of two dialects can understand each other without much effort. Based on the features of Hawu and Dhao that you have observed, are Hawu and Dhao mutually intelligible? Answer in three sentences or less.

Make sure you record your answers in your Answer Sheets!



(M) The Speech Has No End (1/2) [15 points]

Tawala is an Oceanic language spoken in Papua New Guinea by 20,000 people who live in small villages on the East Cape peninsula.

Below and on the next page are some Tawala sentences and their translations. Note that *you(sg)* means “you (referring to one person),” and *you(pl)* means “you (referring to more than one person).” Also note that some of the Tawala words feature *reduplication*, which is a linguistic process in which all or a part of a word is repeated. For this problem, you will not need to figure out *how* a reduplicated form is produced from a non-reduplicated form, but you should pay attention to *when* a reduplicated form is used and when a non-reduplicated form is used.

- 1) **Aem daodaohi ega witewiteihi.**
Your(sg) long legs are never difficult.
- 2) **Ogaleya ma ega igalemi babana itowotowolo.**
You(pl) see him but he does not see you(pl) because he is standing.
- 3) **Apo onanae babana odaladala.**
You(pl) will go because you(pl) are crawling.
- 4) **Apo onadewadewa babana uwitai.**
You(pl) will be good because you(sg) are being difficult.
- 5) **Baha anona ma lawa memehi.**
The speech has substance but the people are settled.
- 6) **Ega Limi natuna babana Limi ega natuna.**
That is not Limi’s child because Limi has no children.
- 7) **Gala gobugobuhi babana tahaya bigabigana.**
The clothes are stained because the path is swampy.
- 8) **Hiduhuduhuna babana baha ega sigasigana.**
They are sitting down because the speech has no end.
- 9) **Hinam ginouli idanedanenehi ma ega dewadewana.**
Your(sg) mother is stealing things and is never good.
- 10) **Inapa ega unananae nu tahaya apo hinapanim.**
If you(sg) aren’t going to the path, they will imprison you(sg).
- 11) **Inapa unapeu apo unagobu.**
If you(sg) fall, you(sg) will be dirty.
- 12) **Lawa pipeuna imae.**
The falling person is staying.
- 13) **Natuhi ega lawa memena natuna.**
Their child is not the settled person’s child.
- 14) **Tahaya ibigabiga ma apo mawa inadao.**
The path is muddy (because it has rained) and the distance will be long.



(M) The Speech Has No End (2/2)

- 15) **Tewela daladalana ogaleya ma gala higobu.**
You(pl) see a crawling child and the clothes are dirty.
- 16) **Tewela towotowolohi hipeu.**
The standing children fall.

M1. Translate the following sentences into English:

- Apo onaduhuna babana aemi higobu.**
- Inapa unatowolo apo unamae.**
- Tewela hidewadewa ma natum dewadewana.**
- Unenae nu tahaya daodaona.**

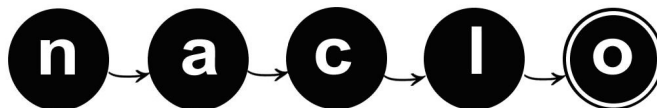
M2. Translate the following sentences into Tawala:

- You(pl) steal a swampy thing.*
- If the path is long, you(sg) will be dirty.*
- Limi's difficult child is staying but is not settled.*
- The children have substance.*

M3. Describe how to translate the English word "child" into Tawala.

M4. As noted in the introduction to this problem, reduplication is a linguistic process in which all or part of a word is repeated. Describe when reduplication is used in Tawala. You do not need to describe *how* the reduplicated form of a word is created; you only need to describe *when* reduplicated forms are used.

Make sure you record your answers in your Answer Sheets!



(N) Do I Care about Duikers? (1/2) [10 points]

Below are some nouns in Dagaare in their singular, plural, and interrogative forms, as well as their definitions. (The interrogative form is used to form questions.) Note that some entries have been left blank.

Singular	Plural	Interrogative	Definition
(a)	biri	bi-bo	<i>seed</i>
bɔŋɔ	bɔnnɪ	bɔŋ-bo	<i>lizard</i>
boŋo	(b)	boŋ-bo	<i>donkey</i>
dɛɛ	dɛɪɪ	(c)	<i>ladder</i>
dunduli	dundulo	dundul-bo	<i>maggot</i>
(d)	(e)	foloŋfug-bo	<i>lung</i>
gbɛbiri	gbɛbie	gbɛbi-bo	<i>toe</i>
gɔɔ	gɔɪɪ	gɔ-bo	<i>thorn</i>
gyili	gyile	gyil-bo	<i>xylophone</i>
(f)	(g)	ir-bo	<i>duiker</i>
ɪɪ	ɪɛ	ɪl-bo	<i>drum</i>
kpankpaŋbieli	kpankpaŋbiele	kpankpaŋbiel-bo	<i>elbow</i>
nimiri	nimie	nimi-bo	<i>eye</i>
nimisɔgɔ	nimisɔgɪɪ	nimisɔg-bo	<i>face</i>
(h)	(i)	nɔɪɪ-bo	<i>shoe</i>
nyagri	nyaga	nyag-bo	<i>intestine</i>
(j)	(k)	ŋmar-bo	<i>the Moon</i>
pɪɛ	pɪɪɪ	pɪ-bo	<i>roof</i>
(l)	pɔɪɪ	pɔ-bo	<i>stomach</i>
sebiri	sebie	sebi-bo	<i>bee</i>
(m)	tannɪ	taŋ-bo	<i>mountain</i>
valeŋvɔɔ	valeŋvɔɪɪ	valeŋvɔ-bo	<i>a kind of solitary wasp</i>
(n)	wɛgɪɪ	wɛg-bo	<i>log</i>
(o)	(p)	wul-bo	<i>bamboo flute</i>
(q)	yaga	yag-bo	<i>cheek</i>



(N) Do I Care about Duikers? (2/2)

Notes on pronunciation:

/i/ is the high front unrounded vowel with advanced tongue root (like English **heat**).

/ɪ/ is the high front unrounded vowel with unadvanced tongue root (like English **bit**).

/e/ is the mid front unrounded vowel with advanced tongue root (like English **mate**).

/ɛ/ is the mid front unrounded vowel with unadvanced tongue root (like English **bet**).

/u/ is the high back rounded vowel with advanced tongue root (like English **but**).

/ʊ/ is the high back rounded vowel with unadvanced tongue root (like English **hook**).

/o/ is the mid back rounded vowel with advanced tongue root (like English **yonder**).

/ɔ/ is the mid back rounded vowel with unadvanced tongue root (like English **caught**).

/a/ is the low front unrounded vowel with unadvanced tongue root (like English **mark**).

/ŋ/ is the velar nasal (like English **king**).

/kp/ and /gb/ are consonants.

The words “high,” “mid,” “low,” “front,” and “back” in these notes refer to the position of the tongue in the mouth, and the words “rounded” and “unrounded” refer to the shape of the lips. “Velar” refers to a specific part of the roof of the mouth.

Although Dagaare is a tonal language, for simplicity all tones (as well as vowel length marks) have been omitted.

N1. Fill in the blanks in the table.

N2. Describe your observations about Dagaare nouns.

Note: The Dagaare language is spoken by around 1.1 million Dagaaba people in Ghana and Burkina Faso. The Dagaaba are a farming people noted for their sophisticated music, usually performed in the form of xylophone (*gyil*) duets accompanied by drums (*iil*); another common form is solo melodies performed on bamboo flute (*wul*). The duiker is a small antelope native to Sub-Saharan Africa famous for its antisocial nature. Maggots are the larvae of flies, typically found in large groups on rotting organic material. The intestines (in humans, comprising the small intestine and the large intestine) are part of the digestive tract.

Make sure you record your answers in your Answer Sheets!



(O) Cameroonian Compounds (1/2) [10 points]

Vengo is the language of the Vengo ethnic group which consists of approximately 14,000 people living predominantly in the village Vengo in Cameroon.¹ The words in this problem are presented in the writing system of the language. The symbol ' represents the glottal stop (the consonantal sound we produce between the vowels in the English expression *uh-oh*).

Compounding is a common word formation process in Vengo. Similarly to English, compounds are variably written as one word, two words, or two hyphenated words in Vengo. Each Vengo word in the column on the left has its English translation (which is explicitly marked as a noun (*n.*) or a verb (*v.*) for clarity) somewhere in the column on the right, but the translations are in a scrambled order.

O1. Match each Vengo entry to its English translation.

- | | |
|-----------------|-----------------------------------|
| 1. tɔ | A. horn (<i>n.</i>) |
| 2. fɔ wi | B. head (<i>n.</i>) |
| 3. nyaa shinjii | C. stay (<i>v.</i>) |
| 4. tɔŋjii | D. hospital (<i>n.</i>) |
| 5. nyaa | E. climb/climb onto (<i>v.</i>) |
| 6. shinjii | F. gun (<i>n.</i>) |
| 7. fɔ | G. medicine (<i>n.</i>) |
| 8. nyaa yiko' | H. house (<i>n.</i>) |
| 9. wi | I. ladder (<i>n.</i>) |
| 10. shi | J. gunpowder (<i>n.</i>) |
| 11. nyaa ndaw | K. roof (<i>n.</i>) |
| 12. ŋjii | L. bull (<i>n.</i>) |
| 13. kɔ' | M. home (<i>n.</i>) |
| 14. ndaw | N. pet (<i>n.</i>) |
| 15. ŋjii-vəfɔ | O. horse (<i>n.</i>) |
| 16. vəkɔ' | P. animal (<i>n.</i>) |

¹ Speakers of the language call their village Vengo and their language Ghang Vengo [ɣáŋ vəŋóo]; however, the village is officially called Babungo and appears with this name on maps of Cameroon, and the language is, therefore, often referred to as Babungo.



(O) Cameroonian Compounds (2/2)

Below are some additional Vengo words, each one written next to its English translation. Note that (*adj.*) indicates that the word is an adjective.

dii	heavy (<i>adj.</i>)
ley	clean, clear, light (<i>adj.</i>)
nsi	ground (<i>n.</i>)
so	to wash (<i>v.</i>)

O2. From the English options Q. through W., choose the most likely meaning of each of the Vengo compound words a. through d., in light of the additional Vengo words and meanings given above.

- | | |
|----------------------|------------------------------|
| a. to ley | Q. housework (<i>n.</i>) |
| b. ko' fi nsi | R. wisdom (<i>n.</i>) |
| c. to dii | S. shame (<i>n.</i>) |
| d. so to | T. courthouse (<i>n.</i>) |
| | U. source (<i>n.</i>) |
| | V. festival (<i>n.</i>) |
| | W. forgiveness (<i>n.</i>) |

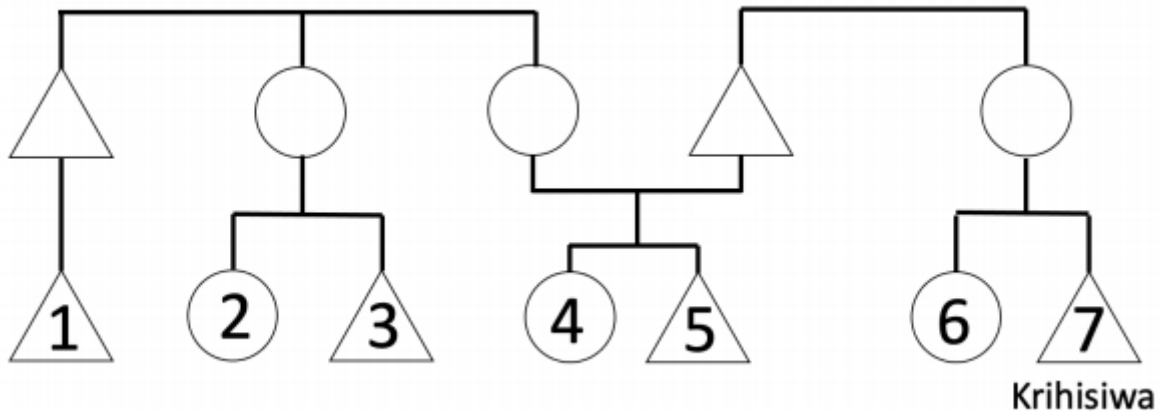
O3. What is the likely English word equivalent of the Vengo word **fi** ?

Make sure you record your answers in your Answer Sheets!



(P) Family Ties (1/2) [10 points]

The Yanomamö are a group of 35,000 indigenous people living in the Amazon Rainforest of South America. Below is a family tree of a Yanomamö family; a triangle denotes a male family member, whereas a circle denotes a female family member (in case you are unfamiliar with family trees, see the end of the problem for further notes on reading them).



A visitor has asked five of the seven children to briefly introduce their brothers, sisters, and cousins. The children's responses are below. Based on these responses, you will need to give the name of the child who corresponds to each numbered position in the family tree; note that we have already provided one match for you (Krihisiwa is 7 in the tree). You may assume that the children labeled 1 through 7 are the only ones in their generation in this family, and you may also assume that any numbers in the children's responses are exact (e.g., "I have two *soriwa*" means "I have exactly two *soriwa*").

Response from Davi:

I have one *suaboya*: Okori

I have two *soriwa*: Rerebawa and Krihisiwa

I have two *amiwa*: Yarima and Bushika

Response from Bushika:

I have two *eiwa*: Mukashe and Davi

I have one *amiwa*: Yarima

I have two *soriwa*: Rerebawa and Krihisiwa

Response from Rerebawa:

I have two *suaboya*: Yarima and Bushika

I have two *soriwa*: Mukashe and Davi



(P) Family Ties (2/2)

Response from Mukashe:

I have one *soriwa*: Rerebawa

I have no *suaboya*

I have two *amiwa*: Yarima and Bushika

Response from Krihisiwa:

I have one *amiwa*: Okori

I have one *suaboya*: Bushika

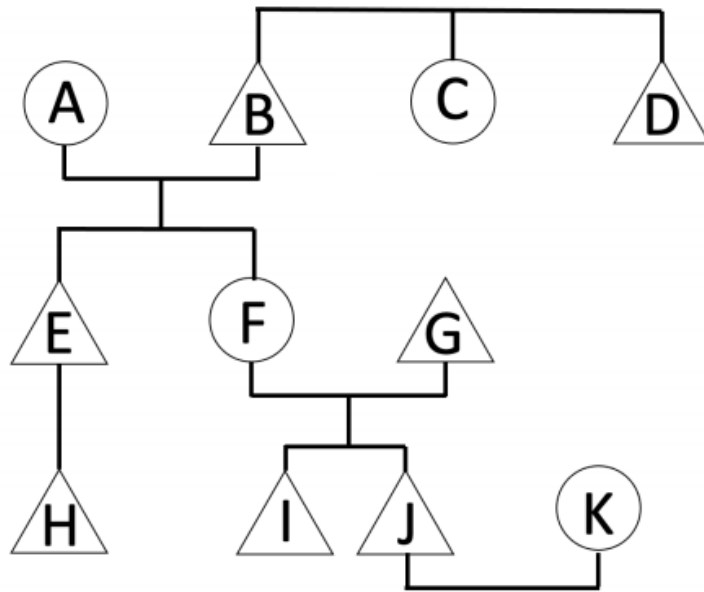
I have one *soriwa*: Davi

P1. Identify which children occupy the numbered positions of the family tree.

P2. Describe the meaning of the following words: *suaboya*, *soriwa*, *amiwa*, *eiwa*.

Note: For simplicity, we have used the singular form of all Yanomamö terms, even when those terms are referring to more than one person.

Example family tree: In case you are unfamiliar with family trees, below is an example. B, C, & D are siblings; so are E & F, and I & J. A & B are married, as are F & G and J & K. The parents of E & F are A & B. Similarly, the parents of I & J are F & G, and E is the father of H. To give a few examples of more distant relationships: A is the grandmother of H, I, & J; F is the aunt of H; and C is the sister-in-law of A.



Make sure you record your answers in your Answer Sheets!



(Q) A Stress Test (1/4) [10 points]

Although languages are different around the world, one feature that is found in many languages is stress. In these languages, each word features one or more syllables with extra emphasis (which linguists call stress).

Let's consider English for a moment. The word *scofflaw* means "a person who openly disregards the law." Try saying this word out loud. Even if you've never seen *scofflaw* before, you probably pronounced it correctly, with the emphasis on the first syllable (*SCOFF.law*), instead of on the second syllable (*scoff.LAW*).

Now take a deep breath and try saying these next two words: *galligaskins* ("loose-fitting breeches") and *ultracrepidarian* ("a person who expresses opinions on matters outside their expertise"). Once again, even if you've never seen these words before, you probably intuitively knew which syllables to stress. The correct pronunciations of these words are *GALL.i.*GAS.kins* and *UL.trä.CRE.pi.*DA.ri.an*.

How should you read this notation? There are 3 things to remember:

- We use periods to mark approximate syllable boundaries.
- We capitalize every stressed syllable.
- We place an asterisk (*) at the start of the syllable with primary stress. (That is, in every English word, there is usually one stressed syllable that receives more emphasis than the other stressed syllables; this is the syllable with primary stress.¹)

How is it that you intuitively know which syllables to stress, even for unfamiliar words? The answer is that English speakers must have some systematic way of assigning stress to novel words. In task **Q1**, we present a simplified version of one theory of how English stress assignment works.

Q1. Based on the data on the next page, fill in the blanks for the following stress assignment algorithm. Each blank corresponds to exactly one word. After filling in the blanks, your algorithm should correctly predict the stress for each of the 9 English words in the table on the next page. (Some blanks can be filled equally well by multiple answers. You only need to provide one correct answer).

1. Assign stress to every **(a)** _____-numbered syllable.²
2. If step (1) made the **(b)** _____ syllable stressed, un-stress it UNLESS the word is **(c)** _____ syllable(s) long.
3. Assign primary stress to the **(d)** _____ syllable farthest to the **(e)** _____.

¹ In case you're wondering how to tell which syllable in a word has primary stress, one technique is called the "Lassie test." To use this technique, pretend that the word is the name of a dog and that you want to call the dog inside. Whichever syllable you elongate when you call out the dog's name is the syllable with primary stress. For example, if your dog were named *Ultracrepidarian*, you would call out something like "Ultracrepi-DAAAA-ri-an!"

² Note that we use the term *odd-numbered syllables* to refer to the first syllable, third syllable, fifth syllable, etc. We use the term *even-numbered syllables* to refer to the second syllable, fourth syllable, sixth syllable, etc.



(Q) A Stress Test (2/4)

Here is the relevant data for Q1.

Word	Stress
elephant	<i>*E.le.phant</i>
crush	<i>*CRUSH</i>
vitamin	<i>*Vi.ta.min</i>
illustration	<i>IL.lu.*STRA.tion</i>
dime	<i>*DIME</i>
scofflaw	<i>*SCOFF.law</i>
galligaskins	<i>GALL.i.*GAS.kins</i>
ultracrepidarian	<i>UL.tra.CRE.pi.*DA.ri.an</i>
supercalafragilisticexpialidocious	<i>SU.per.CA.li.FRA.gi.LI.stic.EX.pi.A.li.*DO.cious</i>

Q2. Stress assignment in English is a complex topic; the algorithm in **Q1** only covers some of the factors that affect English stress. Based on the conversation below (which was annotated for stress by a human), what are some further properties that might need to be added to make the algorithm properly handle English?

Notes: You should only mention factors that are illustrated in the conversation below. If a word has no capital letters in it, that means it has no stress.

Person A: *i'm *HOP.ing to ex.*PORT my *PAINT.ings. *EACH *ONE *SHOWS a *COM.mon *OB.ject in a *STRANGE *SET.ting.*

Person B: *i ob.*JECT to *THAT. we should *IM.port *ART, not *EX.port it!*

Person A: *well, i just *GOT a *PER.mit from the *CUS.toms *OFF.i.cer. she *SAYS that *ART can be an *EX.cell.ent *EX.port.*

Person B: *if *SHE per.*MITS it, then i *GUESS *I must per.*MIT it too.*



(Q) A Stress Test (3/4)

Not all languages stress their words in the same way that English does. However, it turns out that we can still use the same basic algorithm for many other languages; we just need to introduce a few options in the statement of this algorithm. Here is the more general algorithm:

1. Start at the **[left / right]** edge of the word. **[Skip / don't skip]** the syllable at that edge and then assign stress to **[only the first / every alternating]** syllable that you encounter.
2. If the word is longer than one syllable and if step (1) made the word's final syllable stressed, **[leave it that way / un-stress it]**.
3. Assign primary stress to the **[leftmost / rightmost]** stressed syllable.

We refer to these five bolded options as *parameters*. By choosing the right set of parameters each time, we can determine how to stress words in a wide variety of languages!

Q3. For the six languages presented below, examine the examples given to determine which stress assignment parameters the language obeys.³ Select the correct values for each of the parameters mentioned above. (For some languages, there may be multiple correct answers. You only need to provide one correct answer. For simplicity, we have simplified the spellings of some of the example words.)

Mapudungun Word	Stress
wule	wu.*LE
tsipanto	tsi.*PAN.to
elumuyu	e.*LU.mu.YU
eluaenew	e.*LU.a.E.new
kimufaluwulay	ki.*MU.fa.LU.wu.LAY

Maranungku Word	Stress
tiralk	*TI.ralk
merepet	*ME.re.PET
yangarmata	*YANG.ar.MA.ta
langkarateti	*LANG.ka.RA.te.TI
welepemanta	*WE.le.PE.ne.MAN.ta

Weri Word	Stress
ngintip	ngin.*TIP
kulipu	KU.li.*PU
uluamit	u.LU.a.*MIT
akunetepal	A.ku.NE.te.*PAL

Mansi Word	Stress
same	*SA.me
atenel	*A.te.nel
omatanel	*O.ma.TE.nel
pocaganelnel	*PO.ca.GA.nel.nel

³ In practice, it is possible to do this automatically. In fact, one of the desirable properties of parameter-based linguistic theories is that they allow a learner (such as a baby acquiring the language, or a computer model being trained on sentences) to learn properties of the language based on just a few examples. This is because the set of parameters greatly constrains the set of possible systems that the learner has to distinguish between.



(Q) A Stress Test (4/4)

Warao Word	Stress
yapurukitanehase	<i>YA.pu.RU.ki.TA.ne.*HA.se</i>
nahoroahakutai	<i>NA.ho.RO.a.HA.ku.*TA.i</i>
yiwaranae	<i>yi.WA.ra.*NA.e</i>
enahoroahakutai	<i>e.NA.ho.RO.a.HA.ku.*TA.i</i>

Comalapa Kaqchikel Word	Stress
abex	<i>a.*BEX</i>
tinamith	<i>ti.na.*MITH</i>
nubanobel	<i>nu.ba.no.*BEL</i>
shintshuxirisax	<i>shin.tshu.xi.ri.*SAX</i>

Q4. Unfortunately, this type of algorithm does not work for all languages. In **Q2**, we already saw some examples of how it fails to capture certain nuances of English. Below is another language (Selkup) where the algorithm fails. Describe how stress is assigned in this language.

Selkup Word	Stress
qumooqi	<i>qu.*MOO.qi</i>
qumooqlilii	<i>qu.mooq.li.*LII</i>
quminik	<i>*QU.mi.nik</i>
amirna	<i>*A.mir.na</i>
uucikkak	<i>*UU.cik.kak</i>
qolcimpati	<i>*QOL.cim.pa.ti</i>
uucoomit	<i>uu.*COO.mit</i>

Notes on Featured Languages:

1. Mapudungun, also known as Mapuche, is an Araucanian language spoken by approximately 250,000 native speakers in Chile and Argentina.
2. Maranungku is a dialect of Marranj, an Australian Aboriginal language spoken in Northern Australia.
3. Weri is a Kunimaipa language spoken by approximately 14,000 native speakers in Papua New Guinea.
4. Mansi is a Uralic language spoken by about 1000 speakers, most of whom are situated around Russia's Ob River and its tributaries.
5. Warao is spoken by approximately 33,000 native speakers in Venezuela, Suriname, and Guyana. It is a language isolate, meaning that it is not known to be related to any other language.
6. Comalapa Kaqchikel is a variety of Kaqchikel, a Quichean language of the Mayan family spoken by about 450,000 people in Guatemala.
7. Selkup is a member of the Uralic language family spoken by about 1000 people in Siberia.

Make sure you record your Answers in your Answer Sheet!



(R) GloVe Compartment (1/2) [5 points]

How can we represent the meaning of a word in a way that computers can understand? A popular solution is to represent each word as a vector (a list of numbers). For instance, the word *dog* might be represented as [0.7, 1.9, -4.3, 5.6, 0.0, -0.5]. Computers are great at processing numbers, so vectors are a very computer-friendly way to represent information. How do we decide which numbers should go in the vector for a given word? One popular approach was described by the linguist John Firth in 1957: “You shall know a word by the company it keeps.” That is, the vector for a given word is meant to encode which other words tend to occur near this word.

One successful recent approach that builds on this idea is GloVe (Global Vectors for word representation), in which the vectors encode *differences* between words in terms of the other words that occur near them. For example, suppose we were trying to understand what the words *cake* and *cakes* mean. The word *birthday* appears near both of these words, so it is not helpful in distinguishing them. However, if we look at the words *is* and *are*, we would see that *is* appears more often near *cake* than near *cakes*, while *are* appears more often near *cakes* than *cake*. Thus, *is* and *are* can help us distinguish *cake* from *cakes*! Furthermore, we can observe that *candle* and *candles* have a similar relationship to *is* and *are*, so we can conclude that the difference between *cake* and *cakes* is the same as the difference between *candle* and *candles*!

Researchers at NACLO Labs have found some four-dimensional vectors that represent a few English words almost as well as GloVe vectors do! Here are two of the vectors they found:

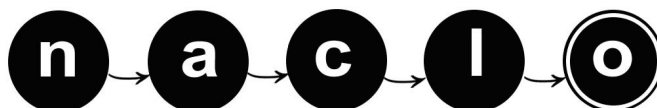
man: [0.5, 0.9, 0.3, 0.3]

woman: [0.5, 0.9, 0.1, 0.5]

daughter: [0.5, 0.7, 0.2, 0.3]

The researchers had more representations, but they somehow mixed up the vectors and the words that they correspond to! Below are some English words and their corresponding vectors, in no particular order:

- | | |
|--------------------------|-------------|
| 1. [0.5, 0.9, 0.2, 0.4] | A. girl |
| 2. [0.5, 0.7, 0.4, 0.1] | B. queen |
| 3. [0.5, 0.9, 0.3, -0.5] | C. prince |
| 4. [0.5, 0.7, 0.2, 1.1] | D. boy |
| 5. [0.5, 0.8, 0.9, 1.3] | E. father |
| 6. [0.5, 0.8, 0.9, 0.5] | F. mother |
| 7. [0.5, 0.8, 0.7, 1.5] | G. person |
| 8. [0.5, 0.8, 0.8, 1.4] | H. king |
| 9. [0.5, 0.9, 0.1, -0.3] | I. ruler |
| 10. [0.5, 0.8, 0.7, 0.7] | J. princess |
| 11. [0.5, 0.7, 0.4, 0.9] | K. son |



(R) GloVe Compartment (2/2)

R1. Match the vectors 1 through 11 to the English words they represent.

Here are a few more word representations, along with their English equivalents, in no particular order. Note that a barometer is a tool for measuring air pressure, while a millibar is a unit of air pressure.

- | | |
|------------------------------|--------------|
| 12. [0.3, 0, -0.6, -0.1] | L. clock |
| 13. [0.2, -0.2, -0.3, -0.2] | M. first |
| 14. [0.4, 0, -0.4, -0.4] | N. second |
| 15. [-0.6, 0.6, 0.2, -0.8] | O. one |
| 16. [-0.6, -0.2, -0.4, -0.4] | P. three |
| 17. [0.4, 0.8, -0.4, -0.4] | Q. third |
| 18. [1.6, 0, 1.8, 0.6] | R. two |
| 19. [0, 0, 0, -0.4] | S. barometer |
| 20. [-0.6, -0.4, -0.2, -0.4] | T. half |
| 21. [1.8, 0, 1.6, 0.4] | U. millibar |

R2. Match the vectors 12 through 21 to their English equivalents. There are two possible answers; either one will receive full credit.

R3. The word *third* actually has two meanings that are relevant to the problem. The vector that is given for *third* above is the average of the vectors that would represent these two meanings. Suppose English used two different words for these two meanings, rather than a single word. What would you expect the vector to be for each meaning?

R4. Below are the two vectors found for the words *doctor* and *nurse*. Even though these words are gender-neutral, the method of defining a word based on the words that occur near it also captures general trends and biases that are in the texts which were used to determine what words occur near each other. Identify which of these vectors goes with which word, and explain how the vectors encode gender-related properties of the corresponding words.

- a) [0.5, 1.3, 0.3, 1.7]
- b) [0.5, 1.3, 0.5, 1.5]

Make sure you record your answers in your Answer Sheets!



(S) Peace Only (1/2) [15 points]

Jamsay is a Dogon language spoken in Mali and Burkina Faso by around 130,000 people. Its name derives from a common greeting, *jâm sǎy*, meaning “peace only.”

Here are some sentences in Jamsay, along with their English translations.

- jèmèm dònjúm sǎy úró bé kùⁿ ∴ fú: é:tóγòba**
Only thin blacksmiths see all the houses.
- mò:ká: kùⁿ sóγótìm sábù gùgùγⁿím ójù bàḡátóγòba**
I locked the door because thieves hide on the road.
- jémèn kùⁿ sùrgòm ∴ fú: jè:rèjè sábù wó gùgùn námátìm**
The blacksmith has already criticized all weavers because I stepped on his melon.
- ànsà:ràn pèγⁿín kùⁿ gùgùn gôn bé ∴ fú: ñé:tóγò**
The old European eats all crooked melons.
- wó ùrò kàná sóγórjèw là:**
You have already unlocked his new house, haven't you?
- ójú kùⁿ námárⁿátìm sábù ijù téré é:tìm**
I took my foot off the road because I saw a fast dog.
- nìwⁿé ùrò pèγⁿ nánárⁿátóγò**
A cat remembers an old house.
- ijú bé ∴ nìwⁿè tēγⁿ bé ∴ sǎy ànà dìgétóγòw**
You follow only dogs and small cats in the village.
- àná kùⁿ jè:rétìw sábù ànsá:rám ∴ sùrgòm ∴ ijú bé kùⁿ sánátìba là:**
You criticized the village because Europeans and weavers fenced in the dogs, didn't you?
- gùgùγⁿín gónìn kùⁿ úró kùⁿ sánárⁿájè là:**
The crooked thief has already taken down the fence around the house, hasn't he?
- sùrgòn téréw èmé úró páγátóγò**
A fast weaver ties sorghum in the house.
- mǎngòlò bé kùⁿ ñé:jèm**
I have already eaten the mangoes.

Note that the diacritics ´, ` , ˇ , ^ represent high, low, rising, and falling tones respectively. The symbol ∴ after a word means that the word is pronounced with “dying-quail intonation,” an exaggerated prolongation of the tone accompanied by an exaggerated drop in pitch.¹ The symbol : after a vowel signifies length, and ⁿ signifies nasalization of the previous sound.

¹ Jeffrey Heath, who studied Jamsay, writes: “The dying-quail intonation contour reminds me of the prosodic pattern of American high-school cheerleaders calling out the letters of their school at sporting events, through their bullhorns (“give me an A...., give me a B...”).”



(S) Peace Only (2/2)

S1. Translate the following Jamsay sentences into English:

- a. wó nìwⁿè kàná bé náhájèm
- b. sùrgôn kùⁿ∴ ànsá:rán kùⁿ∴ màngòlò dǎḡ bé é:tì sábu bànátìw
- c. gùgùⁿím kùⁿ ijù gôn úró jè:rétóyò là:

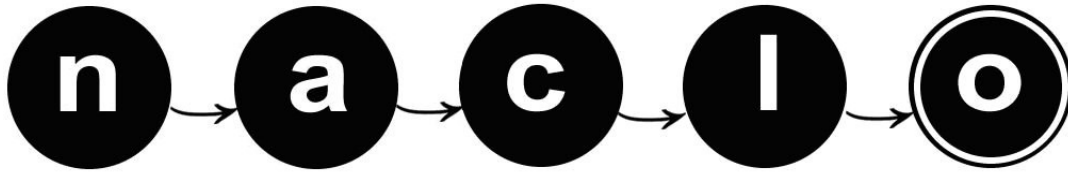
S2. Translate the following English sentences into Jamsay:

- a. A small blacksmith forgot his sorghum and his cat.
- b. Peace follows the old dog, doesn't it?
- c. You have already untied all the melons because you ate the mango.

S3. Describe your observations about Jamsay grammar.

Make sure you record your answers in your Answer Sheets!





The North American Computational Linguistics Open Competition
www.nacloweb.org

Answer Sheets

REGISTRATION NUMBER					

Name: _____

Contest Site: _____

Site ID: _____

City, State: _____

Grade: _____

Please also make sure to **write your registration number and your name on each page of the Answer Sheets**, and **turn in all pages of the Answers Sheets** even if you have left some blank .

SIGN YOUR NAME BELOW TO CONFIRM THAT YOU WILL NOT DISCUSS THESE PROBLEMS WITH ANYONE UNTIL THEY HAVE BEEN OFFICIALLY POSTED ON THE NACLO WEBSITE IN APRIL.

Signature: _____

YOUR NAME:

REGISTRATION #

Answer Sheets (1/10)

(J) A Vintage Sound System

J1. Write the corresponding word for each numbered blank:

(1) _____ (2) _____ (3) _____ (4) _____

(5) _____ (6) _____ (7) _____

J2. Fill in the letter of the correct option in the boxes below:

1. 2. 3. 4. 5. 6.

J3. Fill in the empty cells of the table below:

Old Chinese	Middle Chinese	English Translation
<i>kraj</i>	(1)	'bridle'
<i>n^ɕə</i>	(2)	'violent'
<i>rak</i>	(3)	'female servant'
<i>pre</i>	(4)	'pole'
<i>b^ɕəʔ</i>	(5)	'double'
<i>m^ɕaʃs</i>	(6)	'dust'

J4. Fill in the empty cells of the table below:

Old Chinese	Middle Chinese	English Translation
(7)	<i>ŋjē</i>	'to make a sacrifice to the deity of the soil'
(8)	<i>tōk</i>	'to obtain'



YOUR NAME:

REGISTRATION #

Answer Sheets (2/10)

(J) A Vintage Sound System (continued)

J5. Explain the sound changes that occurred between Old and Middle Chinese:

(K) Putting a Place to a Name

- K1. 1. 2. 3. 4. 5. 6. 7. 8. 9.
10. 11. 12. 13. 14. 15. 16. 17. 18.
19. 20. 21. 22. 23. 24. 25. 26.

K2. Name of the region:

(L) Is This Problem Intelligible?

L1. For each of the sentences, circle the language that it is in. Then, write the letter of its English translation.

- | | | | | | | | |
|----|-------------|-------------|--------------------------|----|-------------|-------------|--------------------------|
| 1. | <i>Hawu</i> | <i>Dhao</i> | <input type="checkbox"/> | 2. | <i>Hawu</i> | <i>Dhao</i> | <input type="checkbox"/> |
| 3. | <i>Hawu</i> | <i>Dhao</i> | <input type="checkbox"/> | 4. | <i>Hawu</i> | <i>Dhao</i> | <input type="checkbox"/> |
| 5. | <i>Hawu</i> | <i>Dhao</i> | <input type="checkbox"/> | 6. | <i>Hawu</i> | <i>Dhao</i> | <input type="checkbox"/> |
| 7. | <i>Hawu</i> | <i>Dhao</i> | <input type="checkbox"/> | 8. | <i>Hawu</i> | <i>Dhao</i> | <input type="checkbox"/> |



YOUR NAME:

REGISTRATION #

Answer Sheets (3/10)

(L) Is This Problem Intelligible? (continued)

L1. (continued)

9. *Hawu* *Dhao*

10. *Hawu* *Dhao*

11. *Hawu* *Dhao*

12. *Hawu* *Dhao*

L2. Write the English translations of each sentence below:

(a)

(b)

(c)

(d)

(e)

L3. Write the Hawu and Dhao translation of each sentence:

(a) *Hawu:*

Dhao:

(b) *Hawu:*

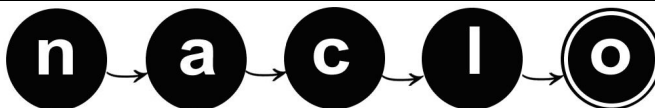
Dhao:

(c) *Hawu:*

Dhao:

L4. Dhao relative of **pedutu**:

L5. Are the languages mutually intelligible? Explain:



YOUR NAME:

REGISTRATION #

Answer Sheets (4/10)

(M) The Speech Has No End

M1. Translate the sentences into English:

a.

b.

c.

d.

M2. Translate the sentences into Tawala:

a.

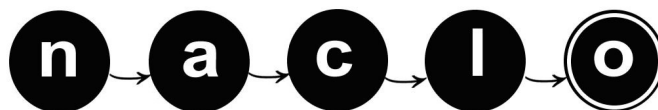
b.

c.

d.

M3. Describe how one would translate “child” in Tawala:

M4. When are reduplicated forms used?



YOUR NAME:

REGISTRATION #

Answer Sheets (5/10)

(N) Do I Care about Duikers?

N1. Write the Dagaare word that goes in each blank:

(a)	<input type="text"/>	(b)	<input type="text"/>	(c)	<input type="text"/>
(d)	<input type="text"/>	(e)	<input type="text"/>	(f)	<input type="text"/>
(g)	<input type="text"/>	(h)	<input type="text"/>	(i)	<input type="text"/>
(j)	<input type="text"/>	(k)	<input type="text"/>	(l)	<input type="text"/>
(m)	<input type="text"/>	(n)	<input type="text"/>	(o)	<input type="text"/>
(p)	<input type="text"/>	(q)	<input type="text"/>		

N2. Describe your observations about Dagaare nouns:



YOUR NAME:

REGISTRATION #

Answer Sheets (6/10)

(O) Cameroonian Compounds

O1. Write the letter of the corresponding English translation for each number:

- 1) 2) 3) 4) 5) 6) 7) 8)
9) 10) 11) 12) 13) 14) 15) 16)

O2. Write the letter, from Q to W, of the English word that is the most likely meaning for each Vengo word:

- a. b. c. d.

O3. What is the likely English word equivalent of **fi** ?

(P) Family Ties

P1. Write the corresponding name for each number:

- 1) 2) 3)
4) 5) 6)

P2. Explain the meaning of the following words:

suaboya:

soriwa:

amiwa:

eiwa:



YOUR NAME:

REGISTRATION #

Answer Sheets (7/10)

(Q) A Stress Test

Q1. Write the appropriate word in each of the blanks:

(a) (b) (c)
(d) (e)

Q2. What are some further properties that should be added to the algorithm in order to handle English?

Q3. For each language, specify the value of the parameter in the correct cell of the table:

	left / right	skip / don't skip	only the first / every alternating	leave it that way / un-stress it	leftmost / rightmost
Mapudungun					
Maranungku					
Weri					
Mansi					
Warao					
Comalapa Kaqchikel					



YOUR NAME:

REGISTRATION #

Answer Sheets (8/10)

(Q) A Stress Test (continued)

Q4. Describe how stress is assigned in Selkup:

(R) GloVe Compartment

R1. Write the letter of the English word corresponding to the vector with each number:

- 1) 2) 3) 4) 5) 6)
7) 8) 9) 10) 11)

R2. Write the letter of the English word corresponding to the vector with each number:

- 12) 13) 14) 15) 16)
17) 18) 19) 20) 21)

R3. What are the two vectors for each of the meanings of the word "third"?

Vector:	<input type="text"/>	Definition:	<input type="text"/>
Vector:	<input type="text"/>	Definition:	<input type="text"/>



YOUR NAME:

REGISTRATION #

Answer Sheets (9/10)

(R) GloVe Compartment (continued)

R4. Identify which vector corresponds to which word, and explain how the vectors encode gender-related notions of these words:

(S) Peace Only

S1. Translate the sentences into English:

a. wó nìwⁿè kàná bé nánájèm

b. sùgôn kùⁿ· ànsá:rán kùⁿ· màngòlò dǝŋ bé é:tì sábù bàŋátìw

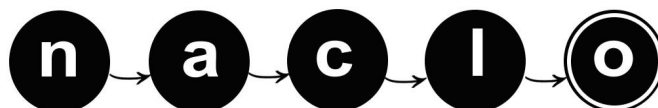
c. gùgùⁿím kùⁿ ijù gôn úrò jè:rétóyò là:

S2. Translate the sentences into Jamsay:

a. A small blacksmith forgot his sorghum and his cat.

b. Peace follows the old dog, doesn't it?

c. You have already untied all the melons because you ate the mango.



YOUR NAME:

REGISTRATION #

Answer Sheets (10/10)

(S) Peace Only (continued)

S3. Describe your observations about Jamsay grammar:

