fi

Kahdennetkymmenennetensimmäiset kansainväliset kielitieteen olympialaiset

Brasília (Brasilia), 23.–31. heinäkuuta 2024

Joukkuekilpailun tehtävä

Leksikostatistiikka on joukko menetelmiä, joiden tarkoituksena on arvioida, kuinka läheistä sukua kielet ovat toisilleen niiden sanastojen perusteella. Näitä menetelmiä sovelletaan tavallisesti pitkiin sanalistoihin, jotka asiantuntijat merkitsevät manuaalisesti ilmoittamalla, oletetaanko tietyn sanaparin olevan peräisin samasta lähteestä. Joskus kuitenkin kielitieteilijät käyttävät leksikostatistisia menetelmiä sanalistoihin, jotka on merkitty automaattisesti. Yksi tällainen automaattinen menettely perustuu *konsonanttiluokkien* käsitteeseen, jonka neuvostoliittolais-israelilainen kielitieteilijä Aharon Dolgopolsky esitti vuonna 1964.

P.	p b ɓ ɸ β f v	K.	k g x ɣ q ɢ χ ɯ	Y.	j ç (juuren alussa)	M.	m ɱ
T.	t d ɗ θ ð ʈ ɖ	R.	r ɾ ɽ ɹ l ɬ ɮ ɭ ʎ ɬ	W.	w ʍ (juuren alussa)	N.	n ɳ ɲ ŋ
S.	s z ʃ ʒ ʂ ʐ ɕ ʑ c ɟ					Q.	t͡ɬ d͡ɮ
H.	ħ ʕ ʜ ʢ ʡ h ɦ ʔ, vokaalit ja j ç w ʍ (paitsi juuren alussa)						

Dolgopolskyn konsonanttiluokat

Alla on merkittyjen sanalistojen katkelmia useista maailman kielikunnista. Merkinnät annetaan alaindeksein. Näiden listojen perusteella on luotu kielten sukupuut käyttäen kahta yksinkertaistettua versiota ns. *StarlingNJ*-algoritmista, minkä lisäksi jokaiselle sanalle on annettu stabiilisuusindeksi. Yläosan puut ja stabiilisuusindeksit perustuvat manuaalisesti merkittyihin sanalistoihin ja alaosan puut ja stabiilisuusindeksit sanalistoihin, jotka on merkitty automaattisesti. Jokaisen sanalistan perusteella on rakennettu kaksi puuta, jotka seuraavat kahta algoritmiversiota: Algoritmi A ja Algoritmi B. Huomatkaa, että joissakin tapauksissa on useita mahdollisia puita, jotka vastaavat yhtä sanalistaa; näissä tapauksissa vain yksi puu valittiin satunnaisesti. Jokaisen puun jokaiselle noodille (solmulle) on määritelty leksikostatistinen etäisyys. Mitä suurempi etäisyys, sitä läheisempi on kielten välinen suhde. Tarkempi termi olisi siis ”käänteinen leksikostatistinen etäisyys”, ei ”leksikostatistinen etäisyys”. Yksinkertaisuuden vuoksi käytämme tässä tehtävässä termiä ”leksikostatistinen etäisyys”.

Sekä stabiilisuusindeksit että leksikostatistiset etäisyydet pyöristetään kahteen desimaaliin. Jos kolmas numero desimaalipilkun jälkeen on pienempi kuin 5, arvo pyöristetään alaspäin; muuten se pyöristetään ylöspäin. Esimerkiksi, 2,836 pyöristetään 2,84:ään, 0,705 pyöristetään 0,71:een, ja 0,703 pyöristetään 0,70:een. Pyöristäminen pätee vain arvoihin, jotka näytetään ihmislukijoille. Toisin sanoen, tietokone, joka käyttää algoritmeja, ”näkee” pyöristämättömät arvot.

Huomatkaa, että joidenkin sanojen tiedetään tai epäillään olevan lainattu muista kielistä. Esimerkiksi kadiveun kielen sana **jokːi** ’suola’ on lainattu guaranin kielen sanasta **jukɨ** ja ’iipain kielen (Mesa Granden murteen) sana **ʔaːnʲ** ’vuosi’ on lainattu espanjan kielen sanasta **ˈaɲo**.

Joissakin tapauksissa sanalistoissa annetaan yhdelle merkitykselle useita synonyymejä, jotka erotetaan toisistaan pilkulla. Esimerkkinä mainittakoon vehvosin kielen ‘jalkaa’ tarkoittavat sanat.

Alla olevissa tiedoissa kaikki prefiksit on erotettu ”=”-merkillä ja kaikki suffiksit ”-”-merkillä. Joitain sanoja käytetään vain prefiksien kanssa. Nämä alkavat ”=”-merkillä.

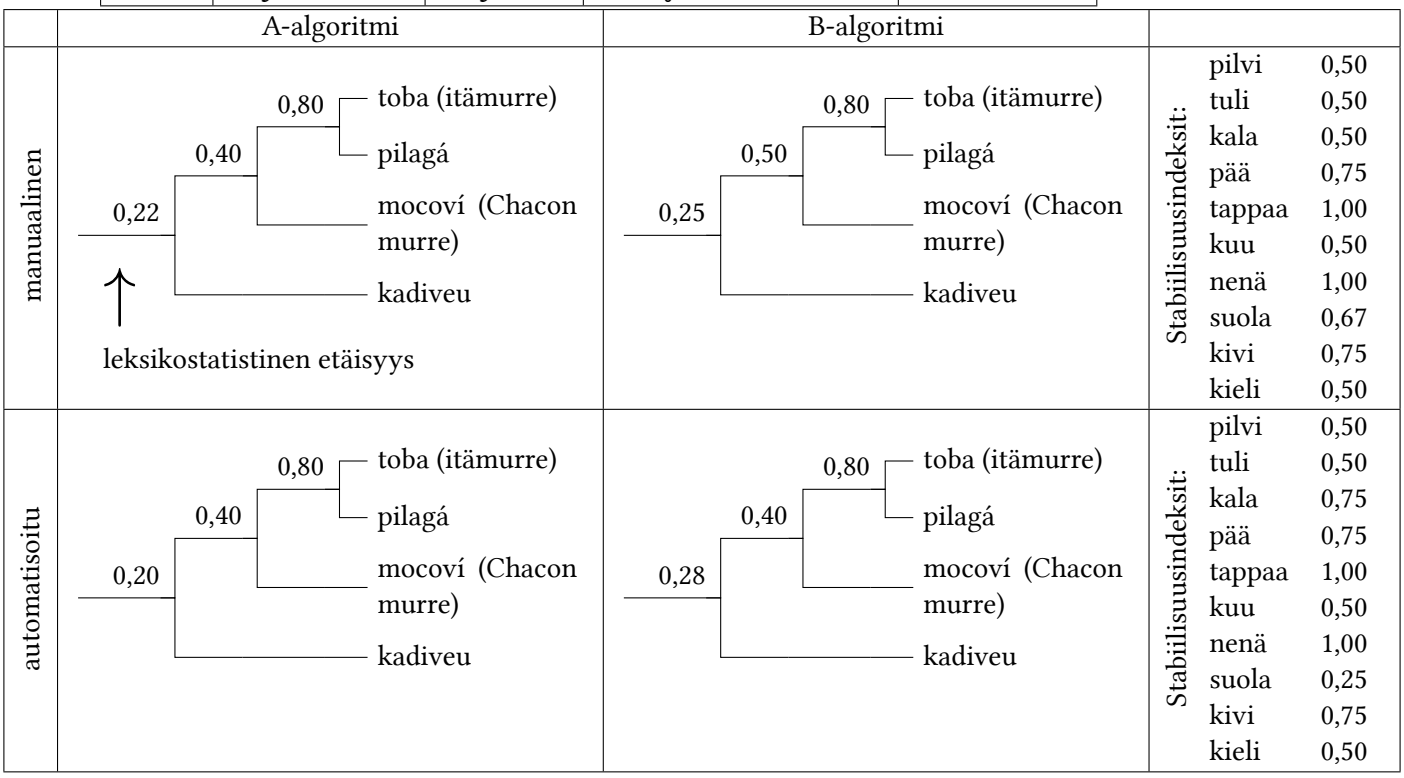
Tiedot on kirjoitettu kansainvälisiä foneettisia aakkosia käyttäen. ˈ = pääpaino, ˌ = sivupaino (pääpainoa heikompi), ◌ː = pitkä äänne, ◌̆ = ylilyhyt äänne, $\widehat{XY}$ =X ja Y äännetään yhtenä äänteenä, ◌́ = korkea tooni, ◌̀ = matala tooni, ◌̂ = laskeva tooni, ˀ◌ = preglottalisoitu äänne (jota edeltää lyhyt ilmavirran estyminen kurkussa), ◌’ = ejektiivi (konsonantti, jossa kurkunpää suljetaan ja sitten avataan),

◌̥ = soinniton konsonantti, ◌̃ = nasalisoitu äänne (nenän kautta lausuttava äänne), ◌̰ = narinaääni, ⁿ◌ osoittaa ilmavirtaa nenän kautta konsonantin edellä, ◌ʰ = aspiroitu (ilmaa puhaltaen lausuttava) konsonantti, ◌ʷ = labialisoitu (pyöristetyillä huulilla lausuttava) konsonantti, ◌ʲ = palatalisoitu (kielen keskiosaa kovaan kitalakeen lähestyen lausuttava) konsonantti. **ɑ**, **æ**, **ɛ**, **ɪ**, **ɨ**, **ɔ**, **ʊ**, **ʉ**, **ə**, **ʌ**, **ɒ**, **ɘ**, **y**, **ɵ**, **ø** ovat vokaaleja. Muut erikoismerkit ovat konsonantteja.

⚠ Minkään tehtävässä mainitun kielen osaamisesta ei ole apua tehtävää ratkaistaessa.

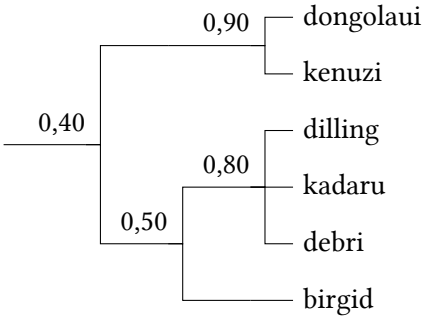
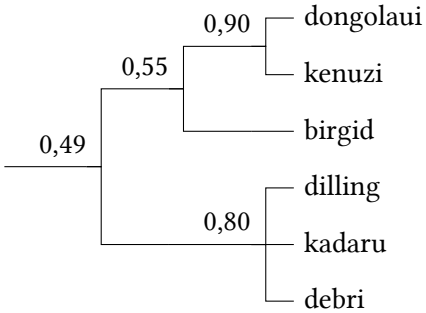
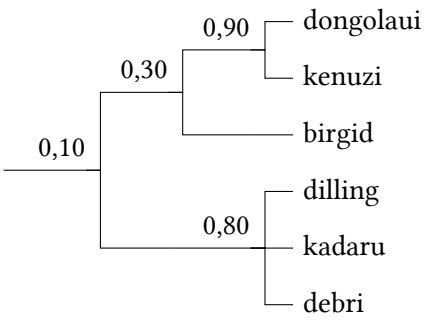
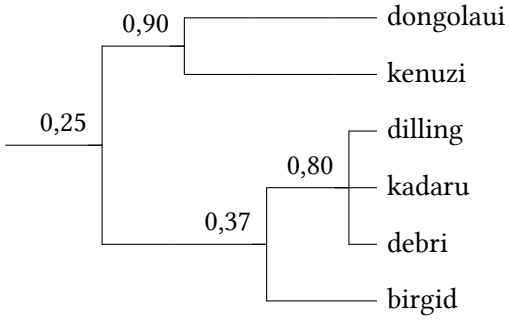
I. osa. Guaikuru-kielet (Argentiina, Brasilia, Paraguay)

	toba (itämurre)	pilagá	mocoví (Chacon murre)	kadiveu
pilvi	l=ʔok$_1$	ˈlo=ʔok$_1$	naweɣelek$_2$	lol:adi$_3$
tuli	nodek$_1$	ˈd=oleʔ$_2$	noɾek$_1$	n=ol:edi$_2$
kala	njaq$_1$	ˈnijaq$_1$	naʎin$_2$	nij:oɢo-d͡ʒegi$_3$
pää	=qajk$_1$	=ˈqajk$_1$	=qaik$_1$	=ak:ilo$_2$
tappaa	=alawat$_1$	=aˈla:t$_1$	=alawat$_1$	=el:owadi$_1$
kuu	ʔawoʁojk$_1$	ʔaˈwoʕojk$_1$	ʃiɾajɣo$_2$	ep:enaj$_3$
nenä	=mik$_1$	=ˈmik$_1$	=mik$_1$	=m:iq:o$_1$
suola	towe$_1$	olˈɣek$_2$	ʔwe$_1$	jok:i$_{-1}$
kivi	qaʔ$_1$	ˈqaʔ$_1$	qaʔ$_1$	wet:iɢa$_2$
kieli	=at͡ʃ-aʁat$_1$	=aˈt͡ʃ-aʕat$_1$	=oʔleɣ-aʁan-aʁat$_2$	=ok:el:i$_3$



II. osa. Nubialainen kielikunta (Egypti, Sudan)

	dongolaui	kenuzi	dilling	kadaru	debri	birgid
tappaa	ˈbɛː$_1$	beː$_1$	hur$_2$	wur-i$_2$	wur-i$_2$	fila:l-e$_1$
kuu	uˈn-at-t$_1$	an-at-ti$_1$	nɔn-ti$_1$	nɔn-tu$_1$	nɔn-to$_1$	ma:l$_2$
vesi	ˈɛss$_1$	essi$_1$	ɔti$_1$	ɔto$_1$	ɔtu$_1$	eɟi$_1$
antaa	ˈtɪr$_1$	tir$_1$	ti$_1$	ti$_1$	ti$_1$	teː-n$_1$
hyvä	ˈsɛrɛː$_1$	sereː$_1$	ken$_2$	kɛn$_2$	kɛŋ$_2$	azze-n$_3$
tuuli	ˈturug$_1$	turug$_1$	irʃ-i$_2$	irʃ-o$_2$	irʃ-o$_2$	kurr-i$_3$
tukka	ˈdɪl-tɪ$_1$	siːr$_2$	tel-ti$_1$	til-tu$_1$	til-tu$_1$	ur=dill-e$_1$
vatsa	ˈtuː$_1$	tuː$_1$	te-te$_2$	to$_1$	to$_1$	tuː$_1$
nukkua	ˈnɛːr$_1$	neːr$_1$	ɟer$_1$	dwallɛli$_2$	ɟer-i$_1$	neːr-i$_1$
aurinko	ˈmasɪl$_1$	masɪl$_1$	ɛɟ$_2$	aɟu$_2$	eŋgal-to$_3$	ʔiːzi$_2$

	A-algoritmi	B-algoritmi	
manuaalinen			Stabiilisuusindeksit: tappaa 0,50 kuu 0,83 vesi 1,00 antaa 1,00 hyvä 0,50 tuuli 0,50 tukka 0,83 vatsa 0,83 nukkua 0,83 aurinko 0,50
automatisoitu			Stabiilisuusindeksit: tappaa 0,33 kuu 0,50 vesi 0,50 antaa 0,67 hyvä 0,50 tuuli 0,50 tukka 0,83 vatsa 1,00 nukkua 0,50 aurinko 0,50

(A) **(2 pistettä)** **ʁ** lausutaan kuin ranskan *r*, kielen takaosassa. Mihin Dolgopolsky-luokkaan se kuuluu ja miten saitte sen selville?

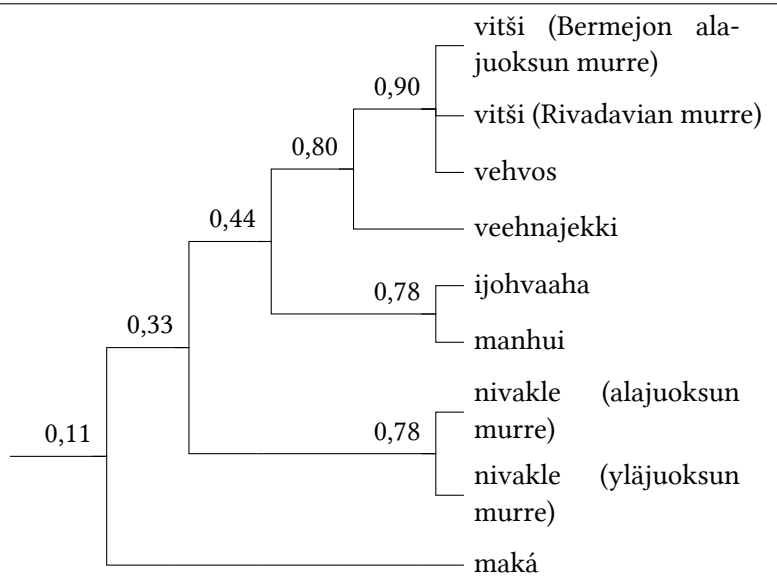
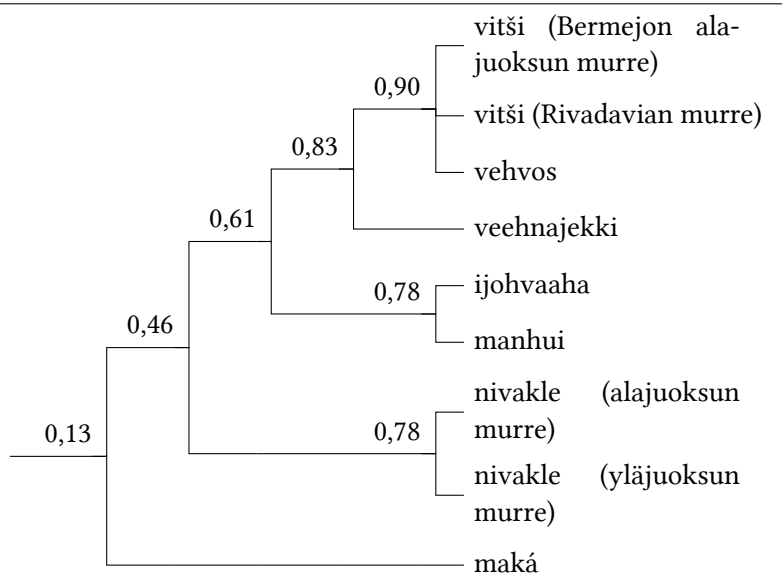
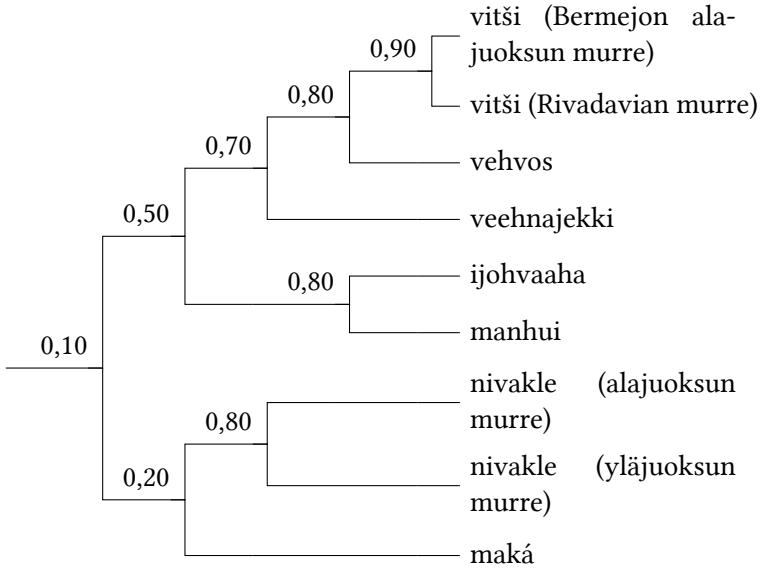
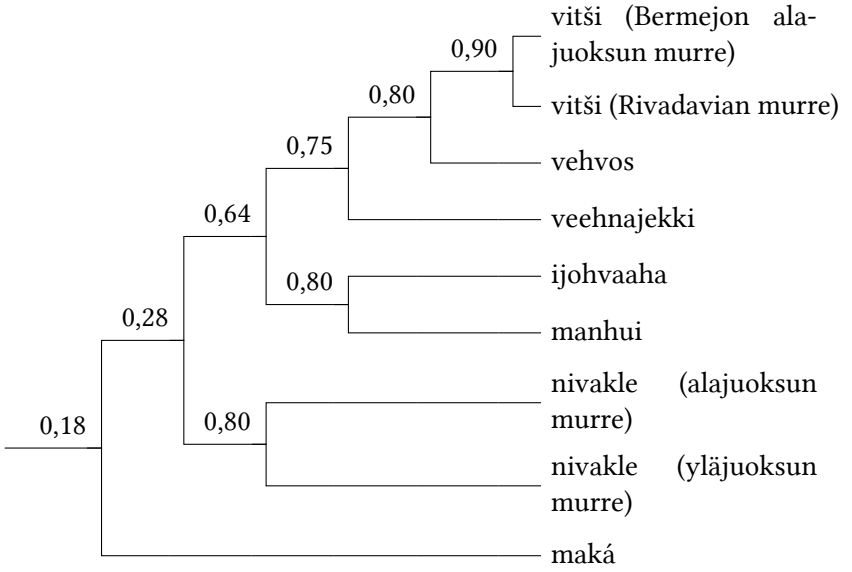
(B) **(2 pistettä)** Vasemmassa yläkulmassa oleva nubialaisten kielten puu on vain yksi kahdesta mahdollisesta puusta tässä algoritmin ja merkintätyypin yhdistelmässä. Piirtäkää toinen mahdollinen puu.

(C) **(2 pistettä)** Vasemmassa alakulmassa oleva nubialaisten kielten puu on vain yksi kahdesta mahdollisesta puusta tässä algoritmin ja merkintätyypin yhdistelmässä. Piirtäkää toinen mahdollinen puu.

(D) **(2 pistettä)** Leksikostatistinen etäisyys 0,49 (joka on määritetty nubialaisten kielten oikeassa yläkulmassa olevan puun juurelle) on pyöristetty kahteen desimaaliin, kuten jotkut muutkin etäisyydet tässä tehtävässä. Mikä on tarkka etäisyys?

III. osa. Mataguaiolainen kielikunta (Argentiina, Bolivia, Paraguay)

	vitši (Bermejon alajuoksun murre)	vitši (Rivadavian murre)	vehvos	veehnajekki	ijohvaaha	manhui	nivakle (alajuoksun murre)	nivakle (yläjuoksun murre)	maká
tuli	ʔitoχ$_{1}$	ʔitɔx$_{1}$	ʔitɑh$_{1}$	ʔiːtɑχ$_{1}$	ˈh͡wat$_{2}$	ˈʔei̯t^{j}e$_{1}$	ʔitɑx$_{1}$	ʔitɑx$_{1}$	feʔt$_{2}$
kala	ʔwahat$_{1}$	wahat$_{1}$	wahat$_{1}$	ʔwaːhat$_{1}$	siˈʔjus$_{-1}$	ʃiˈʔjus$_{-1}$	saxet͡ʃ$_{-1}$	saxet͡ʃ$_{-1}$	seheːt͡s$_{-1}$
jalka	=pat͡ʃu$_{1}$	=qɔlɔ$_{2}$	=pat͡ʃo$_{1}$, =kɑla$_{2}$	=paːkʲ’oʔ$_{1}$	=ˈsat$_{3}$	=kaˈʔlaʔ$_{2}$	=ɸoʔ$_{4}$	=ɸoʔ$_{4}$	=f’iʔ$_{5}$
vesi	ʔinot$_{1}$	ʔinɔt$_{1}$	wah$_{2}$	ʔinɑːt$_{1}$	ʔiˈnʲat$_{1}$	ʔaˈʔnat$_{1}$	jinɑʔt$_{1}$	jinɑʔt$_{1}$	iweliʔ$_{3}$
antaa	=ʔwen̥-u$_{1}$	=wen̥-u$_{1}$	=ʔwen̥-o$_{1}$	=ʔwen̥-oʔ$_{1}$	=ˈwɛh͡n-aʔm$_{2}$	=ˈhajʔ$_{3}$, =ˈwɛn$_{2}$	=xut$_{4}$	=xut-ej$_{4}$	tis-ix$_{5}$
hyvä	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ˈʔes$_{1}$	ˈʔei̯s$_{1}$	ʔis$_{1}$	ʔis$_{1}$	t=ejk’un-ej$_{2}$
tuuli	ʔinwokʷ$_{1}$	ʔinwɔk$_{1}$	ʔihwokʷ$_{1}$	=jaːɬ$_{2}$, =xʷoxʷ$_{3}$	ˈh͡lah͡wuʔ$_{4}$	ˈh͡lah͡wuuʔ$_{4}$	ɬaβiʔm$_{5}$	ɬaβiʔm$_{5}$	t’unik’i$_{6}$
puu	haʔlo$_{1}$	halɔ$_{1}$	haʔlɑ$_{1}$	haʔlɑʔ$_{1}$	ʔaˈʔlaʔ$_{1}$	ʔaˈʔla-k$_{1}$	ʔaʔkxi-juk$_{2}$	jiʔk͡lɑʔ$_{1}$	naxka-k$_{3}$
tukka	=ʔwule-j$_{1}$	=wule-j$_{1}$	=ʔwole-j$_{1}$	=ʔwoːle-ç$_{1}$, hiːlenaχ$_{2}$	=ˈʔwole$_{1}$	=ˈʔwole-j$_{1}$	=ʃateʔt͡ʃ$_{3}$	=jeʔs$_{4}$	=ʔewkux-it͡s$_{5}$
tappaa	=lon$_{1}$	=lɔn$_{1}$	=lɑn$_{1}$	=lɑːn̥$_{1}$	=ˈlaʔan$_{1}$	=ˈlan$_{1}$	=k͡lɑn$_{1}$	=k͡lɑn$_{1}$	=lan$_{1}$

	A-algoritmi	B-algoritmi	
manuaalinen			Stabilisuusindeksit: tuli 0,78 kala 1,00 jalka 0,33 vesi 0,78 antaa 0,44 hyvä 0,89 tuuli 0,33 puu 0,78 tukka 0,67 tappaa 1,00
automatisoitu			Stabilisuusindeksit: tuli 0,78 kala 0,44 jalka 0,33 vesi 0,56 antaa 0,67 hyvä 0,89 tuuli 0,22 puu 0,67 tukka 0,67 tappaa 1,00

IV. osa. Mongolilainen kielikunta (Kiinan kansantasavalta, Mongolia, Venäjä)

(E) (10 pistettä) Tarkastelkaa seuraavaa sanalistaa. Laskekaa sekä manuaalisia että automatisoituja merkintöjä vastaavat stabiilisuusindeksit.

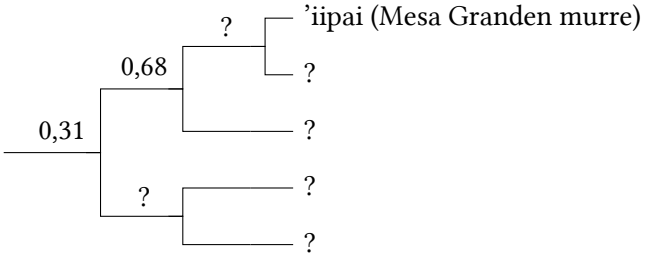
Auttaaksemme teitä, olemme jo laskeneet molemmat stabiilisuusindeksit ’kaikki’-sanalle. Satunnaisessa järjestyksessä nämä ovat 0,36 ja 0,40.

	daguuri (hailarin murre)	hamnigani (mantšun murre)	burjaatti (horin murre)	uus-barguutti	ööldi	hošuutti	kalmukki	halha	ordos	šira-juguuri	baoan
kaikki	hɔː$_{1}$	bʊlt$_{2}$	bʊxɨː$_{3}$	bygd$_{4}$	t͡sug$_{5}$	lug$_{5}$	t͡suk$_{5}$, xamak$_{-1}$	pux$_{3}$, pugt$_{4}$, xamăɢ$_{-1}$	pyɣyte$_{4}$, xamuk^{h}$_{-1}$	t͡ʃ^{h}uq$_{5}$	hanə$_{-2}$
kaarna	hai̯l̥s$_{1}$	qalihʊn$_{1}$	χɔltəhɔn$_{2}$	xal^{j}huː$_{1}$	xolts$_{2}$	xalis$_{1}$	dursn$_{3}$	xɔɮt^{h}ɔ̆s$_{2}$	turusu$_{3}$	χalsən$_{1}$	arasun$_{4}$
vatsa	keːli$_{1}$	gətəhʉn$_{2}$	gedehen$_{2}$	gedyː$_{2}$	geːs$_{2}$	gets$_{2}$	gesn$_{2}$	gitĭs$_{2}$, xiwɮĭj$_{-1}$	ketysy$_{2}$	ketesən$_{2}$	kele$_{1}$
lintu	dəgi$_{-1}$	ɕiwan$_{1}$	ʃubuːn$_{1}$	ʃuwuː$_{1}$	ʃuvuː$_{1}$	ʃuwuː$_{1}$	ʃowun$_{1}$	ʃuwu$_{1}$	ʃuβuː$_{1}$	ʃuːn$_{1}$, pelt͡ʃɔr$_{2}$	bend͡ʑer$_{2}$
tuli	ɢali$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	qal$_{1}$	qal$_{1}$	χal$_{1}$
tie	terg-uːl$_{1}$	qarɢʊ̆i$_{2}$	χarɢɨ$_{2}$, zam$_{-1}$	zam$_{-1}$	d͡zam$_{-1}$	d͡zam$_{-1}$	xaː-lʁə$_{3}$	t͡sam$_{-1}$	t͡ʃam$_{-1}$	mør$_{4}$	mor$_{4}$
suola	hataː$_{1}$	dawhʊn$_{2}$	dabhan$_{2}$	dawuhuː$_{2}$	daws$_{2}$	daws$_{2}$	dawsn$_{2}$	tawsă$_{2}$	taβusu$_{2}$	taːpsən$_{2}$	dabsuŋ$_{2}$
uida	unpa-du$_{1}$	ʊmba-$_{1}$	t^{h}amar-$_{2}$	umb-$_{1}$	sele-$_{3}$	umba-$_{1}$	us-t͡ɕi-$_{4}$, øːm-$_{5}$	siɮĭ-$_{3}$	usu-t͡ʃ^{h}i-la-$_{4}$	umpa-$_{1}$	mba-$_{1}$
vesi	ɔsɔ$_{1}$	ʊxʊn$_{1}$	uhan$_{1}$	uːha$_{1}$	usn$_{1}$	us$_{1}$	usn$_{1}$	ʊsŏ$_{1}$	usun$_{1}$	q^{h}usun$_{1}$	sə$_{1}$
tuuli	kei̯n$_{1}$	halkin$_{2}$	halxin$_{2}$	halxi$_{2}$	sal^{j}xin$_{2}$	salk^{j}i$_{2}$	sal^{j}kn$_{2}$	saɮxĭ$_{2}$	k^{h}iː$_{1}$	k^{h}iː$_{1}$	ki$_{1}$

V. osa. Jumalainen kielikunta (Meksiko, Yhdysvallat)

(F) **(8 pistettä)** Tarkastelkaa seuraavaa sanalistaa. Alla on puu, joka on rakennettu saman sanalistan perusteella. Jotkut tiedot (kielten nimet ja leksikostatistiset etäisyydet) puuttuvat. Täydentäkää aukot. Ilmoittakaa, onko puu manuaalinen vai automaattinen, sekä onko se luotu algoritmilla A vai B.

	mohave	kokopa	javapai	tiipai (Jamulin murre)	’iipai (Mesa Granden murre)
lyhyt	wena=wen-a_1	ˈxʌ=ʔut$_2$	ˈt͡ʃkr=ot-i_2	ləˈʔuɲ$_1$	mə=put-k_3
lintu	ʔit͡ʃ=i=jer$_1$	ˈʃa$_2$	ˈʔ=t͡ʃ=sa$_2$	aʔˈʃa$_2$	ʔaː=ʂaː$_2$
luu	ɲ=a=s=ak$_1$	ˈɲ=j=aːk$_1$	ˈt͡ʃ=j=aːk-a_1	ˈak$_1$	aq$_1$
kuiva	i=roː-v-k$_1$	ˈʂ=ʔar$_2$	ˈru-β-i$_1$	ˈs=ʔaːj$_3$	saːj$_3$
liha	kʷiːkʷay$_1$	ʔi=ˈmaːt͡ʃ$_2$	ˌkʷeːˈθo-β-a_3	ˈkʷak$_4$	kukʷaːj-p$_1$
kaula	maʌaqe$_1$	ˈm=puk$_2$	ˈmlqíˀ$_1$	iːˈpuk$_2$	iː=puk$_2$
nähdä	i=juː-k$_1$	ˈwiː$_2$	ˈʔuː$_1$	ˈwiːw$_2$	ə=wuːw$_2$
häntä	iː=ʔar$_1$	ˈʃ=juʂ$_2$	ˈβ=hé$_3$	ʃəˈjuɬ$_2$	xə=juɬ$_2$
kaksi	havik-k$_1$	ˈx=wak$_1$	ˈh͡wâk-i_1	xəˈwak$_1$	xə=wak$_1$
vuosi	huːðe$_1$	ˌmat-ˈkaːm$_2$	ˈʔ=t͡ʃʰur-a_3	matˈwam$_2$	ʔaːnʲ$_{-1}$



(G) **(20 pistettä)** On luotu jotkut muut jumalaisten kielten puut, joita vastaavat seuraavat puun juuressa (jokaisesta puusta kauimpana vasemmalla) olevat leksikostatistiset etäisyydet:

1. 0,20
2. 0,23
3. 0,24

Piirtäkää jokainen näistä puista. Ilmoittakaa kunkin puun osalta, onko se manuaalinen vai automaattinen, sekä onko se luotu algoritmilla A vai B.

(H) **(3 pistettä)** Kaksi **(G)**-alitehtävässä luetelluista etäisyyksistä on pyöristetty kahteen desimaaliin: 0,23 on pyöristetty 0,225:stä. Mikä muu etäisyys on pyöristetty ja mikä on sen tarkka arvo?

(I) **(4 pistettä)** Selittäkää, miten stabiilisuusindeksit lasketaan.

(J) **(5 pistettä)** Selittäkää, miten leksikostatistiset etäisyydet lasketaan.

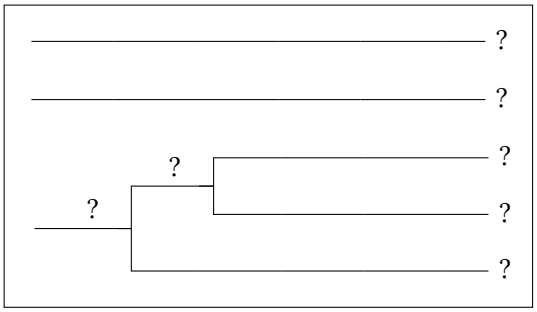
(K) **(4 pistettä)** Selittäkää A- ja B-algoritmien välinen ero.

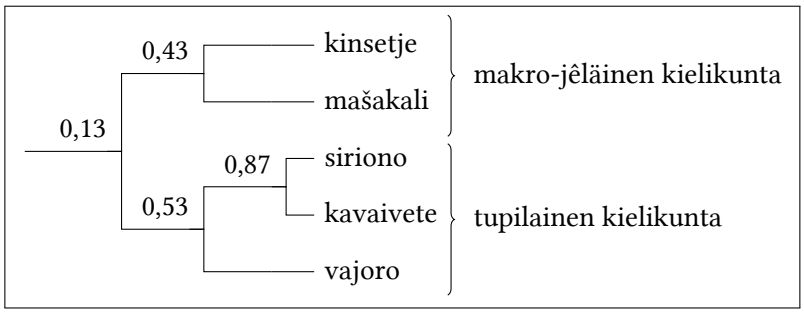
VI. osa. Makro-jêläinen kielikunta ja tupilainen kielikunta (Brasilia, Bolivia)

(L) **(28 pistettä)** Makro-jêläiset ja tupilaiset kielet muodostavat kaksi suurta kielikuntaa, joihin kuuluvia kieliä puhutaan Etelä-Amerikassa. Jotkut kielitieteilijät uskovat niiden olevan kaukaisia sukulaisia. Tarkastelkaa seuraavia sanalistoja.

	A	B	Γ	Δ	E
kaarna	e=ˈe-ke	hʷĩ=ˈkʰʌ	kʉp=ˈpɛ	mĩβ͡m=ˈt͡ɕɑj	=ˈpɛ
vatsa	ˈe=ɾje	=ˈʈʰigi	=ãˈɯ̃n	=ˈtæj	=ɾɛˈwɛk
veri	e=ˈɾuki	=kaˈⁿbɽo	=d͡ʒ=aˈʉ	=ˈhɛβ͡p	=ɾuˈɨ
polttaa	=ˈɾa͡i	=ɾɔ=ˈkʰɹə̃	=poˈkʷa	mũ=...=ˈhaβ͡p	=ɾaˈpɨ
rasva	e=ˈkiɾa	=ˈʈʰwɘmɨ	=ˈd͡ʒ=ap	=ˈtuβ͡p	=ˈkap
jalka	ˈe=i	=ˈhʷaji	=ˈβi	=pɒˈta	=ˈpɨ
käsi	ˈe=o	=ɲɨ̃ˈkʰɹa	=ˈβo	=ˈɲĩβ͡m	=ˈpɔ
raskas	e=ˈusi	=wɨˈʈʰĩ	=poˈti	=β͡pˈtəj	=pɔˈɨj
maksa	ˈe=ja	=ˈⁿba	=piˈa	=t͡ɕiβ͡pkĩˈnãj	=pɨˈʔa
uusi	e=ˈjasu	=ˈⁿdɨwɨ	=paˈgop	=ˈtɨβ͡p	=pɨaˈu
juuri	e=ˈɾao	=jaˈɽe	kʉp=kʉjoˈpɛ	mĩβ͡m=ɲĩβ͡m=t͡ɕaˈti͡ə	=ɾaˈpɔ
iho	ˈe=ɨ	=ˈkʰʌ	=ˈpɛ	=ˈt͡ɕɑj	=ˈpit
häntä	e=ˈɾoko͡i	=ˈⁿbɨ	=d͡ʒ=oˈkʷaj	=nã:=ˈkɨβ͡p	=ˈɾaj
valkoinen	ˈe=ʃɨ̃	=jaˈkʰa	=d͡ʒiˈɾa	=β͡pˈdoɯ	=ˈsĩŋ
siipi	e=ˈheo	=jaˈɽa	=pɛˈo	=ɲĩˈmãɯ	=pɛˈpɔ, =jɨˈwa

Alla on kaksi puuta, jotka on rakennettu samojen sanalistojen perusteella. Jotkut tiedot (kielten nimet ja leksikostatistiset etäisyydet) puuttuvat. Täydentäkää aukot. Ilmoittakaa kunkin puun osalta, onko se manuaalinen vai automaattinen, sekä onko se luotu algoritmilla A vai B.





A	B	Γ	Δ	E
?	?	?	?	?

⚠ Manuaaliset merkinnät ja stabiilisuusindeksit jätettiin tarkoituksella pois tästä alitehtävästä.

(M) (10 pistettä) Dolgopolsky-luokkiin perustuvat automaattiset menettelyt voivat tuottaa vääriä tuloksia. Tässä esimerkissä automaattinen menettely havaitsee enemmän yhtäläisyyksiä sirionon kielen ja tietyn makro-jêläisen kielen (kinsetjen kielen) välillä kuin sirionon ja muiden tupilaisten kielten välillä. Ehdottakaa muokattua automaattista menettelyä, joka tuottaisi oikean makro-jêläisten ja tupilaisten kielten luokittelun, jos sitä sovellettaisiin yllä oleviin sanalistoihin, ja kuvatkaa tämä *lyhyesti*.

⚠ Tämä alitehtävä arvostellaan ainoastaan siinä tapauksessa, että parhaat pisteet saavuttaneet joukkueet päätyvät tasapisteisiin.

Kiitokset Alejandra Vidalille, Maria Konoshenkolle, Ilja Gruntoville ja Jamthô Suyálle muutamia kieliä koskeviin kysymyksiin vastaamisesta. *—Andrei Nikulin, Milena Veneva*

Toimittajat: Ivan Deržanski (tekn. toim.), Hugh Dobbs, Stanislav Gurevitš, Boris Iomdin, Liam McKnight, Andrei Nikulin (päätoimittaja), Aleksejs Pegušev, Jan Petr, Aleksandr Piperski, Marija Rubinštein, Milena Veneva, Elysia Warner.

Suomenkielinen teksti: Andrei Nikulin, Seppo Kittilä.

Onnea kilpailuun!