

NACLO thanks the following for their generous contributions:



Carnegie Mellon University
Language Technologies Institute



and Individual Donors

The Nineteenth Annual

**North American
Computational
Linguistics
Open
Competition**

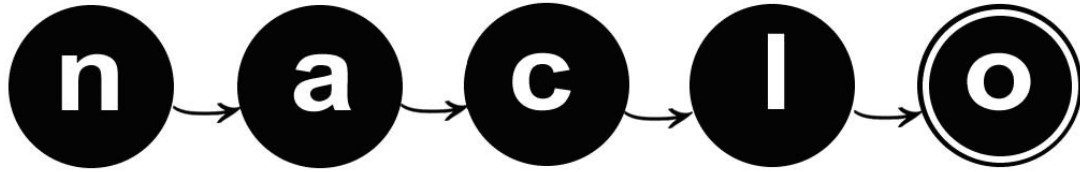
2025

www.naclo.org

**Invitational Round
March 13, 2025**

Serious language puzzles that are surprisingly fun!

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



Welcome to the nineteenth annual North American Computational Linguistics Open Competition! We (the NACLO organizers) are excited for you to participate in this unique event. 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 eight problems, labeled I to P.
2. Follow the facilitators' instructions carefully.
3. If you want clarification on any of the problems, talk to a facilitator. The facilitator will consult with the jury before answering.
4. You may not discuss the problems with anyone except as described in items 3 & 10.
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 at the end of this booklet. **Only the Answer Sheets will be graded.**
7. Write your name and registration number on each page of the Answer Sheets.
Here is an example: Jessica Sawyer #850
8. 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.
9. Don't be discouraged if you don't finish everything! If we have done our job well, very few people will solve all these problems completely in the time allotted.
10. **DO NOT DISCUSS THE PROBLEMS UNTIL THEY HAVE BEEN POSTED ONLINE! THIS MAY BE A COUPLE OF MONTHS AFTER THE END OF THE CONTEST.**

Oh, and have fun!

NACLO 2025 Organizers and Credits (1/2)

NACLO Chair:

Lori Levin — Carnegie Mellon University

Executive Committee:

Lori Levin — Carnegie Mellon University

Patrick Littell — National Research Council Canada

Daniel Lovsted — University of Ottawa

Tom McCoy — Yale University

Logistics Committee:

Nathan Dummitt (Social Media Subchair) — Columbia Grammar and Preparatory School

Elena Filatova (Outreach and University Subchair) — City University of New York Graduate Center

Henry Gustafson — California Institute of Technology

Shuli Jones — Google

Arul Kolla — Massachusetts Institute of Technology

Lori Levin — Carnegie Mellon University

Daniel Lovsted (Canada Coordinator) — University of Ottawa

Cerulean Ozarow (High School Subchair) — Brown University

Luke Sawczak (Website Subchair) — Toronto District Christian High School

Marten van Schijndel (Grading Subchair) — Cornell University

Gloria Yip — State University of New York at Stony Brook

Logistics Committee Co-Chairs:

Shuli Jones — Google

Cerulean Ozarow — Brown University

Program Committee:

Kunaal Chandrashekar — University of Toronto

Ryan Chi — Stanford University

Brandon Cho — Princeton University

Angela Hou — Princeton University

Ken Jiang — University of Waterloo

Arul Kolla — Massachusetts Institute of Technology

Lori Levin — Carnegie Mellon University

Patrick Littell — National Research Council Canada

Daniel Lovsted — University of Ottawa

Tom McCoy — Yale University

Benjamin Yang — California Institute of Technology



NACLO 2025 Organizers and Credits (2/2)

Special thanks to:

High School and University Site hosts in Canada and the USA

Booklet Editors:

Daniel Lovsted and Tom McCoy

Problem Credits:

(I) Arul Kolla

(J) Simi Hellsten

(K) Tom McCoy

(L) Evan Hochstein

(M) Tom Payne, Patrick Littell, and Daniel Lovsted

With thanks to: Ivan Derzhanski

Language consultant: Daniel Herlihy

(N) Arul Kolla

(O) Angela Hou

(P) Alex Wade

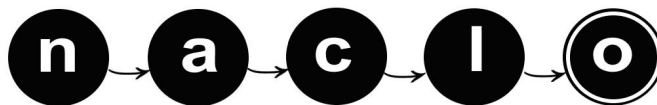
Image Credits:

Tablet in problem (I), page (1/2): British Museum, <https://web.archive.org/web/20090123200156/http://www.mesopotamia.co.uk/writing/explore/stylus.html>

Alligator in problem (K), page (2/4): Wikimedia Commons, [https://commons.wikimedia.org/wiki/File:Alligator_\(PSF\).svg](https://commons.wikimedia.org/wiki/File:Alligator_(PSF).svg)

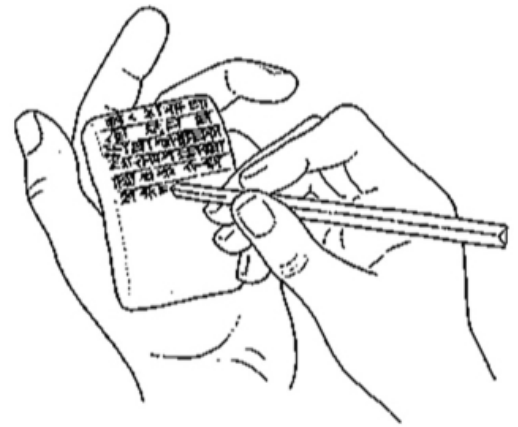
We are grateful for the support of many institutional and individual donors who make this contest possible.

All material in this booklet © 2025, North American Computational Linguistics Open Competition and the authors of the individual problems. Please do not copy or distribute without permission.



(I) A Handheld Tablet (1/2) [10 Points]

Before the advent of modern printing technology, writing was a laborious process that leveraged whatever tools and materials were available. The distinctive wedge-like shapes of Cuneiform writing arose from the shape of the stylus that was pressed into clay. The stylus was made from river reeds that were split lengthwise in a way that gave them a triangular cross section. The image on the right shows what writing in Cuneiform might have looked like.

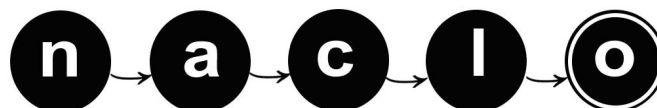


Cuneiform has been used to write several languages over its approximately three-thousand-year lifespan: it was first developed by the **Sumerians** and was later used by the **Babylonians** and even later the **Assyrians** in the region of Mesopotamia.

Below is a glyph bank containing several glyphs in Sumerian, early Babylonian, late Babylonian, and Assyrian, in random order.

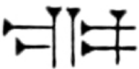


GLYPH BANK

A.	B.	C.	D.	E.	F.
G.	H.	I.	J.	K.	L.
M.	N.	O.	P.	Q.	R.



(I) A Handheld Tablet (2/2)

I1. Fill in the gaps **(1)** - **(18)** on your Answer Sheet with a letter (A-R) from the glyph bank on the previous page. Fill in **(19)** with an English word. Do not fill in the shaded cell.

<u>Sumerian</u>	<u>Early Babylonian</u>	<u>Late Babylonian</u>	<u>Assyrian</u>	<u>English</u>
<i>(do not fill in; see I2)</i>	(1)	(2)	(3)	<i>man</i>
(4)	(5)	(6)	(7)	<i>king</i>
(8)	(9)	(10)	(11)	<i>house</i>
(12)	(13)	(14)		<i>prince</i>
(15)	(16) <i>(all three languages share the same glyph)</i>			<i>bird</i>
(17)	(18)			(19)

I2. Using the information in the table above, you can make a very reasonable prediction as to how the word *man* was written in Sumerian. Draw the Sumerian glyph for *man*.

I3. The city of Nineveh was the capital of the Assyrian empire and the world's largest city for several decades. In Assyrian the name Nineveh was written with the following glyph:



Although the meaning of the name Nineveh is debated, a popular belief is that it can be literally translated to English as “**(a)** of **(b)**”. What are **(a)** and **(b)**?

I4. Describe how the script changed over time from Sumerian to Early Babylonian, from Early Babylonian to Late Babylonian, and from Late Babylonian to Assyrian. You do not need to justify your explanation (that is, you don't need to say why you think your explanation is correct).



(J) Nganasan Idol (1/1) [15 Points]

Nganasan is a Uralic language spoken by the Nganasan people. Traditionally they have been the northernmost people in Eurasia, living in the Taimyr Peninsula in Russia. Today approximately 470 Nganasan people live there, about 50 of whom speak the language.

Below are some Nganasan words written in a phonetic transcription, with their English translations given in alphabetical order. Note that **ð** = *th* as in *this*, **d'** ≈ *j* as in *judge*, **j** = *y* as in *yes*; **č**, **ń**, **ŋ**, and **ʔ** are consonants; **ə**, **ü**, and **ɨ** are vowels. Traditional Nganasan mythology includes several **ŋuəʔ** or gods, such as **bəndubtuə ŋuə** and **ńilɨtɨə ŋuə**. The Nganasan people made small wooden idols (religious statues), some of which would be kept in family homes, and offered them gifts to help please the gods.

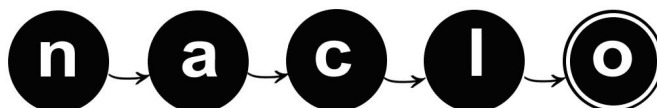
- | | |
|-------------------------|------------------------------------|
| 1. bad'ačimi | A. <i>airplane</i> |
| 2. bad'aʔbəbə | B. <i>boat</i> |
| 3. čimiʔ | C. <i>branches</i> |
| 4. d'üðübad'aʔ | D. <i>churches</i> |
| 5. d'üəbəbə | E. <i>earring</i> |
| 6. d'ütü | F. <i>fire</i> |
| 7. huaad'ütüʔ | G. <i>gunpowder sack</i> |
| 8. kojkəʔ | H. <i>hand</i> |
| 9. kojküʔmaðəʔ | I. <i>hook</i> |
| 10. koubasa | J. <i>idols</i> |
| 11. maðəʔ | K. <i>purse</i> |
| 12. ŋənduj | L. <i>rings (for your fingers)</i> |
| 13. təiriəŋənduj | M. <i>steamboats</i> |
| 14. tuj | N. <i>teeth</i> |
| 15. tuubitiðəʔ | O. <i>tents</i> |
| 16. tuuŋəntuʔ | P. <i>weapons</i> |

J1. Determine the correct correspondences.

J2. Translate into English: 17. **bəbəʔ** 18. **kojkə** 19. **kouʔ**

J3. Translate into Nganasan: 20. *boats* 21. *(piece of) metal*

J4. Suggest a Nganasan translation of the phrase *forest fire*.



(K) Going By the Numbers (1/4) [15 Points]

In artificial intelligence, it is common to represent a word with a vector — that is, a list of numbers. For example, the word *wandered* might be represented with the vector shown below:

$$\textit{wandered} = \begin{array}{|c|c|c|} \hline 3 & 1 & -1 \\ \hline \end{array}$$

How can we make sense of this vector? One approach is to use something called a classifier, which takes in a vector and determines whether that vector belongs to a particular category. As a starting example, we will consider a classifier that makes the following decisions:

Words that are labeled **YES** by the classifier: *wandered, thought, noticed, cheered*

Words that are labeled **NO** by the classifier: *wandering, thinking, noticing, cheering, wanders, thinks, notices, cheers*

K1. What is the category of words that receive a label of YES from this classifier? Write your answer in your Answer Sheet.

Let's take a closer look at how the classifier makes its decisions. The classifier is defined by another vector that we will call w . For our example, w looks like this:

$$w = \begin{array}{|c|c|c|} \hline 1 & -2 & 0 \\ \hline \end{array}$$

A classifier uses its vector w to make a decision as follows:

If $w \cdot \textit{word} > 0$, say **YES**.

Otherwise, say **NO**.

The dot inside " $w \cdot \textit{word}$ " stands for an operation that takes in two vectors and returns a single number. It's easiest to illustrate this operation with an example, and then we'll define it in words later. If we use the vectors shown above for w and *wandered*, then $w \cdot \textit{wandered}$ can be computed as shown:

$$w \cdot \textit{wandered} = 1*3 + -2*1 + 0*-1 = 3 - 2 + 0 = 1$$

That is, we add together the first number of the first vector (1) times the first number of the second vector (3), plus the second number of the first vector (-2) times the second number of the second vector (1), plus the third number of the first vector (0) times the third number of the second vector (-1), which gives the number 1 (if the vectors were longer than 3 numbers, we would keep going until we had used up all the numbers in the vectors). Since 1 is greater than 0, the classifier would say **YES**.



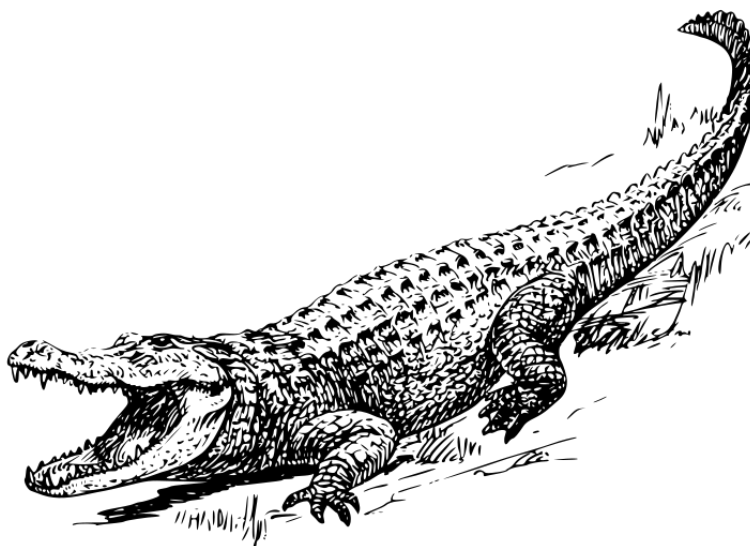
(K) Going By the Numbers (2/4)

On the last page of this problem, we have listed word vectors for 22 words (*alligators* through *visited*). These new vectors contain 7 numbers. We have also listed the w vectors for 9 different classifiers, labeled w_1 through w_9 .

K2. What is $w_4 \cdot \textit{squirrel}$? Your answer should be a single number.

Your task will now be to figure out what type of information each of these nine classifiers (w_1 through w_9) focuses on. But before we get into that, we should say a little more about where our word vectors came from. These vectors were produced by a computer system that was trained to predict the next word in a piece of text. For example, given the input *the books on the _____*, it should try to guess what word could go in the blank; it might guess *table* or *chair*. Next-word prediction lies at the heart of many popular systems in artificial intelligence, such as ChatGPT.

A next-word prediction system is usually not programmed by humans. Instead, it is given some data and learns how to predict words by observing patterns in the data. The prediction process uses word vectors to represent the text that is used as input, and while it is learning the system selects values for the word vectors that will help it perform its task. Therefore, we should expect the word vectors to encode information that is useful for predicting what word could come next in particular contexts. As one example, consider w_1 . If you apply w_1 to the 22 words, it turns out that w_1 says **YES** for *alligators*, *artists*, *envelopes*, *lawyers*, *oranges*, *parrots*, *plates*, *salads*, and *sticks* (you can check this by computing $w_1 \cdot \textit{word}$ for each word). In other words, it seems that w_1 tells us whether each word is a plural noun. How does this relate to next-word prediction? There are some situations where the next word is likely to be a plural noun, such as in the context *I found these _____*. Therefore, to handle such contexts, the system should be able to recognize which words are plural nouns. The information that is identified by w_1 does exactly that.



(K) Going By the Numbers (3/4)

The table below lists the nine classifiers, and for each one it specifies a context where we are likely to see the words that the classifier has labeled **YES** but not the words that it has labeled **NO**.

K3. Fill in the blanks **(a)** - **(h)** in the table below. *Notes:*

- In many cases, there are multiple possible answers. You only need to provide one.
- Each classifier (w_1 through w_9) will be used exactly once.
- When providing potential contexts (in the right column), you should focus on which types of words are most likely, rather than which ones are possible. For instance, for w_1 , the provided context of *I found these _____* is most likely to be followed by a plural noun (which is the correct category that w_1 identifies), but there are settings where this context could be followed by a singular noun (such as in *I found these apple pies*). However, it's more likely for a plural noun to go in that context, meaning that *I found these _____* would be a fully valid answer for w_1 even though technically it could allow singular nouns as well as plural nouns.
- In the contexts that you create (in the right column), you may not use any words from the list of 22 words on the last page of this problem (*alligators, apple, etc.*).

Classifier...	Returns YES for words that are likely in the context...
w_1	I found these _____
(a)	The person was chased by the _____
(b)	I ate the _____
(c)	Yesterday, the person _____
w_5	(d)
w_6	(e)
w_7	(f)
w_8	(g)
w_9	(h)

K4. Below are the vectors for a few more words. For each one, provide a single word that the vector could potentially stand for, assuming that these vectors can also be correctly classified by w_1 through w_9 . There might be multiple right answers; you only need to provide one.

(a)

5	-1	2	-3	-3	-7	4
---	----	---	----	----	----	---

(b)

-5	-2	-1	-2	-12	1	2
----	----	----	----	-----	---	---

(c)

9	-4	5	-9	6	3	7
---	----	---	----	---	---	---



(K) Going By the Numbers (4/4)

<i>alligators</i> =	-5	7	-9	-17	-10	1	3	<i>oranges</i> =	-3	-2	-9	-12	-3	2	8
<i>apple</i> =	5	-4	-3	-5	4	3	2	<i>parrots</i> =	-2	1	-5	-19	-17	2	7
<i>artists</i> =	-3	3	-2	-5	-24	1	9	<i>pizza</i> =	5	-3	-2	-1	1	4	3
<i>book</i> =	5	-1	-4	-14	-5	-4	4	<i>plates</i> =	-1	-3	3	-4	-9	2	3
<i>repaired</i> =	2	-1	-2	2	-14	2	5	<i>purse</i> =	6	-3	-3	-7	-7	-2	4
<i>electrician</i> =	9	2	-1	-2	-12	3	4	<i>salads</i> =	-1	-7	-5	-11	-1	1	5
<i>envelopes</i> =	-4	-2	-1	-9	-7	-2	3	<i>squirrel</i> =	7	2	-2	-5	-10	1	4
<i>interpreter</i> =	7	2	-5	-1	-9	2	2	<i>sticks</i> =	-3	-5	6	-9	-16	1	4
<i>lawyers</i> =	-2	2	-3	-3	-6	1	1	<i>restaurant</i> =	4	-5	-2	-2	-6	-1	3
<i>mug</i> =	3	-2	1	-2	-19	7	2	<i>television</i> =	4	-6	2	-11	-12	3	3
<i>noticed</i> =	1	-8	-4	19	-42	5	9	<i>visited</i> =	1	-2	-2	7	-24	1	9

$$w_1 = \begin{bmatrix} -3 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$w_2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$w_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$w_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 2 \end{bmatrix}$$

$$w_5 = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}$$

$$w_6 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -2 \end{bmatrix}$$

$$w_7 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$w_8 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

$$w_9 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$$



(L) Whistle While You Work (1/1) [15 Points]

Kickapoo is an Algonquian language spoken by approximately 1,000 people in Oklahoma, USA and Coahuila, Mexico. Kickapoo whistle speech is a substitute for spoken Kickapoo using whistles. It was once used for communication between teenagers in Coahuila. Below are some numbers and other words, given both in spoken and whistled Kickapoo. The English meaning is also given for interest, but is not relevant to the solution of the problem.

Notes: In spoken Kickapoo, the mark ´ over a vowel (e.g. **é**) indicates an accented syllable; **θ** = *th* as in *bath*; two adjacent vowels are part of the same syllable. In whistled Kickapoo, **o** represents a mora (a “beat” or unit of time) of a whistled note; the marks ´ and ` on a mora (e.g. **ó** and **ò**) indicate high and low pitch, respectively; **q** indicates a silence between notes; **h** indicates a transition between notes that can be an incomplete silence, a volume change, use of friction, or a pitch glide.

<u>Spoken</u>	<u>Whistled</u>	<u>English</u>	<u>Spoken</u>	<u>Whistled</u>	<u>English</u>
nékoti	óqòqòò	<i>one</i>	nenoh́táweewa	óhòqóhòòhóò	<i>he heard him</i>
níiswi	óòhóò	<i>two</i>	áikaaθoníani	óòqòòhóhóhóò	<i>he is pictured</i>
néθwihi	óhòhóò	<i>three</i>	nesíanaateθikwéeni	(e)	<i>he must be awful</i>
níewi	óòhóò	<i>four</i>	oθáaskwaanéhkweya	òhóòqóóhóqòhóò	<i>his shoes</i>
níananwi	óóhòhóò	<i>five</i>	nemiináaki	óhóóhóòqóò	<i>I gave it to them</i>
nekotwáasika	óqóqóóhòqóò	<i>six</i>	asikáθoona	òhóqóhòòhóò	<i>sock</i>
nóohika	óóhòqóò	<i>seven</i>	kemoopiimwaaki	(f)	<i>in your car</i>
neswáasika	òhóóhòqóò	<i>eight</i>	pakáanani	(g)	<i>pecans</i>
sáaka	óòqóò	<i>nine</i>	oθeθeehwáahi	(h)	<i>their older brothers</i>
metáaθwi	(a)	<i>ten</i>	nesáeθípena	òhóòhóqòhóò	<i>we are hungry</i>
niiswaaπίtaki	(b)	<i>twenty</i>	nemeskoθípena	(i)	<i>we are red</i>
máhwaha	(c)	<i>a bear</i>	nenakamópena	(j)	<i>we sang</i>
níihtiiθe	óóqòòhóò	<i>my friend</i>	nenemaθópena	óhóhòhóqòhóò	<i>we stood</i>
piitikeeno	(d)	<i>enter!</i>	kekoohkoosemwáaki	(k)	<i>your pigs</i>
			eehkatáinaahkwéeki	òòqóqóòhòòqóòqóò	<i>when it was almost noon</i>

L1. On your Answer Sheets, fill in the gaps **(a)** - **(k)** in the table above.

L2. Some linguists believe that the old word for seven in Kickapoo was ***niiswáasika** (the ***** indicates that this is a theorized word, rather than one directly observed). What would this have been in whistled speech?

L3. Explain how the whistled form of a Kickapoo word is formed. You do not need to justify your explanation (that is, you don't need to say why you think your explanation is correct).



(M) Relative Clauses (1/1) [10 Points]

Irish, also known as Gaeilge or Irish Gaelic, is spoken fluently by approximately 80,000 people in Ireland. It is a Celtic language, distantly related to English and many other languages of Europe. Here are five Irish sentences:

- A. Tá garmhac santach ag deartháir cliste.
- B. Tá iníon chineálta ag nia cúthail.
- C. Tá deirfiúr chliste ag garmhac cineálta.
- D. Tá deartháir cineálta ag deirfiúr bheag.
- E. Tá iníon shantach ag deirfiúr chúthail.

M1. Below are three of the translations. To what Irish sentences do they correspond? Write your answers on your Answer Sheets.

A short sister has a kind brother.

A smart brother has a greedy grandson.

A shy nephew has a kind daughter.

M2. Translate the remaining sentences, indicating the letter for each one.

M3. Translate the following sentences to Irish:

(a) A smart grandson has a shy daughter.

(b) A greedy sister has a short nephew.

M4. *Justify* your answers. That is, by referencing the evidence provided in the problem, explain why you think your answers are correct. You should include any observations you made about the Irish language.



(N) Something That Is Solved (1/1) [10 Points]

Nancowry (also known as Mūöt) is an Austroasiatic language spoken by around 950 people in the Nicobar Islands, India, specifically on the islands of Nancowry, Katchal and Kamorta. Below are some words in Nancowry with possible English translations, given in a phonetic transcription. Note that **u**, **ə**, **ɛ**, **ia**, **ua**, **ua**, **uã**, **ã**, and **ẽ** are vowels; **ŋ**, **y**, and **ʔ** are consonants.

caluak	<i>to swallow</i>	ta	<i>level, flat</i>
fanue	<i>string</i>	teh	<i>to float</i>
hakuãt	<i>to hang (something) up</i>	tikoʔa	<i>something that is prodded</i>
haneh	<i>an airplane</i>	tinkoʔ	<i>a prodding stick</i>
harac	<i>to heat (something)</i>	tiyäh	<i>new</i>
hata	<i>to level (something) out</i>	top	<i>to drink (something)</i>
hateha	<i>something that is floated</i>	tumyäh	<i>to renew (something)</i>
haʔëh	<i>to approach (something)</i>	ʔinttät	<i>a wipe</i>
hinkuãt	<i>a hook (e.g. on a wall)</i>	ʔittät	<i>to wipe (something)</i>
humʔëh	<i>to cause (something) to approach</i>	fue	(a)
kanap	<i>a tooth</i>	fuea	(b)
kapa	<i>something that is bitten</i>	hapin	(c)
kasu	<i>to trap (something)</i>	kuãt	(d)
kintal	<i>a saw</i>	haruk	(e)
kuah	<i>to shave</i>	(f)	<i>a beverage</i>
pahuaʔ	<i>to fear, be scared</i>	(g)	<i>a cup</i>
paloʔ	<i>loose</i>	(h)	<i>to float (something)</i>
pin	<i>thick</i>	(i)	<i>to fly</i>
pumhuaʔa	<i>a scared person</i>	(j)	<i>a knife</i>
pumloʔ	<i>to loosen (something)</i>	(k)	<i>to make (something) prod</i>
rac	<i>hot</i>	(l)	<i>near</i>
ranuk	<i>a vehicle</i>	(m)	<i>to saw (e.g. wood)</i>
ruk	<i>to arrive</i>	(n)	<i>something loose</i>
sahuan	<i>cool</i>	(o)	<i>a throat</i>
saput	<i>to turn (oneself) over</i>	(p)	<i>a trap</i>
sinhuan	<i>(some) ice</i>	(q)	<i>to turn (something) over</i>
sumhuan	<i>to cool (something) down</i>		

N1. Fill in the gaps **(a) - (q)**. In exactly one answer, you should be unable to deduce a particular letter. In this case, replace it with an **X**, e.g. “**anXwer**”.

N2. Below are four more Nancowry words, and their English translations in a random order. Determine the correct correspondences.

- | | | | |
|--|-------------------------|------------------------------------|---|
| 1. kahay | 2. kumfec | 3. mahelepa | 4. tanian |
| A. <i>to feel empty (in the heart)</i> | B. <i>a file (tool)</i> | C. <i>to make (something) tiny</i> | D. <i>something that is made to fit</i> |

N3. Explain your observations about how words are formed in Nancowry. You do not need to justify your explanation (that is, you don't need to say why you think your explanation is correct).

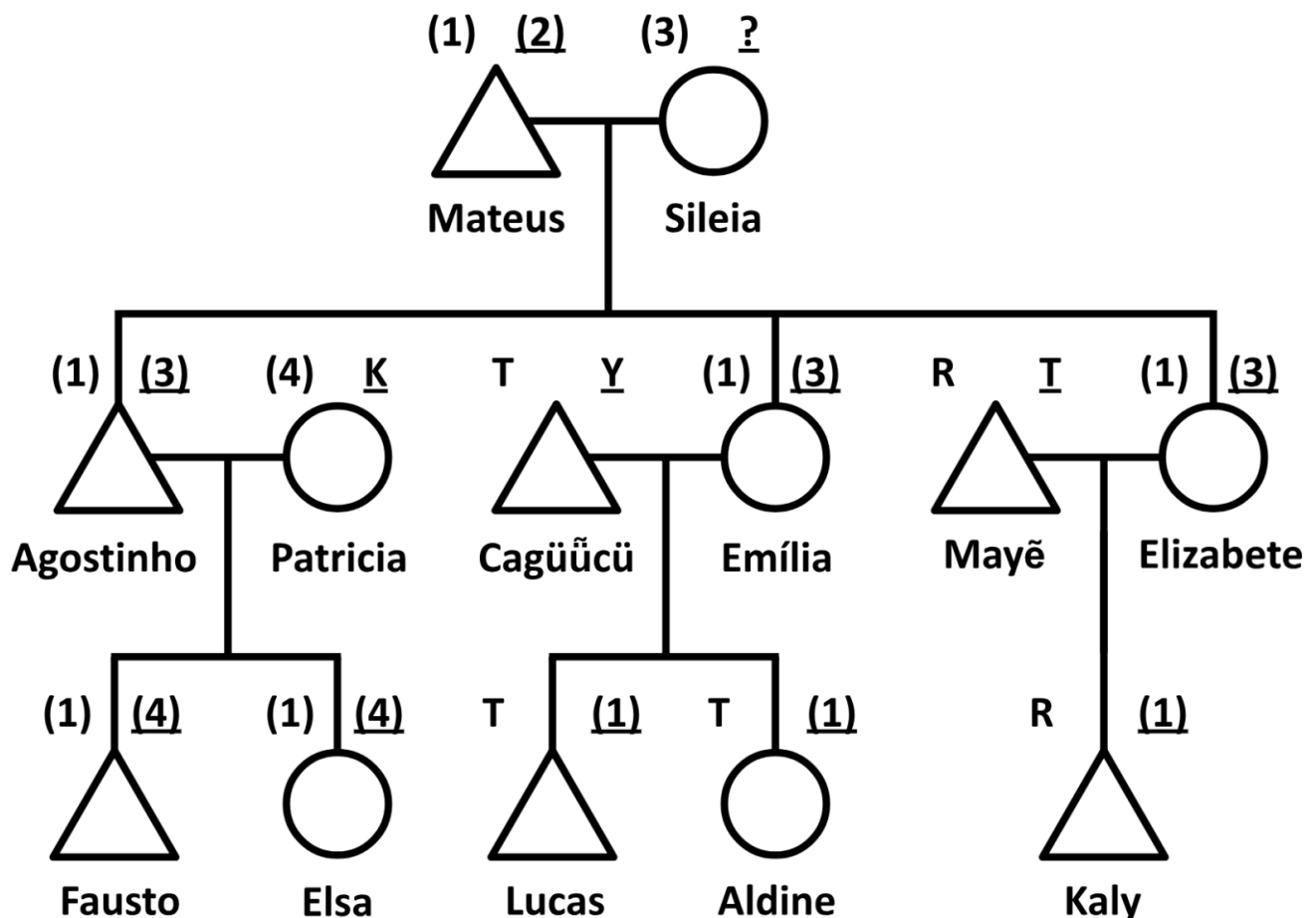


(O) I Speak, Therefore I Am (1/3) [10 Points]

Along the Vaupés River in Brazil and Colombia today are 17 indigenous groups whose languages belong to the Eastern Tukanoan language family, including Kotiria, Tukano, and Tuyuka. The exception is Tariana, which is part of the Arawakan language family. There are around 30,000 people belonging to these groups, but less than half still speak their ethnic language. In this problem, all speakers will only speak their ethnic language.

The Vaupés groups are patrilineal, so a child's native language is their father's native language, which thereby becomes known as their patrilect. Their mother's native language becomes their matrilect. Furthermore, the majority of the Vaupés groups practice a marriage system known as exogamy, whereby marriages occur between individuals who speak different patrilects.

You are given a family tree showing the patrilect and matrilect of some family members and their relationship to each other. Triangles represent males, and circles represent females. A horizontal line between two people represents marriage, and a vertical line represents offspring. For example, Mateus and Sileia's children are Agostinho, Emília, and Elizabete. For each family member, their patrilect is indicated on the left, while their matrilect is indicated on the right and underlined. Note that **K** is Kotiria, **R** is Tariana, **T** is Tukano, and **Y** is Tuyuka.



n → **a** → **c** → **l** → **o**

(O) I Speak, Therefore I Am (2/3)

In this problem, all speech is given in a speaker's patrillect. You are given five statements along with their English translations. Note that **ɨ** and **ʉ** are vowels halfway between pizza and food. The tilde **̃** indicates that air is escaping through the nose. The accents **´** and **ˆ** both mark stress. In the English translations, note that 'speak' refers solely to the patrillect, while 'understand' refers to the matrilect.¹

Speaker	Language	Statement	English Translation
Patricia	Tuyuka	Yʉ Docapúaraye wedea	<i>I speak Tuyuka</i>
Aldine	Tukano	Yí'î pako Akotí yee uumí	<i>My mom speaks Kotiria</i>
Agostinho	Kotiria	Yʉ Phaana thʉ'oeraha	<i>I do not understand Tariana</i>
Emilia	Kotiria	Yʉ phukoro Dahsea thʉ'oerara	<i>My dad does not understand Tukano</i>

O1. On your Answer Sheets, fill in the gaps (1) - (4) on the family tree with the correct languages using the above information. Each language is used exactly once.

O2. You are given seven more statements, one from each of Cagüücü, Mayë, Elizabete, Fausto, Elsa, Lucas, and Kaly in their patrillects. However, you are not told which speaker gives each statement, and their English translations are in scrambled order. Match each statement to its English translation.

Kotiria, Tariana, Tukano, or Tuyuka

English

- | | |
|--|----------------------------------|
| A. Yʉ phukoro Dahseakoro hira | i. <i>I speak Kotiria</i> |
| B. Nohá Taliá yarupe nu-tsaápeni-noka | ii. <i>I speak Tariana</i> |
| C. Yí'î pako Di'î yee uumí | iii. <i>I speak Tukano</i> |
| D. Yʉ Kotiriakoro hiha | iv. <i>I understand Tuyuka</i> |
| E. Yʉ Tuyuka thʉ'oha | v. <i>My mom speaks Kotiria</i> |
| F. Yí'î Ye'pâ-masa yee uúkũ | vi. <i>My mom speaks Tukano</i> |
| G. Nohádoa Panumápe yarupe du-tsaápeni-noka | vii. <i>My mom speaks Tuyuka</i> |

¹ In Kotiria, to "speak" is to "be"; "I am Kotiria" would be a closer translation than "I speak Kotiria". If your father is Kotiria you cannot say "I speak Tukano", but you could say "I understand Tukano".








(O) I Speak, Therefore I Am (3/3)

03. While matching the statements to their speakers, you find that you are stuck. Luckily, Fausto's twin kids Ricardo and María are reading over your shoulder. María reads aloud, **Yu Kotiriakoro hiha**. Her brother Ricardo repeats after her, but María laughs and tells him he can't say that. Ricardo agrees and corrects himself, **Yu Kotiriro hiha**. Suddenly, you have an epiphany and realize how to complete the task. Match the statements A - G to their speakers.

04. Ricardo needs your help saying, "My mom does not speak Tuyuka." On your Answer Sheets, write the corresponding sentence in Ricardo's patrilect.

05. There are several ways to convey the same information. For example, "I speak Kotiria" and "My dad speaks Kotiria" convey the same information because a speaker's patrilect must be the same as their father's patrilect. Give a sentence in Ricardo's patrilect that conveys the same information as "My mom does not speak Tuyuka."

06. Recall that it is generally forbidden to marry a partner who speaks the same patrilect. All of the marriages in this problem conform to this rule. However, there is another condition required for a marriage to be considered ideal. All of the marriages are also ideal, meaning that they meet this condition, except for the marriage between Mayë and Elizabete. Fausto, Elsa, Lucas, Aldine, and Kaly's spouses have challenged you by lining up in scrambled order as shown below. Each of them also tells you their patrilect and matrilect. Assuming that all marriages in this problem take place between two people of the opposite sex, there is exactly one way to correctly match each of the speakers to their spouse. On your Answer Sheets, identify the correct correspondences.

A. K <u>Y</u> 	B. Y <u>K</u> 	C. Y <u>K</u> 	D. Y <u>R</u> 	E. K <u>I</u> 
--	--	--	--	--

07. You are told that the marriage between Mateus and Sileia is ideal. On your Answer Sheets, identify Sileia's matrilineal language, indicated by a ? on the family chart.



(P) The Greatest Thing Since Boiled Breadfruit (1/1) [15 Points]

Kwomtari is a Senu River language spoken by several hundred people in Papua New Guinea. Below are some sentences in Kwomtari, alongside their English translations:

Meli blulu nelabe.	<i>The water will be boiling soon.</i>
Nalime nafomelee.	<i>He was telling himself a story a while ago.</i>
Nalile natemefe.	<i>I will be telling them a story in a while.</i>
Melife walafe.	<i>I will be going for water soon.</i>
Meneme naomeie.	<i>I was telling myself a story a while ago.</i>
Meli blulu nelalee.	<i>The water is boiling.</i>
Nibie blulu nefomefe.	<i>I will be boiling the breadfruit in a while.</i>
Menele naolane.	<i>They are telling me a story.</i>

Notes: Breadfruit is the large fruit of the breadfruit tree, related to jackfruit. Usually eaten cooked, it is a staple food in many tropical areas. In this problem, “breadfruit” is singular everywhere it occurs.

P1. Translate the following sentences from Kwomtari to English:

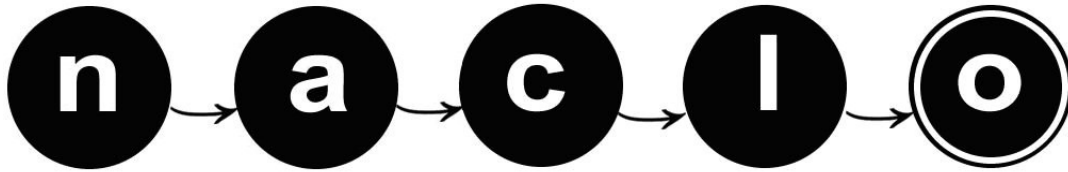
- (a) Menele naomelee.
- (b) Nalime natelane.
- (c) Wameie.

P2. Translate the following sentences from English to Kwomtari:

- (a) He will be going for breadfruit soon.
- (b) He was going to him a while ago.
- (c) The breadfruit will be boiling in a while.

P3. Explain your observations about the Kwomtari language. You do not need to justify your explanation (that is, you don't need to say why you think your explanation is correct).





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

Answer Sheets

REGISTRATION NUMBER					

Name: _____

Contest Site: _____

Site ID: _____

City, State/Province: _____

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 #

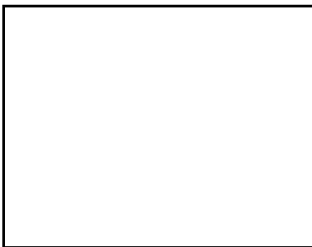
Answer Sheets (1/9)

(I) A Handheld Tablet

I1. Fill in the gaps **(1)** - **(18)** with a letter (A-R) from the glyph bank on page (1/2) of the problem. Fill in **(19)** with an English word. *Do not fill in the shaded cell.*

<u>Sumerian</u>	<u>Early Babylonian</u>	<u>Late Babylonian</u>	<u>Assyrian</u>	<u>English</u>
<i>(do not fill in)</i>	(1)	(2)	(3)	<i>man</i>
(4)	(5)	(6)	(7)	<i>king</i>
(8)	(9)	(10)	(11)	<i>house</i>
(12)	(13)	(14)		<i>prince</i>
(15)	(16) <i>(all three languages share the same glyph)</i>			<i>bird</i>
(17)	(18)			(19)

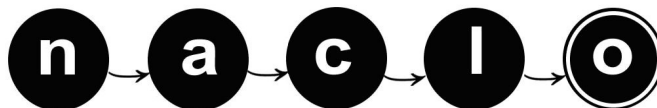
I2. In the box, draw the Sumerian glyph for *man*.



I3. Fill in the blanks:

A popular belief is that *Nineveh* can be literally translated to “(a) _____ of (b) _____.”

Problem (I) continues on the next page.



YOUR NAME:

REGISTRATION #

Answer Sheets (2/9)

(I) A Handheld Tablet (continued)

I4. Describe how the script changed over time from Sumerian to Early Babylonian, from Early Babylonian to Late Babylonian, and from Late Babylonian to Assyrian. You do not need to justify your explanation.

(J) Nganasan Idol

J1. Determine the correct correspondences by writing one letter (A-P) in each box.

1.	<input type="text"/>	2.	<input type="text"/>	3.	<input type="text"/>	4.	<input type="text"/>	5.	<input type="text"/>	6.	<input type="text"/>	7.	<input type="text"/>	8.	<input type="text"/>
9.	<input type="text"/>	10.	<input type="text"/>	11.	<input type="text"/>	12.	<input type="text"/>	13.	<input type="text"/>	14.	<input type="text"/>	15.	<input type="text"/>	16.	<input type="text"/>

J2. Translate into English:

17. *bəbəʔ*

18. *kojkə*

19. *kouʔ*

J3. Translate into Nganasan:

20. *boats*

21. *(piece of) metal*

J4. Suggest a Nganasan translation of the phrase *forest fire*:



YOUR NAME:

REGISTRATION #

Answer Sheets (3/9)

(K) Going By the Numbers

K1. What is the category of words that receive a label of YES from the classifier?

K2. What is $w_4 \cdot \text{squirrel}$?

K3. Fill in the blanks (a) - (h) in the table below.

Classifier...	Returns YES for words that are likely in the context...
w_1	I found these _____
(a)	The person was chased by the _____
(b)	I ate the _____
(c)	Yesterday, the person _____
w_5	(d)
w_6	(e)
w_7	(f)
w_8	(g)
w_9	(h)

K4. Write one word in each blank.

(a)

(b)

(c)



YOUR NAME:

REGISTRATION #

Answer Sheets (4/9)

(L) Whistle While You Work

L1. Fill in the words from positions (a) through (k) in the table.

(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"/>		

L2. What would *niiswáasika have been in whistled speech?

L3. Explain how the whistled form of a Kickapoo word is formed. You do not need to justify your explanation.



YOUR NAME:

REGISTRATION #

Answer Sheets (5/9)

(M) Relative Clauses

M1. Next to each English translation, write the letter of the Irish sentence to which it corresponds.

Letter	English translation
	A short sister has a kind brother.
	A smart brother has a greedy grandson.
	A shy nephew has a kind daughter.

M2. Translate the remaining sentences, indicating the letter for each one.

Letter	English translation

M3. Translate the following sentences to Irish:

(a) A smart grandson has a shy daughter.

(b) A greedy sister has a short nephew.

M4. Justify your answers and explain your observations about the Irish language.



YOUR NAME:

REGISTRATION #

Answer Sheets (6/9)

(N) Something That Is Solved

N1. Write the appropriate answer next to each letter (a) through (q).

Nancowry	English
fue	(a)
fuea	(b)
hapin	(c)
kuāt	(d)
haruk	(e)
(f)	<i>a beverage</i>
(g)	<i>a cup</i>
(h)	<i>to float (something)</i>
(i)	<i>to fly</i>
(j)	<i>a knife</i>
(k)	<i>to make (something) prod</i>
(l)	<i>near</i>
(m)	<i>to saw (e.g. wood)</i>
(n)	<i>something loose</i>
(o)	<i>a throat</i>
(p)	<i>a trap</i>
(q)	<i>to turn (something) over</i>

Problem (N) continues on the next page.



YOUR NAME:

REGISTRATION #

Answer Sheets (7/9)

(N) Something That Is Solved (continued)

N2. Next to each Nancowry word, write the letter (A - D) of the correct English translation.

1. kahay

2. kumfec

3. mahelepa

4. tanian

N3. Explain your observations about how words are formed in Nancowry. You do not need to justify your explanation.

(O) I Speak, Therefore I Am

O1. Next to each number, write the language that it corresponds to.

(1)

(2)

(3)

(4)

O2. Next to each letter, write the number of its English translation (from i to vii).

A.

B.

C.

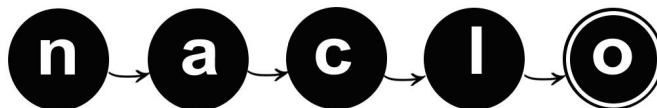
D.

E.

F.

G.

Problem (O) continues on the next page.



YOUR NAME:

REGISTRATION #

Answer Sheets (8/9)

(O) I Speak, Therefore I Am (continued)

03. Next to each name, write the letter of the sentence (A - G) which that person uttered.

Cagüücü	<input type="text"/>	Mayë	<input type="text"/>	Elizabeth	<input type="text"/>	Fausto	<input type="text"/>
	Elsa	<input type="text"/>	Lucas	<input type="text"/>	Kaly	<input type="text"/>	

04. Write "My mom does not speak Tuyuka" in Ricardo's patilect.

05. Give a sentence in Ricardo's patilect that conveys the same information as "My mom does not speak Tuyuka."

06. Next to each name, write the letter (A - E) of the corresponding person.

Fausto	<input type="text"/>	Elsa	<input type="text"/>	Lucas	<input type="text"/>	Aldine	<input type="text"/>	Kaly	<input type="text"/>
--------	----------------------	------	----------------------	-------	----------------------	--------	----------------------	------	----------------------

07. What is Sileaia's matrilineal language?

(P) The Greatest Thing Since Boiled Breadfruit

P1. Translate the following sentences from Kwomtari to English:

(a) Menele naomelee.

(b) Nalime natelane.

(c) Wameie.

Problem (P) continues on the next page.



YOUR NAME:

REGISTRATION #

Answer Sheets (9/9)

(P) The Greatest Thing Since Boiled Breadfruit (continued)

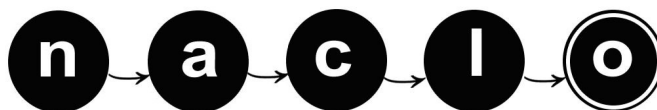
P2. Translate the following sentences from English to Kwomtari:

(a) He will be going for breadfruit soon.

(b) He was going to him a while ago.

(c) The breadfruit will be boiling in a while.

P3. Explain your observations about the Kwomtari language. You do not need to justify your explanation (that is, you don't need to say why you think your explanation is correct).



YOUR NAME:

REGISTRATION #

Additional Answer Space (1/1)

If you use this additional space, please do both of the following:

- 1. On this sheet, clearly indicate which question(s) you are answering. E.g., write "Problem P3."*
- 2. In the regular answer space(s) for the question(s) you are answering, note that you are using additional answer space. E.g., at the end of the answer space for problem P3 in the regular Answer Sheets, add a note saying "See additional answer space."*

