



Olympiade linguistique canadienne

DEUXIÈME ÉTAPE

24 AVRIL 2021

Bienvenue à la deuxième étape de la cinquième Olympiade linguistique canadienne francophone! Nous te félicitons pour tes résultats à la première étape!

Pour que le concours se déroule de façon juste pour toutes celles et tous ceux qui participent à l'épreuve, nous te demandons de lire attentivement et de suivre à la lettre les règlements suivants.

Règlements

1. Le concours dure trois heures et inclut cinq problèmes, ordonnés de F à J.
2. Cette épreuve est virtuelle. Tu as dû recevoir des instructions détaillées sur la procédure à suivre.
3. Si tu as besoin de clarifier le contenu de n'importe quel problème, communique avec le jury, qui sera en ligne pendant l'épreuve.
4. Tu ne peux avoir accès à aucune documentation ni discuter des problèmes avec personne, sauf dans le cas mentionné au point 3.
5. Toutes tes réponses doivent être écrites sur la feuille de réponse, qui doit être numérisée et envoyée par courriel au jury immédiatement après l'épreuve, à *ling.olymp.canada@gmail.com*.
6. Remplis tes renseignements personnels sur la feuille de réponse.
7. Chaque problème de cette épreuve a été vérifié en détail par des linguistes, ainsi que par des étudiants, pour s'assurer que les problèmes sont énoncés clairement, qu'ils ne contiennent pas d'erreurs et qu'ils peuvent être résolus. Certains problèmes sont plus difficiles que d'autres, mais ils peuvent tous être résolus par le raisonnement logique et en employant ta créativité. Aucune connaissance préalable n'est requise.
8. Si nous avons bien accompli notre tâche, très peu de candidat.e.s réussiront à résoudre tous les problèmes proposés dans le temps imparti. Ne te décourage pas si tu n'arrives pas à répondre à tout.
9. Les résultats de la deuxième étape seront annoncés par courriel vers la mi-mai.

Amuse-toi!

Crédits, remerciements et bande-annonce sur la linguistique

Organisateurs :

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

Édition et correction des épreuves :

Andrés Pablo Salanova

Auteurs des problèmes :

F : Andrés Pablo Salanova

G : Olga Ourioupina

H : Andrés Pablo Salanova

I : Olga Fiodorova

J : Andrés Pablo Salanova

Réviseurs (un très grand merci à vous!):

Daniel Lovsted, Clara Foz, Andrey Nikulin, Emad Alansary, Gustavo Beritognolo

Commanditaires en 2021 :

North American Association for Computational Linguistics (NAACL)

Chacune des quelque 7000 langues parlées dans le monde dans l'actualité —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 en même temps un réservoir de connaissances du peuple qui la parle et un système extrêmement complexe de règles inconscientes 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 « dans la nature ». Parfois avec des informations incomplètes, le linguiste essaie de mettre au jour les règles qui décrivent le phénomène et de comprendre le fonctionnement du langage humain de façon plus approfondie. En appliquant des principes semblables à ceux dont tu t'es servi 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.

Visite notre site web pour trouver plus de ressources : <http://aix1.uottawa.ca/~asalanov/OLCF>.

Tu peux aussi nous laisser des commentaires sur les feuilles de réponse.

(F) Un problème classique sur l'arabe classique

L'arabe classique est la forme de la langue arabe employée dans des textes du 7^{ème} au 9^{ème} siècle. C'est la variété d'arabe sur laquelle est basée la langue arabe standard contemporaine, qui peut être assez différente des variétés parlées localement. Nous fournissons une translittération phonétique de quelques mots de l'arabe classique, avec des traductions approximatives vers le français :

ta' liim	enseigner
'allama	il a enseigné quelque chose à quelqu'un
sallama	il a donné quelque chose à quelqu'un
ḏakkara	il a rappelé quelque chose à quelqu'un
tasarru'	se presser (se dépêcher)
tasallum	recevoir
tasrii'	presser (accélérer)
tadakkara	il s'est rappelé quelque chose
tasliif	prêter

F1. Sur la feuille de réponse, donne les formes ou traductions qui manquent dans le tableau suivant :

tasliim	(a)
(b)	il a prêté quelque chose à quelqu'un
(c)	il a appris quelque chose
(d)	il a reçu quelque chose
(e)	emprunter
ta'allum	(f)
(g)	il a pressé quelqu'un
(h)	rappeler
tasallafa	(i)
tadakkur	(j)
tasarra'a	(k)

Note : Les symboles ' et **ḏ** représentent des consonnes de l'arabe classique qui n'existent pas en français; la première se prononce comme un *r* articulé très au fond de la gorge, tandis que la deuxième se prononce comme le *th* de l'anglais dans le mot *then*.

(G) Sasamoan

Le samoan est la langue maternelle d'environ 500 000 personnes aux Samoa occidentales et Samoa américaines, dans l'océan Pacifique, ainsi que dans d'autres territoires polynésiens. Elle appartient à la famille malayo-polynésienne qui s'étend de Madagascar à l'ouest jusqu'à l'Île de Pâques à l'est, et de Hawai'i au nord à la Nouvelle-Zélande au sud.

Voici quelques phrases en samoan avec leurs traductions en français.

- | | |
|--|---------------------------------|
| 1. 'Ua lafi le pua'a. | Le cochon s'est caché. |
| 2. 'Ua tutuli e tagata maile. | Les gens ont chassé les chiens. |
| 3. 'Ua pupu'e e le pusi 'isumu. | Le chat a attrapé les souris. |
| 4. 'Ua pu'e e le tama le pusi. | Le garçon a attrapé le chat. |
| 5. 'Ua fefefe teine. | Les filles ont eu peur. |
| 6. 'Ua fasi e tama le 'isumu. | Les garçons ont tué la souris. |

G1. Traduis du samoan vers le français :

- a. **'Ua fefe le pusi.**
- b. **'Ua tuli e 'isumu le pusi.**

G2. Traduis du français vers le samoan :

- c. Les garçons se sont cachés.
- d. Les souris ont attrapé le chien.
- e. La fille a tué les cochons.

Note : Le ' est une consonne du samoan, le « coup de glotte », qui se prononce comme le silence entre les voyelles dans *oh-oh*.

(H) Fais le bon quechua

Voici la liste de tous les mots d'un dictionnaire du quechua de Cuzco qui contiennent les lettres suivantes : **ch', chh, k', kh, p', ph, q', qh, t', th** (même si on écrit deux ou trois lettres, chacune de ces séquences est considérée comme étant une seule lettre en quechua). La prononciation de ces lettres est expliquée à la page suivante.

aqha	ch'umpi	hump'i	k'ichiy	mach'aqway	phaqcha	qhata	q'utu	t'aqllay
chhachu	ch'unchull	hunt'a	k'ikuy	mach'ay	phasanqalla	qhatu	rakhu	t'aqsay
chhaku	ch'unqay	huq'ullu	k'illay	mak'as	phasi	qhaway	raphi	t'ika
chhalay	ch'uñu	husut'a	k'illichu	map'a	phasku	qhichwa	rap'i	t'impuy
chhalla	ch'upa	hutk'u	k'illimsa	maq't'a	phata	qhichincha	rawk'ana	t'inkay
chhallchay	ch'upu	ichhu	k'intu	markhu	phatay	qhichipra	rirp'u	t'inri
chhalluy	ch'uqñi	ichhuna	k'iri	marq'ay	phatmay	qhilli	rit'iy	t'ipana
chhapa	ch'uru	iphu	k'iski	mast'ay	phaway	qhipa	ruk'ana	t'ipqay
chhapchiy	ch'usaq	itha	k'ispiñu	mat'ay	phichu	qhispi	rump'u	t'iray
chhillchi	ch'usiq	khachuy	k'uchi	mat'i	phillillu	qhuna	ruphay	t'isay
chhulli	ch'uspa	khaka	k'uchu	mawk'a	phina	qhuña	ruqt'u	t'iyu
chhullunku	ch'uspi	khakuy	k'ullu	maych'a	phiña	qhupu	sach'a	t'uqpina
chhuqa	ch'uwi	khallay	k'umpa	mayt'u	phiri	qhurquy	sank'ayu	t'uqu
chhuruna	ch'uya	khallwa	k'umuchiy	mich'a	phiwi	qhuru	saphi	t'uqyay
ch'aki	ch'uyanchay	khamuy	k'upa	mikhuy	phukuna	qhutu	sayk'uy	t'utura
ch'allay	hak'akllu	khanka	k'urpa	millp'uti	phukuy	qhuya	simp'ay	t'uru
ch'aqchuy	hak'u	khasay	k'usillu	millp'uy	phullu	q'achu	siq'iy	t'uuyur
ch'aqi	hallp'a	khatatatay	k'usi	minchha	phuqchiy	q'allay	sirk'a	ukhu
ch'aqiy	hamak'u	khaya	k'utuy	mink'a	phuru	q'apay	such'i	uphakuy
ch'aqu	hamk'a	kipu	k'uychi	mink'akuy	phuska	q'apiy	such'u	wach'i
ch'aqway	hamk'ay	kipuna	k'uyka	misk'ay	phuspu	q'asa	sullk'a	wallwak'u
ch'arki	hamp'atu	khiwiy	k'uyuy	misk'i	phusullu	q'asuy	sunkha	wamp'u
ch'aska	hamut'ay	khuchi	laqha	mit'a	phuturiy	q'awa	suphu	wank'ar
ch'atay	hank'a	khullu	laq'akuy	mitk'ay	phuyu	q'ayma	suphullu	wanthi
ch'awar	hank'u	khuyuy	laq'ay	muchhay	p'acha	q'ichwa	surq'an	waskha
ch'away	hanuk'ay	k'acha	lat'ay	much'ay	p'akisqa	q'illay	suskhuy	wayk'uy
ch'ichi	hapt'ay	k'aki	lirp'u	much'u	p'ampay	q'illu	sut'i	wayq'u
ch'ichiy	hap'iy	k'alla	lirq'u	muk'uy	p'aqu	q'inti	sut'inchay	wayt'ay
ch'ikikuy	hark'ay	k'allampa	llamkhay	mullpha	p'asña	q'ipa	thalay	wichq'ay
ch'illiku	hasp'iy	k'allana	llamk'ay	muskhiy	p'isaq	q'ipi	thallay	wikch'uy
ch'in	hawaq'ullay	k'allma	llamp'u	mutk'a	p'isqi	q'ipichay	thaskiy	wik'uña
ch'inyachiy	hawch'a	k'amiy	llamt'a	mut'i	p'itay	q'ipiy	thasnuy	wiphala
ch'iqchi	hayk'a	k'analla	llank'i	mut'uy	p'uqchu	q'isa	thatkiy	wirp'a
ch'iqtay	hayk'ap	k'anchay	llanthu	nak'ay	p'urku	q'isana	thiqni	wisch'uy
ch'iquy	hayt'ana	k'anca	llik'iy	naq'iy	p'uku	q'isqintu	thiqtina	wisk'acha
ch'irway	hayt'ay	k'antiy	llimp'i	ñak'ariy	p'unchaw	q'isqis	thiqtiy	wisq'an
ch'isi	hich'ay	k'aray	llipt'a	ñaqch'a	p'uñu	q'iya	thupay	wisq'ay
ch'isiyay	hik'iy	k'arpa	lluk'iy	ñaqch'ay	p'uqay	q'umalliy	thuqay	wist'u
ch'iya	hillp'uy	k'arku	llunk'uy	ñaqha	p'usqu	q'umir	thuta	yamt'a
ch'uklla	hiq'ipay	k'askay	lluq'i	ñusqhun	p'uyñu	q'uncha	t'akay	yuthu
ch'ulla	hisp'ay	k'aspi	lluchk'a	phaka	qhachun	q'uñi	t'anta	
ch'ullpa	huch'uy	k'ata	lluskhuy	phalika	qhapaq	q'upa	t'apsay	
ch'ullu	huk'ucha	k'ayra	lluthu	phanchiy	qhaquy	q'uruta	t'aqa	
ch'umay	humint'a	k'ayrapin	llut'ay	phapu	qhari	q'usñi	t'aqllakuy	

En quechua, le **q** se prononce comme un *k* mais plus vers l'arrière de la bouche, le **ch** se prononce comme

en anglais (« chip »), le **ll** est un son proche du *li* dans « lion », et le **ñ** se prononce comme le *gn* de « compagnon ». Les lettres terminées par un **'** représentent des sons *glottalisés*, c'est-à-dire prononcés avec la fermeture complète des cordes vocales, et une légère explosion lors de l'ouverture de la bouche produite par l'air poussé par le larynx. Les lettres terminées par un **h** (sauf pour **ch**) représentent des sons *aspirés*, c'est-à-dire prononcés avec une sortie d'air qui dure assez longtemps après l'ouverture de la bouche.

Voici quelques exemples de mots qui ne contiennent aucune des consonnes en question :

papa utku tayta qillqay pinchi aka hatun wañuy harawiy chukulli chukcha aqsu

Note que le **h** ne peut apparaître qu'au début d'un mot.

Les sons aspirés et glottalisés en quechua de Cuzco ont un usage plus restreint que les sons représentés par les lettres **ch**, **k**, **p**, **q** et **t**, qui représentent des sons sans aspiration ni glottalisation (les lettres **k**, **p** et **t** sont prononcées comme les lettres *k*, *p* et *t* en français).

H1. En te basant sur les patrons que tu peux identifier dans la liste, lesquels des mots inventés suivants sont des mots possibles en quechua ? Pour chaque mot, indique *oui* ou *non* sur la feuille de réponse. Piste : environ la moitié de ces mots sont des mots possibles.

- | | | | |
|----|-----------------|----|-------------------|
| a. | suq'tiña | i. | paqhari |
| b. | ayllulla | j. | q'umirin |
| c. | p'ut'uni | k. | phuych'u |
| d. | wit'apha | l. | rumaq'u |
| e. | uq'ira | m. | chhupincha |
| f. | illwanaq | n. | haqarin |
| g. | hachhuni | o. | apharqa |
| h. | hut'irpa | p. | ithnuni |

Pour t'aider à réfléchir à ça, demande-toi si les mots suivants sont possibles en français : *moucd*, *dsrou*, *pmar*. Compare-les aux mots suivants : *losse*, *dourgue*, *trave*. Aucun de ces mots n'existe dans le dictionnaire, mais tu t'attendrais sans doute à trouver les trois derniers plus facilement que les trois premiers. Qu'est-ce qui fait que les premiers ne semblent pas être du vrai français, même si toutes les lettres sont des lettres employées en français ?

H2. Énonce les règles qui décrivent où peuvent apparaître les sons glottalisés et aspirés dans un mot en quechua. Piste : tes règles ne feront pas de différence entre les sons à l'intérieur d'une même classe (c'est-à-dire, tous les sons aspirés auront les mêmes possibilités de placement, et pareillement pour les sons glottalisés). Sois concis(e).

(I) Œil, nez, bouche...

Voici des termes pour des parties du visage en trois langues : laotien, tiriýó, et kuuk-thaayore. Les parties du visage sont les suivantes : 1. œil, 2. cils, 3. nez, 4. narine, 5. bouche, 6. lèvre inférieure, 7. lèvre supérieure. Il faut savoir que les mots **khon**, **ihpoti** et **panjr** ont la même traduction vers le français, et pareil pour les mots **huu**, **oota** et **ranth**.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
laotien	dang	paak	khon taa	him sop	taa	huu dang	him sop
tiriýó	oona	mita	enu ihpoti	poti	enu	oona oota	ehpi
kuuk-thaayore	koow	thaaw	meer-panjr	koo-petan	meer	koo-ranth	thaa-petan

I1. Place les numéros des parties du visage dans la case correspondante. Explique tes choix en quelques mots.

I2. Le mot **nam** signifie *eau* en laotien. Comment dit-on *larme* en laotien ?

I3. Le mot **apuru** signifie *couvercle* en tiriýó. Comment dit-on *paupière* en tiriýó ?

I4. Quelle partie du visage dans la langue kuuk-thaayore s'appelle **koo-mut-panjr** ?

Note : Le laotien est une langue de la famille tai, parlée par environ 30 millions de locuteurs en Asie du Sud-Est, en particulier au Laos, où elle est la langue officielle.

Le tiriýó est une langue de la famille caribe, parlée par 2000 personnes dans la région frontalière entre le Surinam et le Brésil.

Le kuuk-thaayore est une langue australienne de la famille pama-nyungan parlée par environ 200 personnes. Les mots en kuuk-thaayore et en tiriýó sont présentés dans l'orthographe habituelle. Les mots en laotien sont présentés dans une translittération en écriture latine. Le **ĩ** est une voyelle propre au tiriýó, intermédiaire entre le *i* et le *ou* du français.

(J) Relativement parlant

Le bouli est une langue gour parlée par environ 150 000 personnes au Ghana. Voici quelques phrases en bouli, fournies dans une transcription simplifiée. Dans cette transcription, le **j** représente le *gn* d'*oignon*.

N ja Atim li te nuru way mango ku la.	J'ai vu l'homme à qui Atim a donné la mangue.
Amok da n li te Atim na buy la.	Amok a acheté la vache que j'ai donnée à Atim.
N de nuru wa li ko na buy la.	J'ai mangé la vache que l'homme a tuée.
Nuru wa ja na bu.	L'homme a vu la vache.

J1. Traduis les phrases suivantes vers le français :

- Atim ja n li da mango kuy la.**
- Na bu de Amok li da mango kuy la.**
- N li da na buy la de mango ku.**

J2. Comment dirait-on en bouli ?

- J'ai donné la mangue à Amok.
- La vache que j'ai achetée a vu Atim.
- L'homme à qui Atim a donné la vache a vu Amok.

O le i'uga, fa'afetai lava (C'est la fin, merci!)



Carte de localisation des langues évoquées dans cette épreuve