



Olympiade linguistique canadienne

PREMIÈRE ÉTAPE

25 MARS 2023

Bienvenu.e à la septième Olympiade linguistique canadienne francophone ! Nous te félicitons ! Tu participes à un évènement unique. Pour que le concours se déroule de façon juste pour toutes celles et tous ceux qui participent partout au Canada et ailleurs dans le monde, nous te demandons de lire attentivement et de suivre à la lettre les règlements suivants.

Règlements

1. Cette épreuve est d'une durée de trois heures et comprend cinq problèmes, ordonnés de A à E.
2. Certains de ces problèmes s'étendent sur deux pages. Assure-toi de les compléter jusqu'à la fin.
3. Cette épreuve est virtuelle. Tu devrais avoir reçu par courriel des instructions détaillées sur la démarche à suivre.
4. Si tu as besoin de précisions sur le contenu de n'importe quel problème, communique avec le jury suivant les instructions qui t'ont été transmises.
5. Aucune documentation ne peut être consultée, et personne ne peut être consulté, sauf dans le cas mentionné au point 4.
6. Il faut écrire toutes les réponses sur les feuilles de réponse, numériser toutes les feuilles de réponse, et les envoyer par courriel au jury à la fin de l'épreuve.
7. Fournis tes renseignements personnels sur la première feuille du cahier de réponses.
8. Chaque problème de cette épreuve a été soigneusement vérifié par des linguistes, ainsi que par des étudiant.e.s, pour assurer que les problèmes sont énoncés clairement, qu'ils ne contiennent pas d'erreurs et qu'ils peuvent être résolus. Certains d'entre eux sont plus difficiles que d'autres, mais ils peuvent tous être résolus avec un raisonnement logique et de la créativité. Aucune connaissance préalable n'est requise.
9. Si nous avons bien réussi notre tâche, très peu de candidat.e.s réussiront à résoudre tous les problèmes proposés dans le temps accordé. Ne te décourage pas si tu n'arrives pas à répondre à tout!
10. Tu peux garder ta copie du cahier de problèmes, mais nous te prions de ne pas la montrer ni d'en discuter avec qui que ce soit avant la fin avril.
11. Les résultats de la première étape seront annoncés par courriel en avril.

Amuse-toi!

Crédits et remerciements

Responsables :

Andrés Pablo Salanova (Université d'Ottawa), Gustavo Beritognolo (Université de Montréal)

Édition et correction des épreuves :

Andrés Pablo Salanova

Auteur des problèmes :

Andrés Pablo Salanova

(les sources des données utilisées dans les problèmes seront divulguées avec les réponses)

Réviseurs :

Gustavo Beritognolo

Marc Brunelle

Nathalie Dion

Clara Foz

Daniel Lovsted

André Nikouline

Surveillants de la première étape :

Gustavo Beritognolo

Corinne Soucy

Commanditaires en 2023 :

Département de linguistique, Université d'Ottawa
North American Association for Computational Linguistics (NAACL)

La composition graphique de ce document a été faite avec Lua^AT_EX.

(A) Wayúu wan?

Le wayúu est une langue arawak parlée par environ 400 000 personnes dans la péninsule de Guajira, dans la frontière entre la Colombie et le Vénézuéla.

En wayúu, l'accent tombe sur une des trois premières syllabes d'un mot, suivant certaines règles. Voici quelques exemples des mots avec l'accent indiqué par un accent aigu (´) sur la voyelle qui le reçoit. Puisque la source de ces données ne fournit pas de traduction pour certains de ces mots, nous ne présentons ici que les formes. Leurs traductions ne sont pas importantes pour résoudre ce problème.

tatúna	kamáneewaa	ayónnajaa
chochó	ashé'etaa	wayúu
chóochoo	wátta'a	éimalawaa
tóusajeein	áashajawaa	átpana
sha'watáa	a'ülújaa	nu'wayúuse
te'ejéna	me'rotóo	pü'yalája
jo'úukot	e'ínnaa	máa'ala

A1: Indique l'emplacement de l'accent dans les mots suivants :

- | | |
|----------------------|--------------------------|
| a. ta'waatüin | f. aja'itaa |
| b. susu | g. ka'wayuuseeshi |
| c. chü'ü | h. pütchimaajatü |
| d. iipüna | i. ka'apülasü |
| e. shaakuma | j. ee'irajaa |

Note : Le ' représente une consonne, le « coup de glotte », prononcé comme le silence entre les voyelles de *oh-oh*. Le **ü** représente une voyelle à mi chemin entre le *u* et le *ou* du français, mais prononcé avec les lèvres non-arrondies. Le **ch** et le **sh** représentent chacun une seule consonne : le *tch* et le *ch* du français, respectivement. Le **j** représente le son du *h* de l'anglais. Le **r** est un son entre *r* et *l*, propre au wayúu. Les autres lettres ont une prononciation proche de leur prononciation en français.

(B) Je me frotte la tête

Le guarani est une des langues officielles du Paraguay, et est parlé par environ 7 millions de personnes dans ce pays et dans les pays voisins. Comme tu le verras dans ce problème, en guarani un seul mot peut signifier une phrase complète.

omo'ã	Il/elle le couvre.
añehãikutu	Je me nettoie les dents avec un cure-dents.
rejepohéi	Tu te laves les mains.
añeakâpichy	Je me frotte la tête.
ajejuruhéi	Je me lave la bouche.
opirakutu	Il/elle pêche (perce des poissons).
ojepe'a	Il/elle (p.ex., une porte) s'ouvre.
reñeakâky'o	Tu t'enlèves les poux de la tête.
ahesamo'ã	Je couvre les yeux de quelqu'un.
rejepoka	Tu te tords.

B1: Que veulent dire les expressions suivantes ?

- ahesape'a**
- rejepichy**
- ojekutu**
- reñemo'ã**
- ajejurupe'a**
- ojejurumo'ã**

B2: Comment dirais-tu en guarani les phrases suivantes ?

- il se frotte les mains
- il se couvre la tête
- tu t'épouilles (enlèves les poux)
- je me perce les mains

Note : Le ' représente une consonne, le « coup de glotte », prononcé comme l'interruption dans la vibration des cordes vocales qu'on produit entre les voyelles de *oh-oh*. Le y représente une voyelle à mi chemin entre le *u* et le *ou* du français, mais prononcé avec les lèvres non-arrondies. Le ñ se prononce comme le *gn* dans *oignon*. Le j représente le son au début du mot *hier*, mais prononcé avec la langue plus rapprochée du palais. Les voyelles qui ont un tilde (~) sont des voyelles nasales, tandis que l'accent tonique est indiqué par un accent aigu (´), mais seulement quand il ne tombe pas sur la dernière syllabe.

(C) Dakelh te semble facile à lire ?

Le dakelh, aussi appelé porteur ou carrier, est une langue athabascane parlée par environ 1 000 personnes en Colombie Britannique. Le dakelh s'écrit aujourd'hui avec l'alphabet latin, mais le premier système d'écriture utilisé pour l'écrire était semblable à celui employé dans certaines régions pour écrire le cri et l'inuktitut. Cependant, en raison d'un plus grand nombre de sons, l'écriture du dakelh contient des symboles qui ne se trouvent pas dans ces autres langues.

C1: Voici quelques mots écrits dans le syllabaire. Dans la colonne de droite, tu trouves les transcriptions de ces mêmes mots en alphabet latin, mais en ordre aléatoire. Associe chaque mot à la bonne transcription.

a.	ᑕᑕ	homme	1.	ts'eke
b.	ᑕᑕ	original	2.	dzulh
c.	ᑕᑕ	couteau	3.	tilh
d.	ᑕᑕ	panier	4.	be'ooduget
e.	ᑕᑕᑕ	forêt	5.	yests'e
f.	ᑕᑕᑕ	arbre	6.	datsancho
g.	ᑕᑕ	femme	7.	duni
h.	ᑕ	canot	8.	tintah
i.	ᑕᑕᑕᑕᑕ	fourchette	9.	duchun
j.	ᑕᑕᑕ	chevreuil	10.	tes
k.	ᑕᑕ	plongeon	11.	dadzi
l.	ᑕᑕᑕ	caribou	12.	musdus
m.	ᑕᑕ	montagne	13.	dune
n.	ᑕᑕᑕᑕ	corbeau	14.	ts'i
o.	ᑕᑕᑕᑕ	vache	15.	whudzih

C2: Voici deux nouvelles publiées dans un journal dakelh paru en 1891 :

ᐱᐅᐅᐅ ᐅᐅ ᐅᐅᐅᐅᐅ, ᐅᐅ ᐅᐅᐅᐅ ᐅᐅᐅᐅ ᐅᐅᐅᐅ ᐅᐅᐅᐅ, ᐅᐅᐅ ᐅᐅᐅᐅ ᐅᐅᐅᐅ ᐅᐅᐅᐅ ᐅᐅᐅᐅ ᐅᐅᐅᐅ.

Pendant cinq jours, le Père Maurice a été incapable de prêcher la parole de Dieu car il avait mal à la dent.

ᐅᐅ ᐅᐅᐅ ᐅᐅᐅᐅᐅ ᐅᐅᐅᐅᐅ ᐅᐅᐅ.

Le Père Marchal est devenu chef de l'Okanagan.

ᐅᐅᐅ ᐅᐅᐅᐅᐅ ᐅᐅᐅᐅᐅ ᐅᐅᐅᐅ. ᐅᐅᐅᐅ ᐅᐅᐅᐅ.

La fille de Julia est née. Elle s'appelle Philomène.

Identifie les mots suivants dans ces extraits : (p) Maurice, (q) Marchal, (r) Okanagan, (s) Philomène, (t) Julia. Il faut considérer que la prononciation de ces noms en dakelh pourrait être différente de leur prononciation en français ou en anglais.

C3: Transcris les mots suivants avec l'alphabet latin :

- u. ᐅᐅᐅ grue du Canada
- v. ᐅᐅᐅᐅ terre, sol
- w. ᐅᐅᐅᐅ foyer
- x. ᐅᐅᐅᐅᐅᐅᐅ tisonnier
- y. ᐅᐅᐅᐅᐅᐅᐅᐅ cerisier sauvage

Note: Certaines combinaisons de lettres représentent un seul son en dakelh : **ts** (prononcé comme dans *tsé-tsé*), **ts'** (comme un **ts** mais avec glottalisation), **ch** (le *tch* du français), **dz** comme le *zz* dans *pizza*), **lh** (un *l* sans vibration des cordes vocales), **wh** (un *w* sans vibration des cordes vocales).

(D) Jaguar, crapaud, tamanoir, tinamou

Ce petit texte est un extrait légèrement simplifié d'un récit raconté à l'auteur du problème par un aîné du peuple mēbēngôkre, un groupe autochtone du centre du Brésil. Le tamanoir, aussi appelé *fourmilier géant*, est un gros mammifère qui se nourrit de termites et qui a des pattes antérieures très puissantes qui terminent par des ongles longs et courbés qui lui permettent de détruire les nids de termites. Un tinamou est un oiseau qui ressemble à une perdrix.

rop ne bri puma je prõt, nē pī pa o amino nhõpõk,

le jaguar court par peur du crapaud, et perce son propre œil avec une branche d'arbre

nē kam tu nō, nhỹm pàt ne 'yry bõx

et alors il [reste] juste couché, et un tamanoir arrive près de lui

nē kam — « tãmdjwýje'õ, mýj ne ano 'ã? »

et alors [le tamanoir dit] « petit-fils, [c'est] quoi sur ton œil ? »

— « ba bri puma je iprõt nhỹm pī pa ne ino nhõpõk »

« j'ai couru par peur du crapaud et une branche d'arbre m'a percé l'œil »

nhỹm kam — « amrẽ ba amã ano kanhĩn kaba »

et alors [le tamanoir dit] « [viens] vers ici, je t'enlève la crotte de ton œil »

nhỹm kam kum — « kati, dja ga ino kaba »

et alors [le jaguar] lui [dit] — « non, tu enlèveras mon œil »

nhỹm kum — « kati, dja ba ano kanhĩn bit kaba »

et [le tamanoir] lui [dit] — « non, j'enlèverai seulement la crotte de ton œil »

nē kam no mã pa owadjà nē « tongȳk », no kuni kaba

et alors il rentre sa patte dans l'œil et « plouc », enlève l'œil entier

nē mã tẽ nhỹm nō nhỹm kam atoroti bõx

et [le tamanoir] part et [le jaguar] reste couché, et alors un tinamou arrive

nē kum — « tãmdjwýje'õ, mýj ne ajã? »

et [le tinamou] lui [dit] « petit-fils, c'est quoi sur toi ? [qu'as-tu ?] »

nhỹm kam kum — « ba iprõt nē pī o amino nhõpõk,

et alors [le jaguar] lui [dit] « j'ai couru et j'ai percé mon propre œil avec un arbre,

nhỹm kam pàt ne ino kaba »

et alors le tamanoir m'a [complètement] enlevé l'œil. »

nhỹm kum — « go amrẽ dja » nhỹm atoroti 'yry dja

et [le tinamou] lui [dit] « mets-toi debout vers ici [approche-toi] », et [le jaguar] se met près de lui,

nhỹm o nhỹ nē no kane

et [le tinamou] s'assoit avec lui et lui guérit l'œil

Les traductions de chaque ligne du récit sont très littérales. Cependant, les mots en mēbēngôkre ne suivent pas nécessairement le même ordre qu'en français, et parfois les mots d'une langue ne sont pas présents dans l'autre. Quand quelque chose est entre crochets dans la traduction, il s'agit soit d'une clarification, soit de l'ajout d'un mot qui est sous-entendu mais qui n'est pas exprimé dans le texte mēbēngôkre, et qui est nécessaire pour bien le comprendre.

Piste : les noms des animaux dans ce récit sont les suivants : **rop, bri, pàt, atoroti**.

D1: Traduis les expressions suivantes vers le français :

- | | |
|----------------------------------|-----------------------------------|
| a. rop ne pĩ mã pa owadjà | f. ba pĩ pa o pĩ nhõpõk |
| b. ga pàt 'yry aprõt | g. ba tu nhỹ |
| c. ba rop pyma je mã tẽ | h. bri ne mã tẽ |
| d. ba pĩ pa kuni kaba | i. rop ne bõx |
| e. atoroti ne ipa kane | j. pàt no kanhĩn ne apa 'ã |

D2: Il y a deux mots différents dans le récit qui sont traduits par *et*. Lesquels, et quand est-ce qu'on utilise l'un ou l'autre ?

Note : Les voyelles « accentuées » représentent des voyelles particulières du mēbēngōkre : celles avec un tilde (~) sont des voyelles nasales, le **y** est une voyelle entre *ou* et *u*, mais avec les lèvres non-arrondies, le **ê** est comme le *é* du français, le **ÿ** est une voyelle un peu plus ouverte que **y**, et le **à** une voyelle plus ouverte que **ÿ** mais moins que **a**. Le **dj**, **nh** sont des consonnes qui se prononcent comme le *j* dans *jean* et comme le *gn* dans *oignon*, respectivement. Le ' représente le « coup de glotte », prononcé comme l'interruption dans la vibration des cordes vocales qu'on produit entre les voyelles de *oh-oh*. Le **r** représente un *r* comme celui de *cara* en espagnol ou comme les *tt* de *better* en anglais.

(E) Un peu de maths n'est guèze de trop

Le guèze (ou ge'ez) est un système d'écriture employé dans un ensemble de langues parlées dans la Corne de l'Afrique, dont les principales sont l'amharique et le tigrigna. En guèze, comme en latin et en grec classique, les lettres s'utilisent aussi pour écrire des chiffres.

Voici quelques comptes où les chiffres ont été remplacés par leurs équivalents guèzes (les symboles +, −, × et = ne sont pas utilisés en guèze).

- a. $\bar{\rho} \times \bar{\xi} = \bar{\xi} \bar{\rho}$
- b. $\bar{\xi} \times \bar{\iota} = \bar{\psi}$
- c. $\bar{\psi} + \bar{\psi} = \bar{\rho}$
- d. $\bar{\iota} \times \bar{\iota} = \bar{\rho}$
- e. $\bar{\rho} + \bar{\psi} = \bar{\rho} \bar{\psi}$
- f. $\bar{\iota} + \bar{\iota} = \bar{\pi}$
- g. $\bar{\rho} \bar{\xi} + \bar{\pi} = \bar{\rho} \bar{\pi} \bar{\xi}$

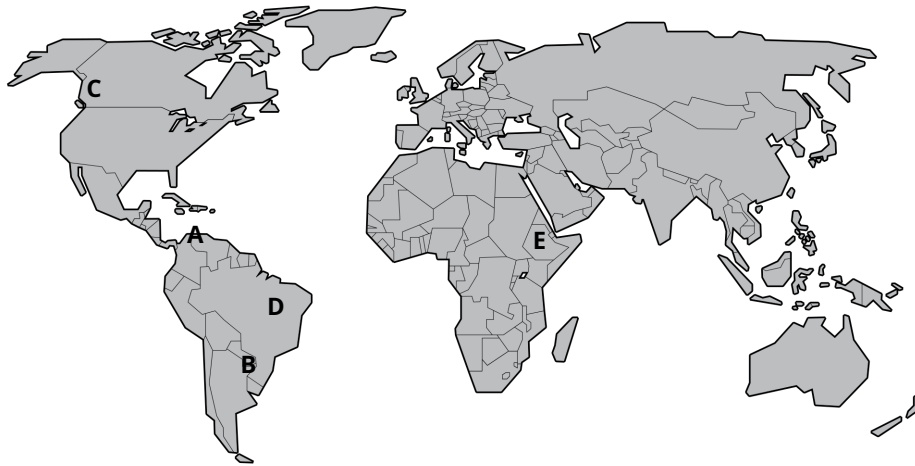
E1: La valeur la plus élevée dans ces expressions est 500. Donne la valeur qui correspond à chaque expression (seulement le total de chaque côté du symbole d'égalité).

E2: Donne les valeurs qui manquent (en chiffres) dans les expressions suivantes :

$$\begin{array}{l} \bar{\iota} - \bar{\xi} = \text{(h)} \\ \text{(j)} + \bar{\iota} = \bar{\iota} \bar{\xi} \end{array} \quad \begin{array}{l} \bar{\psi} \times \text{(i)} = \bar{\xi} \bar{\rho} \\ \bar{\xi} \times \text{(k)} = \bar{\rho} \bar{\pi} \bar{\xi} \end{array}$$

Note : Pour bien représenter en guèze les quantités qui sont représentées par plus d'un caractère, il faut relier les lignes horizontales au-dessus et en-dessous de chaque caractère qui les compose, à la manière de la numérotation romane. Cependant, en raison de limitations typographiques, ces lignes ne sont pas reliées dans la plupart des ouvrages imprimés.

C'est tout, merci!



Emplacement des langues évoquées dans cette épreuve

Chacune des quelque 7 000 langues parlées dans le monde de nos jours —ainsi que chacune des centaines de langues qui ne sont désormais plus parlées mais sur lesquelles nous disposons de quelques informations— est une vraie merveille : elle est à la fois un réservoir de connaissances du peuple qui la parle et un système de règles inconscientes extrêmement complexes qui s’appliquent de façon régulière et systématique pour permettre aux locuteurs d’exprimer des pensées sur pratiquement toute chose imaginable.

Les linguistes examinent les langues dans une optique scientifique, qui ne juge pas (il n’y a pas un *meilleur parler*, ni des langues plus avancées que d’autres), et qui traite le phénomène linguistique de façon objective, tel qu’il est observé « sur le terrain » ou en laboratoire. Parfois avec des informations incomplètes, le linguiste essaie de dégager les règles qui décrivent le phénomène et comprendre le fonctionnement du langage humain de façon plus approfondie. En appliquant des principes semblables à ceux dont tu t’es servi.e pour résoudre ces problèmes, les linguistes décodent des textes anciens, documentent et décrivent des langues jamais écrites auparavant, examinent la variation géographique et sociale des langues et construisent des modèles de fonctionnement du langage dans le cerveau humain. Les linguistes collaborent aussi à l’élaboration de matériaux d’apprentissage de langues et de logiciels pour la traduction automatique, ainsi qu’à l’amélioration des traitements pour les troubles de langage, entre autres.

Si ces problèmes t’ont intéressé(e), visite notre site web pour trouver plus de ressources : <https://linguisti.ca/OLCF>.

Tu peux aussi nous laisser des commentaires sur la feuille de réponse.