



Olimpiada Lingüística d'Espanya
Olimpiada Lingüística de España
Espainiako Olinpiada Linguistikoa
Olimpiada Lingüística de España

Exercicis entrenament OLE 2014

01

Novembre 2013



Informació

Este dossier d'exercicis arreplega una selecció d'exercicis d'Olimpiades lingüístiques realitzades en cursos anteriors. La finalitat d'aquests problemes és que us exerciteu en la realització dels mateixos, aplicant la lògica i l'enginy.

És molt important llegir detingudament els enunciats dels problemes, ja que en ells trobareu la clau per a resoldre els problemes. Tingueu en compte que la informació que us oferim en els enunciats no està “perquè sí”, sinó que pot servir-vos de guia per a una correcta resolució de l'exercici.

En la prova de l'Olimpiada, tindreu tres hores per a resoldre sis problemes. És important que organitzeu be el vostre temps per a obtenir la màxima puntuació.

Com podreu comprovar, els exercicis varien en dificultat. Però, fins i tot en un mateix exercici, la dificultat també varia. No us desanimeu per açò. És probable que al principi us coste respondre a totes les parts, però ja veureu com, amb una mica d'entrenament, els vostres resultats milloraran molt. A més, tampoc esperem que tots contesteu tot el dossier. El normal és que contesteu una part, i a poc a poc sigueu capaços d'arribar més lluny. Però tot el que pugueu contestar serà avaluat. Per poc que sigueu capaços de resoldre d'un problema, eixa resposta puntuarà de cara a la final.

Recordeu que no fa falta que tingueu nocions de lingüística o d'altres llengües, tan sol que apliqueu la lògica i el sentit comú, i prengueu-vos-ho amb ganes.

Ànim!

Equip OLE



Índex

PROBLEMA	AUTOR	PÀGINA
(A) El cas del pali.....	Babette Verhooven.....	3
(B) La gramàtica mana!.....	Andrea Schalley i Patrick Littell.....	6
(C) La cara o creu de Huffman	John DeNero.....	9
(D) Inuit intuïtiu.....	Harold Somers.....	11
(E) Sopa de lletres en kwak'wala.....	Patrick Littell.....	14
(F) Fans del fenici!.....	Harold Somers.....	17

Exercicis adaptats de la *North American Computational Linguistics Olympiad*, traduïts per l'Olimpiada Lingüística d'Espanya



(A) El cas del pali (1/3)

El pali és una llengua morta, com el llatí. Era una llengua literària relacionada amb el sànscrit, el predecessor dels idiomes moderns parlats en el nord de l'Índia, com l'hindi. El pali es va començar a escriure al voltant de l'any 100 a. C. a Sri Lanka per monjos budistes per a preservar els ensenyaments de Buda. El pali se solia escriure utilitzant l'escriptura brami, però també s'escriu usant l'alfabet llatí (que és el que nosaltres anem a usar ací). El pali encara és usat per monjos budistes i estudiosos (de la mateixa manera que el llatí encara és usat al Vaticà per religiosos catòlics i teòlegs).

El pali és una llengua molt flexiva; significa que a les paraules principals com a noms i verbs se'ls afig una gran varietat de sufixos i prefixos que aclareixen què paper exerceix la paraula en l'oració.

L'espanyol és també una llengua flexiva, però no tant com el pali. Ací teniu uns exemples de flexió en català:

- xiqueta – xiquet; el final en –a ens indica el gènere de la paraula.
- cotxe – cotxes; afegint una –s s'indica el nombre plural de la paraula.
- salt – saltarem – saltaven – ha saltat; els sufixos afegit al verb salt(ar) informen de la persona que salta, si és singular o plural, o el moment en què se salta: passat, present o futur.

A més, segurament coneixereu altres idiomes en els quals s'usen prefixos i sufixos per a aportar altres informacions respecte a una determinada paraula; per exemple, en anglés:

- John – John's coat. La 's s'utilitza després d'un substantiu per a indicar possessió.

El pali té diferents sons consonàntics i vocàlics, que expliquen l'ús de signes diacrítics (signes especials) en lletres específiques. Estos signes diacrítics no són necessaris per a resoldre l'exercici. A més, el pali no empra ni majúscules ni signes de puntuació.

Ací teniu algunes oracions en pali amb la seua traducció en català:

Pali	Traducció al català
mahāmatto nisīdati	El ministre se senta
mahāmattaṃ upasaṃkamanti	Els visiten el ministre
samaṇo tathāgat hoti	El filòsof està il·luminat
samaṇi atthaṃ pucchanti	Els pregunten el significat als filòsofs
upāsako pucchati	El deixeble pregunta
loko mahāmattassa	El món del ministre



(A) El cas del pali (2/3)

A1. Traduiu les següents oracions en català a pali:

1. El ministre pregunta als filòsofs																			
2. El filòsof se senta																			
3. Ells se senten																			

A2. Traduiu les següents frases al català, usant el vocabulari que es mostra en el format d'un diccionari (tenen tots la forma del subjecte, sense sufixos):

Pali	Traducció a l'espanyol
rājo	rei
devo	déu
gāmo	poble

1. rājo nisīdati																			
2. rājo gāmassa devo hoti																			



(A) El cas del pali (3/3)

A3. Tradueix al pali, introduint una lletra en cada cel·la, ignorant els signes diacrítics:

1. El ministre pregunta als reis														
2. El poble del deixeble														
3. El significat del món és déu														



(B) La gramàtica mana! (1/3)

Una de les maneres en què els ordinadors entenen el llenguatge és mitjançant l'anàlisi d'oracions per a descobrir la funció que exerceix cada paraula. Una gramàtica lliure de contextos (GLC) (també cridada gramàtica de l'estructura de la frase) consisteix en una sèrie de regles per a formar oracions. Només les oracions que poden ser generades usant una sèrie de regles són considerades gramaticalment correctes i 'ben construïdes'. La informàtica i la lingüística usen les GLC per a definir i analitzar llengües, on una 'llengua' és definida com totes i cadascuna de les oracions que la GLC pot generar. O és el símbol amb el qual comença cada oració.

Les següents regles confeccionen una GLC molt senzilla:

- O → N V
- N → xiquets
- N → esquirols
- V → canten
- V → mengen

Cada regla indica que l'element a l'esquerra de la fletxa es pot expandir en els elements de la dreta de la fletxa. Mitjançant una substitució repetida de símbols, esta GLC pot expandir el símbol O en 'esquirols canten', 'xiquets canten', 'esquirols mengen' i 'xiquets mengen'. No poden, però, generar 'xiquets mengen esquirols' o 'esquirols mengen xiquets', ni tan sols 'xiquets' o 'mengen' –pots comprovar com no hi ha cap seqüència de substitucions que expandisca O en una d'elles–.

A continuació us oferim una altra GLC simple. Les regles s'han numerat per a facilitar-te el treball, però els números no són part de les regles.

- | | | | |
|--------------------|-------------------|---------------------|---------------|
| 1. O → SN SV | 2. SV → SV SP | 3. SP → P | 4. VI → corre |
| 5. SN → N | 6. SV → SV COP SV | 7. SP → P SN | 8. CONJ → que |
| 9. SN → D N | 10. N → esquirol | 11. VT → persegueix | 12. P → en |
| 13. SN → SN COP SN | 14. N → ell | 15. VT → (es) menja | 16. P → fora |
| 17. SV → VI | 18. N → Joan | 19. VT → atrapa | 20. COP → i |
| 21. SV → VI SP | 22. N → Maria | 23. VT → compta | 24. D → el |
| 25. SV → VT SN | 26. N → gos | 27. VT → veu | 28. D → la |
| 29. SV → VT CONJ O | 30. N → arbre | 31. VI → (se) senta | |



(B) La gramàtica mana! (2/3)

BI. Ací teniu una història simple. Moltes de les oracions, d'acord amb la GLC anterior, no estan ben construïdes, és a dir, no poden ser derivades des d'una O a partir d'una substitució repetida de símbols.

Indiqueu les oracions que la GLC anterior pot generar en les cel·les següents; ignora els punts i les contraccions.

- A. Joan veu el gos i Maria veu el gos.
- B. El gos veu Joan i Maria.
- C. El gos veu un esquirol.
- D. L'esquirol se senta en l'arbre.
- E. Aqueix esquirol veu el gos.
- F. L'esquirol va ser vista pel gos.
- G. El gos corre.
- H. L'esquirol a l'arbre corre.
- I. El gos atrapa l'esquirol i es menja l'esquirol.
- J. El gos menja.
- K. Joan veu que el gos es menja l'esquirol.
- L. Joan explica Maria que el gos es menja l'esquirol.
- M. El gos veu que Joan veu que ell es menja l'esquirol.
- N. I el gos corre fora.
- O. Maria i Joan persegueix el gos.
- P. Joan alcança i atrapa el gos.
- Q. Joan menja gos.



(B) La gramàtica mana! (3/3)

B2. No totes les oracions que la GLC pot generar són, en realitat, oracions en català. Per exemple, “Joan i Maria menja” pot ser generada, però no és una oració correcta en valencià.

Indiqueu altres tres exemples d'oracions que poden ser generades per aquesta GLC, però que no són correctes en valencià; introduïu una paraula per cel·la. Ignora els punts i el final de les oracions.

A.						
B.						
C.						

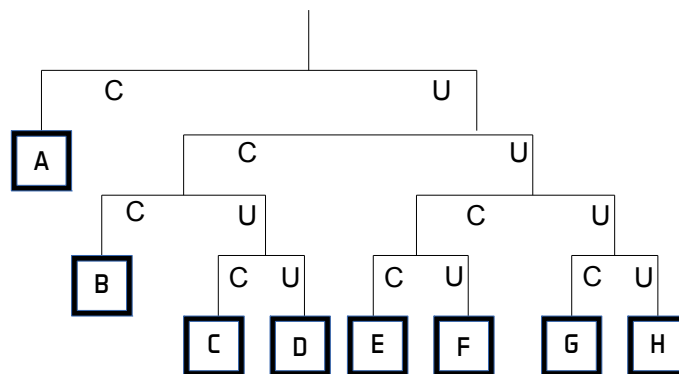
B3. Una de les regles de la GLC anterior és redundant: totes les oracions que poden ser generades usant aqueixa regla poden ser al mateix temps generades per la combinació d'altres regles. Escriviu el número de la regla redundant en la cel·la.



(C) La cara o creu de Huffman (1/2)

Quan Ada es tornà boja, va enviar a la seua amiga Gadea missatges en codi utilitzant monedes d'un cèntim, mostrant en cadascuna d'elles o bé la cara (C) o la creu (U) de la moneda. Per exemple: CUCUCCUCCCUUUCUCUUUCCCC.

Ada també va enviar un arbre de xifrat, que indicava com llegir el missatge codificat per les monedes. Un arbre de codis s'inicia amb dues branques, marcades amb (C)ara o cre(U). Cada branca dirigeix bé a una lletra del missatge, bé a una altra branca. Este tipus d'arbre és conegut com a Arbre de Huffman, el seu inventor.



Les monedes són llegides d'esquerra a dreta, i cadascuna indica cap a quina branca de l'arbre de xifrat s'ha de seguir. Quan s'aconsegueix una lletra, la següent lletra es codifica tornant a començar des de l'alt de l'arbre. Per exemple, el missatge anterior "CUCUCCUCCCUUUCUCUUUCCCC", és representat separant individualment les lletres en la taula següent:

A	C	A	B	A	G	A	D	E	A
C	UCUC	C	UCC	C	UUUC	C	UCUU	UUCC	C



(C) La cara o creu de Huffman (2/2)

C1. Descodifiqueu els següents missatges utilitzant l'arbre d'a dalt:

A.	UUCUUUCCUCUUUUCCUUCUCUCUUC																				
B.	UCUCCUUCUUUCCUCUUUUCCUUUCCUCUUUUCCC																				

C2. A la següent paraula anglesa de Ada li falta una moneda en el seu interior. Localitzeu el lloc en que manca, indiqueu l'orientació (cara o creu) del cèntim perdut i descodifiqueu el missatge:

CUCUCCUCCCCUUC																				
----------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Lloc de la nova moneda (contant des de l'esquerra):

Orientació:

C3. Ada no vol gastar-se tot els seus diners en missatges. Dissenyau un arbre de xifrat i escriviu el codi per a les lletres de baix, de tal manera que requereisca del mínim nombre de cèntims possible, però que els missatges es continuen descodificant de manera correcta. Considereu que el missatge només té les lletres de l'exemple (IMPS en el primer exemple i ABCDR en el segon). En una codificació Huffman, el codi d'una lletra no pot començar amb el codi d'una altra. Per exemple, si una lletra es codifica amb C, cap altra es pot codificar amb CU. De fet, si alguna lletra es codifica amb C, la resta de lletres hauran de començar per U.

MISSISSIPPI

LLETRA	CODI
I	
M	
P	
S	

Nombre total de monedes _____

ABRACADABRA

LLETRA	CODI
A	
B	
C	
D	
R	

Nombre total de monedes _____



(D) Inuit intuitiu (3/3)

D2. Donada la informació que heu extret del text, com escriuríeu les següents paraules utilitzant el sistema d'escriptura de l'inuktitut? Comenceu amb dues paraules per a la paraula 'neu'² Introduïu un caràcter en cada cel·la.

A.	qanniq 'neu quan cau'														
B.	aput 'neu quan està en el sòl'														
C.	mukluk 'botes de pell de foca'														
D.	umiaq 'canoa'														

D3. Finalment, podeu identificar la paraula que el català ha pres prestada de l'inuktitut en A, i identificar els llocs B i C? Escriviu la paraula en català, una lletra en cada cel·la.

A.	ᑕᑖᑭᑕᑖ (un mitjà de transport)														
B.	baC														
C.	ᑕᑕᑖ														

² Alguna vegada haureu llegit que els inuit tenen moltes paraules diferents per a referir-se a la 'neu'. En realitat, és una espècie de llegenda urbana. En inuktitut hi ha dues paraules per a referir-se a la neu, encara que molts dels seus significats poden expressar-se afegint sufixos –t'hauràs adonat que les paraules en inuktitut són molt llargues–.



(E) Sopa de lletres en Kwak'wala (1/3)

El Kwak'wala és la llengua dels Kwakwaka'wakw, una de les nombroses nacions indígenes canadenques. Es parla en la zona de Vancouver, però només sobreviuen uns pocs centenars de parlants. Hem amagat 30 paraules en Kwak'wala en la següent sopa de lletres, en horitzontal, vertical i diagonal, però només us mostrem 10. El vostre repte és descobrir les 20 paraules que falten i escriure-les al costat del seu significat.

K'W	A	KW	N	I	U	A	KW	A	'M	K'	XW
G	X	T	A	Y	A	Y	A	G	A	Y	U
Y	I	DL	A	X	P'	U	G	D	TL	A	Y
'L	GW	D	U	K'	I	KW	A	DL	I	K'	A
B	A	K	A	Y	U	GW	Y	T'S	DZ	I	B
K'	T'S	'M	A	KW	A	L	A	S	K'	N	I
K	I	DL	A	T'S	I	B	XW	Y	T'L	U	L
Y	'M	K'W	I	P'	A	U	A	T'	U	XW	I
KW	G	A	I	T'S	N	Y	S	'L	B	U	XW
X	TS	N	L	I	P	A	A	A	TS	K	U
D	U	K'	T'	T'S	L	GW	I	G	XW	A	N
XW	S	A	L	A	K	A	Y	A	A	'M	I
I	K'	D	P	G	GW	'M	B	KW	K'W	T'S	T'L
A	K	A	'L	M	U	I	I	A	A	U	I
A	T'S	KW	DZ	A	'L	X	TS	P	T'	K'	K
KW	A	K	I	TL	A	GW	A	K	M	A	TL



(E) Sopa de lletres en Kwak'wala (2/3)

Nota: Algunes seqüències de dos caràcters, com tl, dl, kw, gw i ts, són tractades com a lletres simples en l'alfabet Kwak'wala. A representa el so “uh”. K i G es pronuncien com a “k” i “g” forta, però amb la llengua molt cap a arrere en la boca, cap a la gola. X es pronuncia com una “h” lleugerament aspirada (sense arribar al so de la “j” castellana), i X com la “ch” en “Bach” (una “h” aspirada més pròxima a la “j”). Un apòstrof indica que el so es pronuncia amb un increment de la pressió el final de la gola.

EI. Introduiu cada caràcter, tal com apareix en la paraula de la sopa de lletres, en cada cel·la.

Planxa													
Pastís de fruites del bosc	'L	<u>A</u>	<u>G</u>	<u>A</u>	KW								
Bol d'oli de bacallà													
Granera	<u>X</u>	I	GW	A	I	O							
Baralla de naips	L	I	B	A	I	O							
Recollidor													
Sobre													
Expert jugador de naips													
Expert teixidor de punt de calça													
Pescador	K	I	T'L	I	N	O	<u>X</u> W						
Vaixell de pesca													
Menjar per a mullar en oli	T'S	<u>A</u>	P	A	L	A	S						
Cistell de punt de calça	I	<u>A</u>	<u>G</u>	A	T'S	I							
Agulles de fer punt de calça													
Carta (correspondència)	K'	A	D	<u>A</u>	KW								
Llapis o bolígraf													
Roba de punt, p. e. un jersei													
Ser orgullós	TL	<u>A</u>	M	<u>K</u>	A								
Pescar amb xarxa													
Mullar menjar en oli de bacallà													
Planxar	'M	<u>A</u>	KW	A									



(E) Sopa de lletres en Kwak'wala (3/3)

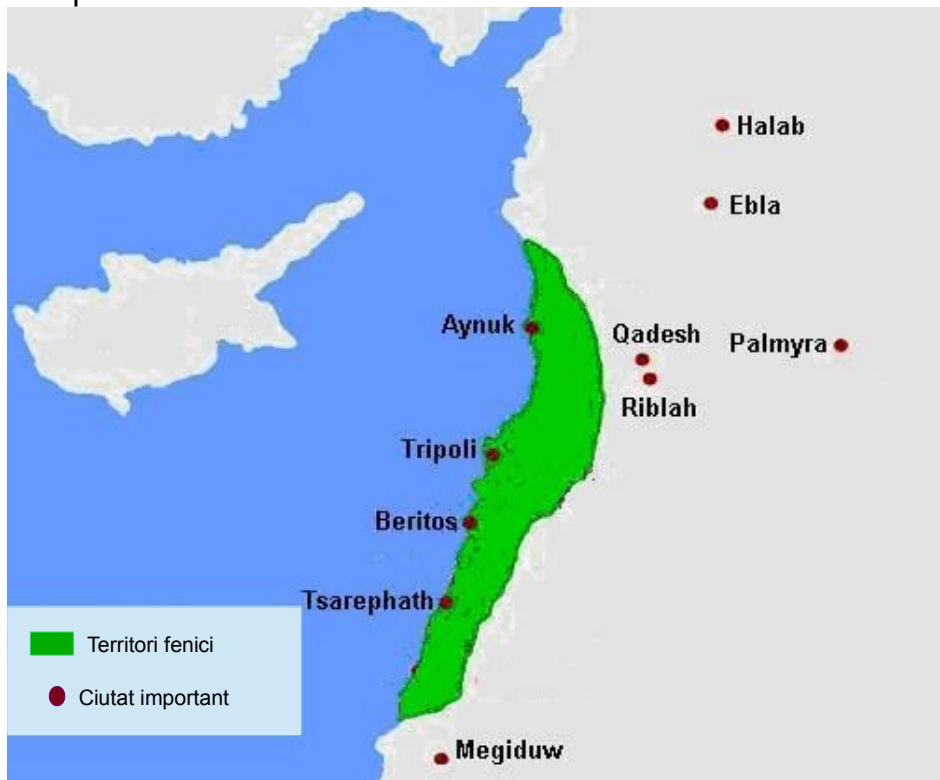
Fer punt de calça										
Fer pastissos de fruites del bosc										
Jugar a naips										
Agranar										
Escriure	K'	A	T	A						
Vaixell d'esbarjo										
Llana										
Roba arrugada										
Escriptor										



(F) Fans del fenici! (1/2)

L'escriptura fenícia pot ser datada al voltant del 1050 a. de C., i d'ella van evolucionar l'escriptura àrab, l'hebrea, i per extensió, la grega, la llatina i la ciríl·lica.

La civilització fenícia es concentrava al llarg de la costa mediterrània, en una zona coneguda com Canaan. El mapa següent mostra un nombre de ciutats fenícies i altres properes que van ser importants socis comercials. Els noms reflecteixen la pronunciació fenícia. No obstant això, dos de les ciutats del mapa es mostren amb els seus noms moderns, que són molt diferents de com eren conegudes en temps dels fenicis.



Este mapa té copyright © mapasof.net i es reproduceix sota la llicència Creative Commons Attribution-ShareAlike 1.0. En algunes impressores en blanc i negre, els colors del mar i del territori fenici són difícils de distingir. El territori fenici és l'àrea que inclou les següents ciutats: Aynuk, Tripoli, Beritos i Tsarephat.

FI. Relaciona els noms fenicis de la llista següent amb els de el mapa. Recorda, dos dels noms no estan relacionats, per la qual cosa dos hauran de quedar-te sense marcar.

1	⊕ 7 4 ʳ	6	ʳ ʳ ʳ ʳ
2	ʳ ʳ ʳ	7	ʳ ʳ ʳ
3	ʳ ʳ ʳ ʳ	8	ʳ ⊕ ʳ
4	ʳ ʳ ʳ ʳ	9	ʳ ʳ ʳ
5	ʳ ʳ ʳ ⊕	10	ʳ ʳ ʳ ʳ



(F) Fans del fenici! (2/2)

Emplena amb el nombre corresponent, o usa la lletra X si es tracta d'alguna d'aquestes ciutats no relacionades.

Aynuk	
Beritos	
Ebla	
Halab	
Megiduw	
Palmyra	
Qadesh	
Riblah	
Tripoli	
Tsarephath	

F2. Escriu com es pronunciaven els noms de les dues ciutats fenícies que han quedat sense relacionar. Explica breument quin element manca i per què.