de

Einundzwanzigste Internationale Linguistik-Olympiade

Brasília (Brasilien), 23.–31. Juli 2024

Aufgabe des Gruppenwettbewerbs

Die Lexikostatistik ist eine Sammlung von Methoden, die entwickelt wurden, um auf der Grundlage des Wortschatzes abzuschätzen, wie eng verwandt beliebige Sprachen sind. Normalerweise werden diese Methoden auf langen, manuell von Experten annotierten Wortlisten angewendet, wobei die Experten auch angeben, ob zwei bestimmte Wörter vermutlich aus derselben Quelle stammen. Manchmal allerdings verwenden Linguisten lexikostatistische Methoden mit Wortlisten, die durch automatisierte Verfahren annotiert werden. Ein solches automatisiertes Annotationsverfahren basiert auf dem Konzept der *Konsonantenklassen*, die 1964 vom sowjetisch-israelischen Linguisten Aharon Dolgopolsky eingeführt wurden.

P.	p b ɓ ɸ β f v	K.	k g x ɣ q ɢ χ ʯ	Y.	j ç (am Anfang der Wurzel)	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 (am Anfang der Wurzel)	N.	n ɳ ɲ ŋ
S.	s z ʃ ʒ ʂ ʐ ɕ ʑ c ɟ					Q.	t͡ɬ d͡ɮ
H.	ħ ʕ ʜ ʢ h ɦ ʔ, Vokale und j ç w ʍ (außer am Anfang der Wurzel)						

Dolgopolskys Konsonantenklassen

Unten sind annotierte Fragmente von Wortlisten aus mehreren Sprachfamilien der Welt. Die Annotationen werden durch tiefgestellte Ziffern dargestellt. Auf der Grundlage dieser Listen wurden Sprachfamilienstammbäume mithilfe zweier vereinfachter Versionen des sogenannten *StarlingNJ*-Algorithmus gebaut, und ein *Stabilitätsindex* wurde jedem Wort zugewiesen. Die Bäume und Stabilitätsindizes oben basieren auf manuell generierten Wortlisten, und die unten basieren auf automatisch annotierten Wortlisten. Für jede Wortliste gibt es zwei Bäume, die mithilfe zweier Versionen des Algorithmus – Algorithmus A und Algorithmus B – gebaut wurden. Beachtet, dass in einigen Fällen mehrere mögliche Bäume einer Wortliste entsprechen; in solchen Fällen wurde nur ein Baum zufällig ausgewählt. Jedem Knoten in jedem Baum wird eine lexikostatistische Distanz zugewiesen. Je höher die Distanz, desto enger die Verwandtschaft zwischen den Sprachen. Deswegen wäre es präziser den Begriff „umgekehrte lexikostatistische Distanz“ als den Begriff „lexikostatistische Distanz“ zu verwenden. In dieser Aufgabe wird der Begriff „lexikostatistische Distanz“ der Einfachheit halber verwendet.

Sowohl die Stabilitätsindizes als auch die lexikostatistischen Distanzen werden auf zwei Dezimalstellen gerundet. Wenn die dritte Ziffer nach dem Komma kleiner als 5 ist, rundet man ab; andernfalls rundet man auf. Zum Beispiel wird 2,836 auf 2,84, 0,705 auf 0,71, und 0,703 auf 0,70 gerundet. Das Runden gilt nur für die Werte, die menschlichen Lesern gezeigt werden. Anders gesagt kann der Computer, der die Algorithmen ausführt, „sehen“, was die ungerundeten Werte sind.

Beachtet, dass einige Wörter bekanntermaßen oder vermutlich aus anderen Sprachen entlehnt sind. Zum Beispiel wurde das Wort **jokːi** ‚Salz‘ der Kadiwéu-Sprache aus dem Guaraní-Wort **jukɨ** und **ʔaːnʲ** ‚Jahr‘ des Ipai (Mesa Grande) aus dem spanischen **ˈaɲo** entlehnt.

In einigen Fällen werden in den Wortlisten mehrere Synonyme einer einzigen Bedeutung durch Kommata getrennt angegeben. Ein Beispiel ist ‚Fuß‘ im Vejoz.

In den Daten unten werden alle Präfixe durch das „=“-Zeichen und alle Suffixe durch das „-“-Zeichen getrennt. Einige Wörter kommen nur mit Präfixen vor. Diese Wörter beginnen mit dem Zeichen „=“.

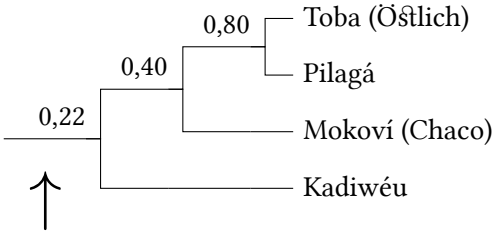
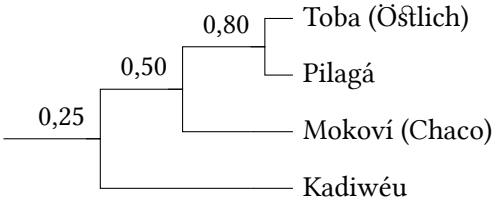
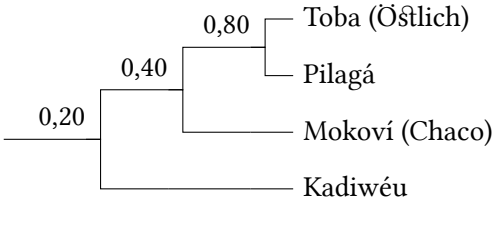
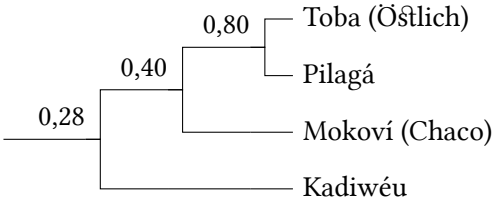
Die Daten werden mit dem Internationalen Phonetischen Alphabet transkribiert. ˈ = Hauptbetonung, ˌ = Nebenbetonung (schwächer als Hauptbetonung), ◌ː = langer Laut, ◌̆ = sehr kurzer Laut, X͡Y =

X und Y werden als ein einziger Laut ausgesprochen, ◌́ = hoher Ton, ◌̀ = tiefer Ton, ◌̂ = fallender Ton, ˀ◌ = vorglottalisierter Laut (nach vorhergehender Blockierung des Luftstroms im Hals ausgesprochen), ◌’ = ejektiver Laut (durch eine kurze Blockierung des Luftstroms im Hals ausgesprochen), ◌̥ = stimmloser Laut, ◌̃ = nasalisierter Laut (durch die Nase ausgesprochen), ◌̰ = knarrende Stimme (ein tiefer, kratziger Stimmklang), ⁿ◌ zeigt an, dass vor dem Konsonanten etwas Luft durch die Nase strömt, ◌ʰ = aspirierter Konsonant (mit Luftstoß ausgesprochen), ◌ʷ = labialisierter Konsonant (mit gerundeten Lippen ausgesprochen), ◌ʲ = palatisierter Laut (durch Hebung eines Teils der Zunge in Richtung des harten Gaumens ausgesprochen). **ɑ**, **æ**, **ɛ**, **ɪ**, **ɨ**, **ɔ**, **ʊ**, **ʉ**, **ə**, **ʌ**, **ɒ**, **ɘ**, **ʏ**, **ɵ**, **ø** sind Vokale. Andere Sonderzeichen sind Konsonanten.

⚠ Kenntnisse irgendeiner der genannten Sprachen bringen bei der Lösung dieser Aufgabe keinen Vorteil.

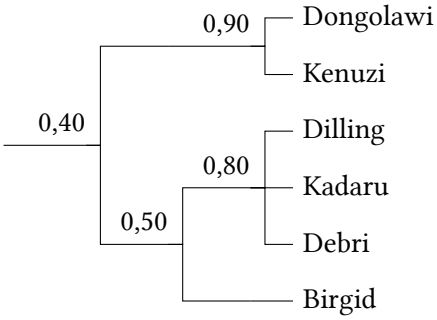
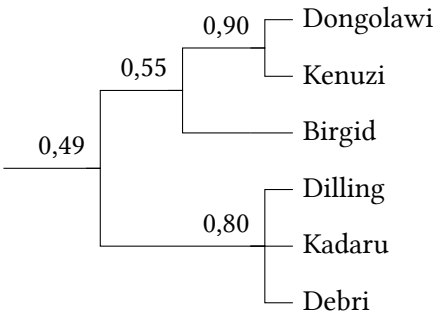
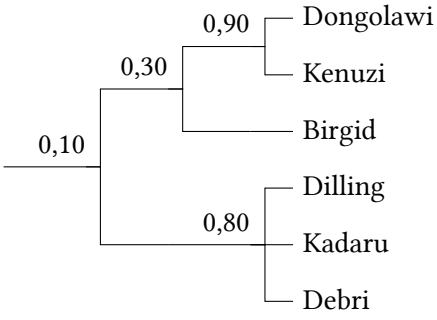
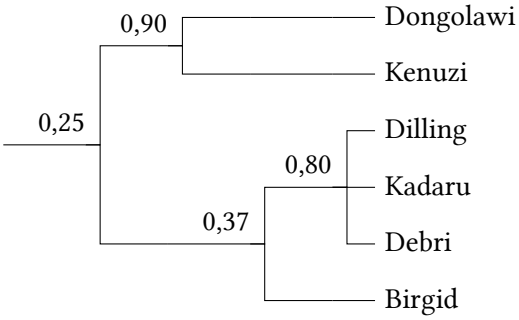
Teil I. Guaikurú-Sprachfamilie (Argentinien, Brasilien, Paraguay)

	Toba (Östlich)	Pilagá	Mokoví (Chaco)	Kadiwéu
Wolke	l=ʔok$_1$	ˈlo=ʔok$_1$	naweɣelek$_2$	lol:adi$_3$
Feuer	nodek$_1$	ˈd=oleʔ$_2$	noɾek$_1$	n=ol:edi$_2$
Fisch	njaq$_1$	ˈnijaq$_1$	naʌin$_2$	nij:oɢo-d͡ʒegi$_3$
Kopf	=qajk$_1$	=ˈqajk$_1$	=qaik$_1$	=ak:ilo$_2$
töten	=alawat$_1$	=aˈla:t$_1$	=alawat$_1$	=el:owadi$_1$
Mond	ʔawoʁojk$_1$	ʔaˈwoʃojk$_1$	ʃiɾajɣo$_2$	ep:enaj$_3$
Nase	=mik$_1$	=ˈmik$_1$	=mik$_1$	=m:iq:o$_1$
Salz	towe$_1$	olˈɣek$_2$	ʔwe$_1$	jok:i$_{-1}$
Stein	qaʔ$_1$	ˈqaʔ$_1$	qaʔ$_1$	wet:iɢa$_2$
Zunge	=at͡ʃ-aʁat$_1$	=aˈt͡ʃ-aʃat$_1$	=oʔleɣ-aʁan-aʁat$_2$	=ok:el:i$_3$

	Algorithmus A	Algorithmus B	
manuell			Stabilitätsindizes: Wolke 0,50 Feuer 0,50 Fisch 0,50 Kopf 0,75 töten 1,00 Mond 0,50 Nase 1,00 Salz 0,67 Stein 0,75 Zunge 0,50
automatisiert			Stabilitätsindizes: Wolke 0,50 Feuer 0,50 Fisch 0,75 Kopf 0,75 töten 1,00 Mond 0,50 Nase 1,00 Salz 0,25 Stein 0,75 Zunge 0,50

Teil II. Nubische Sprachfamilie (Ägypten, Sudan)

	Dongolawi	Kenuzi	Dilling	Kadaru	Debri	Birgid
töten	ˈbɛː$_1$	beː$_1$	hur$_2$	wur-i$_2$	wur-i$_2$	filaːl-e$_1$
Mond	uˈn-at-t$_1$	an-at-ti$_1$	nɔn-ti$_1$	nɔn-tu$_1$	nɔn-to$_1$	maːl$_2$
Wasser	ˈɛss$_1$	essi$_1$	ɔti$_1$	ɔto$_1$	ɔtu$_1$	eɟi$_1$
geben	ˈtɪr$_1$	tir$_1$	ti$_1$	ti$_1$	ti$_1$	teː-n$_1$
gut	ˈsɛrɛː$_1$	sereː$_1$	ken$_2$	kɛn$_2$	kɛŋ$_2$	azze-n$_3$
Wind	ˈturug$_1$	turug$_1$	irʃ-i$_2$	irʃ-o$_2$	irʃ-o$_2$	kurr-i$_3$
Haare	ˈdɪl-tɪ$_1$	siːr$_2$	tel-ti$_1$	til-tu$_1$	til-tu$_1$	ur=dill-e$_1$
Bauch	ˈtuː$_1$	tuː$_1$	te-te$_2$	to$_1$	to$_1$	tuː$_1$
schlafen	ˈnɛːr$_1$	neːr$_1$	ɟer$_1$	dwallɛli$_2$	ɟer-i$_1$	neːr-i$_1$
Sonne	ˈmasɪl$_1$	masɪl$_1$	ɛɟ$_2$	aɟu$_2$	eŋgal-to$_3$	ʔiːzi$_2$

	Algorithmus A	Algorithmus B	
manuell			Stabilitätsindizes: töten 0,50 Mond 0,83 Wasser 1,00 geben 1,00 gut 0,50 Wind 0,50 Haare 0,83 Bauch 0,83 schlafen 0,83 Sonne 0,50
automatisiert			Stabilitätsindizes: töten 0,33 Mond 0,50 Wasser 0,50 geben 0,67 gut 0,50 Wind 0,50 Haare 0,83 Bauch 1,00 schlafen 0,50 Sonne 0,50

(A) **(2 Punkte)** Der Konsonant **ʁ** wird wie *r* im Wort *rot* ausgesprochen, am hinteren Teil der Zunge. Zu welcher Dolgopolsky-Klasse gehört dieser Konsonant, und wie habt ihr das festgestellt?

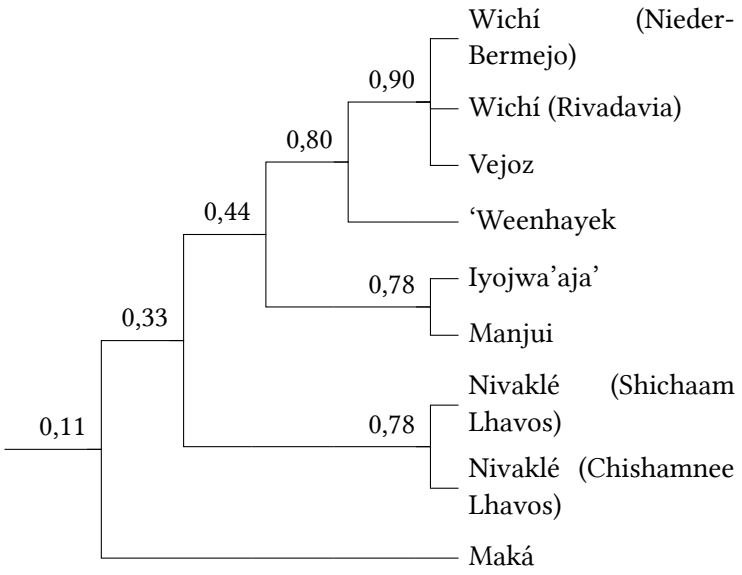
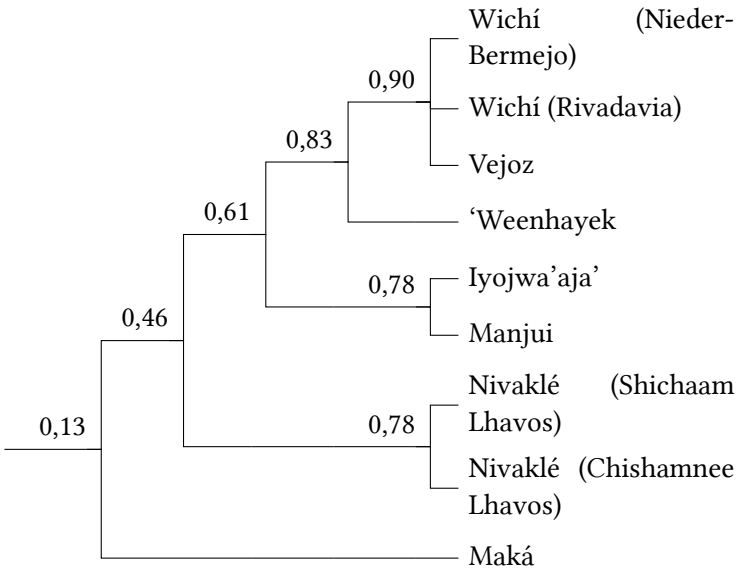
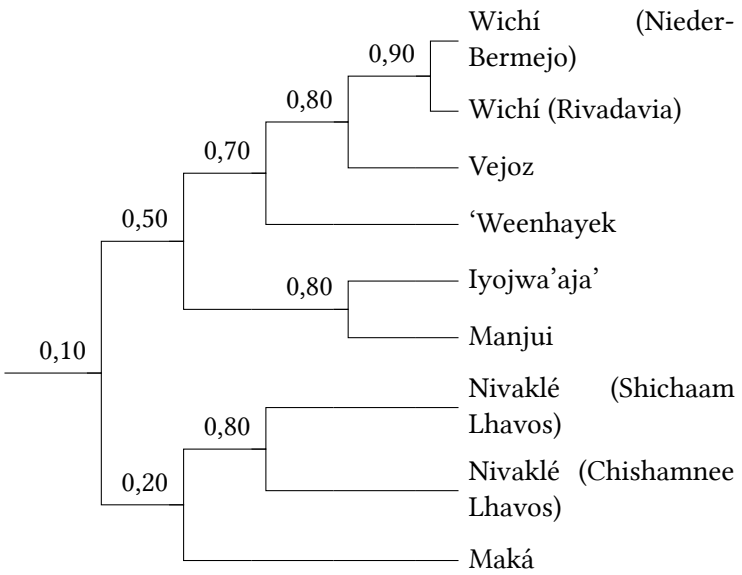
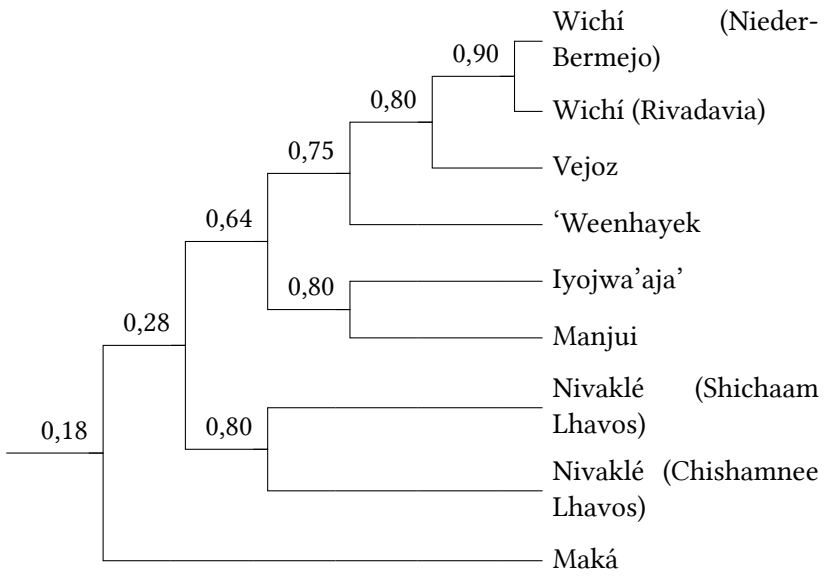
(B) **(2 Punkte)** Der nubische Baum oben links ist nur einer von zwei möglichen Bäumen für diese Kombination von Algorithmus und Annotationstyp. Zeichnet den anderen möglichen Baum.

(C) **(2 Punkte)** Der nubische Baum unten links ist nur einer von zwei möglichen Bäumen für diese Kombination von Algorithmus und Annotationstyp. Zeichnet den anderen möglichen Baum.

(D) **(2 Punkte)** Wie andere Distanzen in dieser Aufgabe wurde die (der Wurzel des nubischen Baumes oben rechts zugewiesene) lexikostatistische Distanz 0,49 auf zwei Dezimalstellen gerundet. Was ist die exakte Distanz?

Teil III. Mataguayische Sprachfamilie (Argentinien, Bolivien, Paraguay)

	Wichí (Nieder-Bermejo)	Wichí (Rivada-via)	Vejoz	‘Weenhayek	Iyojwa’aja’	Manjui	Nivaklé (Shichaam Lhavos)	Nivaklé (Chis-hamnee Lhavos)	Maká
Feuer	ʔitoχ$_{1}$	ʔitɔx$_{1}$	ʔitɑh$_{1}$	ʔiːtɑχ$_{1}$	ˈh͡wat$_{2}$	ˈʔe͡itʲe$_{1}$	ʔitɑx$_{1}$	ʔitɑx$_{1}$	feʔt$_{2}$
Fisch	ʔwahat$_{1}$	wahat$_{1}$	wahat$_{1}$	ʔwaːhat$_{1}$	siˈʔjus$_{-1}$	ʃiˈʔjus$_{-1}$	saxet͡ʃ$_{-1}$	saxet͡ʃ$_{-1}$	sehet͡s$_{-1}$
Fuß	=pat͡ʃu$_{1}$	=qɔlɔ$_{2}$	=pat͡ʃo$_{1}$, =kɑla$_{2}$	=paːkʲʼoʔ$_{1}$	=ˈsat$_{3}$	=kaˈʔlaʔ$_{2}$	=ɸoʔ$_{4}$	=ɸoʔ$_{4}$	=fʼiʔ$_{5}$
Wasser	ʔinot$_{1}$	ʔinɔt$_{1}$	wah$_{2}$	ʔinɑːt$_{1}$	ʔiˈnʲat$_{1}$	ʔaˈʔnat$_{1}$	jinɑʔt$_{1}$	jinɑʔt$_{1}$	iweliʔ$_{3}$
geben	=ʔweˌn-u$_{1}$	=weˌn-u$_{1}$	=ʔweˌn-o$_{1}$	=ʔweˌn-oʔ$_{1}$	=ˈwɛh͡n-aʔm$_{2}$	=ˈhajʔ$_{3}$, =ˈwɛn$_{2}$	=xut$_{4}$	=xut-ej$_{4}$	tis-ix$_{5}$
gut	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ʔis$_{1}$	ˈʔes$_{1}$	ˈʔe͡is$_{1}$	ʔis$_{1}$	ʔis$_{1}$	t=ejkʼun-ej$_{2}$
Wind	ʔinwokʷ$_{1}$	ʔinwɔk$_{1}$	ʔihwokʷ$_{1}$	=jaːɬ$_{2}$, =xʷoxʷ$_{3}$	ˈh͡lah͡wuʔ$_{4}$	ˈh͡lah͡wuuʔ$_{4}$	ɬaβiʔm$_{5}$	ɬaβiʔm$_{5}$	tʼunikʼi$_{6}$
Baum	haʔlo$_{1}$	halɔ$_{1}$	haʔlɑ$_{1}$	haʔlɑʔ$_{1}$	ʔaˈʔlaʔ$_{1}$	ʔaˈʔla-k$_{1}$	ʔaʔkxi-juk$_{2}$	jiʔk͡lɑʔ$_{1}$	naxka-k$_{3}$
Haare	=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$_{5}$
töten	=lon$_{1}$	=lɔn$_{1}$	=lɑn$_{1}$	=lɑːn̥$_{1}$	=ˈlaʔan$_{1}$	=ˈlan$_{1}$	=k͡lɑn$_{1}$	=k͡lɑn$_{1}$	=lan$_{1}$

<table>
<tr><th></th><th>Algorithmus A</th><th>Algorithmus B</th><th></th></tr>
<tr><td>manuell</td><td> 
</td><td> 
</td><td>Stabilitätsindizes:
Feuer 0,78
Fisch 1,00
Fuß 0,33
Wasser 0,78
geben 0,44
gut 0,89
Wind 0,33
Baum 0,78
Haare 0,67
töten 1,00</td></tr>
<tr><td>automatisiert</td><td> 
</td><td> 
</td><td>Stabilitätsindizes:
Feuer 0,78
Fisch 0,44
Fuß 0,33
Wasser 0,56
geben 0,67
gut 0,89
Wind 0,22
Baum 0,67
Haare 0,67
töten 1,00</td></tr>
</table>

Teil IV. Mongolische Sprachfamilie (Volksrepublik China, Mongolei, Russland)

(E) (10 Punkte) Analysiert die folgende Wortliste. Berechnet die Stabilitätsindizes, die den manuellen und den automatisierten Annotationen entsprechen.

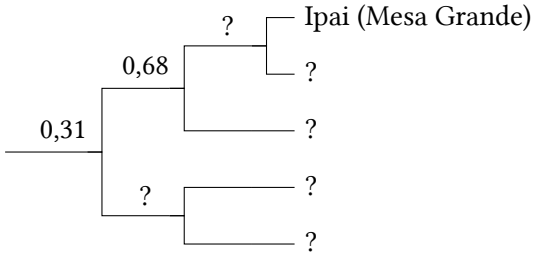
Als Hilfestellung wurden beide Stabilitätsindizes für das Wort ‚alle' schon berechnet. In zufälliger Reihenfolge sind diese Indizes 0,36 und 0,40.

	Dagurisch (Hailar)	Chamnigan (Mandschu)	Burjatisch (Chori)	Neu-Bargutisch	Ööld	Choschut	Kalmückisch	Chalcha	Ordos	Ost-Yugur	Bonan
alle	hɔː$_{1}$	bʊlt$_{2}$	bʊxiː$_{3}$	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$_{-2}$
Rinde	hai̯l̥s$_{1}$	qalihʊn$_{1}$	χɔltəhɔn$_{2}$	xaʲhuː$_{1}$	xolts$_{2}$	xalis$_{1}$	dursn$_{3}$	xɔɮt^{h}ɔ̆s$_{2}$	turusu$_{3}$	χalsən$_{1}$	arasun$_{4}$
Bauch	keːli$_{1}$	gətəhʉn$_{2}$	gedehen$_{2}$	gedyː$_{2}$	geːs$_{2}$	gets$_{2}$	gesn$_{2}$	gitĭs$_{2}$, xiwɮʒĭj$_{-1}$	ketysy$_{2}$	ketesən$_{2}$	kele$_{1}$
Vogel	dəgi$_{-1}$	ɕiwan$_{1}$	ʃubuːn$_{1}$	ʃuwuː$_{1}$	ʃuvuː$_{1}$	ʃuwuː$_{1}$	ʃowun$_{1}$	ʃuwu$_{1}$	ʃuβuː$_{1}$	ʃuːn$_{1}$, pelt͡ʃɔr$_{2}$	bend͡ʑer$_{2}$
Feuer	ɢali$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	ɢal$_{1}$	qal$_{1}$	qal$_{1}$	χal$_{1}$
Weg	terg-uːl$_{1}$	qarɢʊ̯i$_{2}$	χarɢɨ$_{2}$, zam$_{-1}$	zam$_{-1}$	d͡zam$_{-1}$	d͡zam$_{-1}$	xaː-lʁə$_{3}$	t͡sam$_{-1}$	t͡ʃam$_{-1}$	mør$_{4}$	mor$_{4}$
Salz	hataː$_{1}$	dawhʊn$_{2}$	dabhan$_{2}$	dawuhuː$_{2}$	daws$_{2}$	daws$_{2}$	dawsn$_{2}$	tawsă$_{2}$	taβusu$_{2}$	taːpsən$_{2}$	dabsuŋ$_{2}$
schwimmen	unpa-du$_{1}$	ʊmba-$_{1}$	t^{h}amar-$_{2}$	umb-$_{1}$	sele-$_{3}$	umba-$_{1}$	us-t͡ɕi-$_{4}$, øːm-$_{5}$	siɮʒĭ-$_{3}$	usu-t͡ʃ^{h}i-la-$_{4}$	umpa-$_{1}$	mba-$_{1}$
Wasser	ɔsɔ$_{1}$	ʊxʊn$_{1}$	uhan$_{1}$	uːha$_{1}$	usn$_{1}$	us$_{1}$	usn$_{1}$	ʊsŏ$_{1}$	usun$_{1}$	q^{h}usun$_{1}$	sə$_{1}$
Wind	kei̯n$_{1}$	halkin$_{2}$	halxin$_{2}$	halxi$_{2}$	salʲxin$_{2}$	salkʲi$_{2}$	salʲkn$_{2}$	saɮʒxĭ$_{2}$	k^{h}iː$_{1}$	k^{h}iː$_{1}$	ki$_{1}$

Teil V. Yuma-Sprachfamilie (Mexiko, USA)

(F) **(8 Punkte)** Analysiert die folgende Wortliste. Unten ist ein Baum, der mit derselben Liste gebaut wurde. Einige Daten (Namen von Sprachen und lexikostatistische Distanzen) fehlen. Füllt die Lücken aus. Gebt an, ob der Baum manuell oder automatisch generiert ist, sowie ob er von Algorithmus A oder B generiert ist.

	Mohave	Kokopa	Yavapai	Tipai (Jamul)	Ipai (Mesa Grande)
kurz	wena=wen-a_1	ˈxʌ=ʔut$_2$	ˈt͡ʃkr=ot-i_2	ləˈʔuɲ$_1$	mə=put-k_3
Vogel	ʔit͡ʃ=i=jer$_1$	ˈʃa$_2$	ˈʔ=t͡ʃ=sa$_2$	aʔˈʃa$_2$	ʔaː=ʂaː$_2$
Knochen	ɲ=a=s=ak$_1$	ˈɲ=j=aːk$_1$	ˈt͡ʃ=j=aːk-a_1	ˈak$_1$	aq$_1$
trocken	i=roː-v-k_1	ˈʂ=ʔar$_2$	ˈru-β-i_1	ˈs=ʔaːj$_3$	saːj$_3$
Fleisch	k^wiːk^way$_1$	ʔi=ˈmaːt͡ʃ$_2$	ˌk^weːˈθo-β-a_3	ˈk^wak$_4$	kuk^{w}aːj-p_1
Hals	maʌaqe$_1$	ˈm=puk$_2$	ˈmlqíˌ$_1$	iːˈpuk$_2$	iː=puk$_2$
sehen	i=juː-k_1	ˈwiː$_2$	ˈʔuː$_1$	ˈwiːw$_2$	ə=wuːw$_2$
Schwanz	iː=ʔar$_1$	ˈʃ=juʌ̊$_2$	ˈβ=hé$_3$	ʃəˈjuɬ$_2$	xə=juɬ$_2$
zwei	havik-k_1	ˈx=wak$_1$	ˈh͡wâk-i_1	xəˈwak$_1$	xə=wak$_1$
Jahr	huːðe$_1$	ˌmat-ˈkaːm$_2$	ˈʔ=t͡ʃʰur-a_3	matˈwam$_2$	ʔaːnʲ$_{-1}$



(G) **(20 Punkte)** Einige andere Bäume wurden für die Yuman-Sprachen generiert und zeigen die folgenden lexikostatistischen Distanzen an der Wurzel jedes Baumes (ganz links an jedem Baum):

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

Zeichnet jeden dieser Bäume. Gebt für jeden der Bäume an, ob er manuell oder automatisch generiert ist, sowie ob er von Algorithmus A oder B generiert ist.

(H) **(3 Punkte)** Zwei Distanzen in Teilaufgabe **(G)** wurden auf zwei Dezimalstellen gerundet: 0,23 wurde von 0,225 hochgerundet. Welche andere Distanz wurde gerundet, und wie hoch ist ihr präziser Wert?

(I) **(4 Punkte)** Erklärt, wie die Stabilitätsindizes berechnet werden.

(J) **(5 Punkte)** Erklärt, wie die lexikostatistische Distanzen berechnet werden.

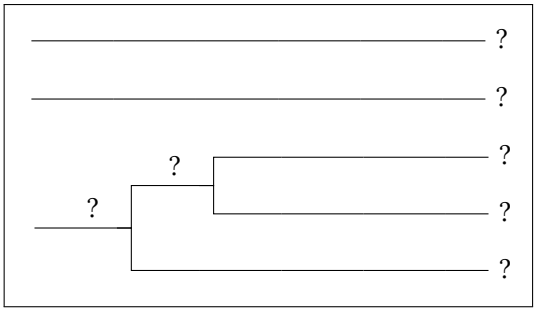
(K) **(4 Punkte)** Erklärt den Unterschied zwischen den Algorithmen A und B.

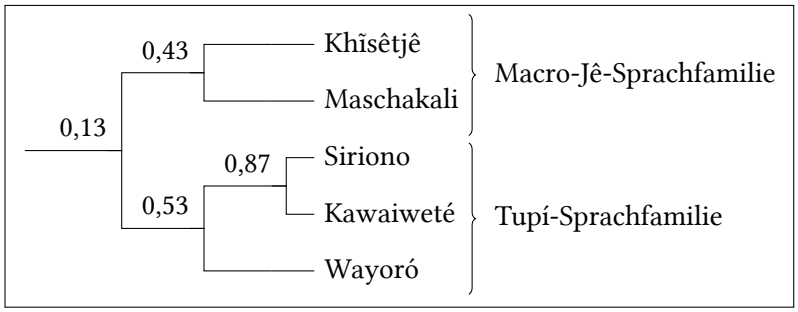
Teil VI. Macro-Jê-Sprachfamilie und Tupí-Sprachfamilie (Brasilien, Bolivien)

(L) (28 Punkte) Die Macro-Jê-Sprachen und Tupí-Sprachen sind zwei bedeutende Sprachfamilien Südamerikas. Einige Linguisten denken, dass sie entfernt verwandt sind. Analysiert die folgenden Wortlisten.

	A	B	Γ	Δ	E
Rinde	e=ˈe-ke	hʷĩ=ˈkʰʌ	kʉp=ˈpɛ	mĩβ͡m=ˈt͡ɕaj	=ˈpɛ
Bauch	ˈe=ɾje	=ˈtʰigi	=ãˈʉ̃n	=ˈtæj	=ɾɛˈwɛk
Blut	e=ˈɾuki	=kaˈⁿbɽo	=d͡ʒ=aˈʉ	=ˈhɛβ͡p	=ɾuˈɨ
verbrennen	=ˈɾai̯	=ɾɔ=ˈkʰɹə̃	=poˈkʷa	mũ=...=ˈhaβ͡p	=ɾaˈpɨ
Fett	e=ˈkiɾa	=ˈtʰwəmɨ	=ˈd͡ʒ=ap	=ˈtuβ͡p	=ˈkap
Fuß	ˈe=i	=ˈhʷaji	=ˈβi	=pɒˈta	=ˈpɨ
Hand	ˈe=o	=ɲɨ̃ˈkʰɹa	=ˈβo	=ˈɲĩβ͡m	=ˈpɔ
schwer	e=ˈusi	=wɨˈtʰĩ	=poˈti	=β͡pˈtəj	=pɔˈɨj
Leber	ˈe=ja	=ˈⁿba	=piˈa	=t͡ɕɨβ͡pkɨ̃ˈnɑ̃j	=pɨˈʔa
neu	e=ˈjasu	=ˈⁿdɨwɨ	=paˈgop	=ˈtɨβ͡p	=pɨaˈu
Wurzel	e=ˈɾao	=jaˈɽe	kʉp=kʉjoˈpɛ	mĩβ͡m=ɲĩβ͡m=t͡ɕaˈti̯ə	=ɾaˈpɔ
Haut	ˈe=ɨ	=ˈkʰʌ	=ˈpɛ	=ˈt͡ɕaj	=ˈpit
Schwanz	e=ˈɾoko͡ɨ	=ˈⁿbɨ	=d͡ʒ=oˈkʷaj	=nãː=ˈkɨβ͡p	=ˈɾaj
weiß	ˈe=ʃɨ̃	=jaˈkʰa	=d͡ʒiˈɾa	=β͡pˈdoɯ̨	=ˈsĩŋ
Flügel	e=ˈheo	=jaˈɽa	=pɛˈo	=ɲɨ̃ˈmãɯ̨	=pɛˈpɔ, =jɨˈwa

Unten sind zwei Bäume, die mit denselben Listen gebaut wurden. Einige Daten (Namen von Sprachen und lexikostatistische Distanzen) fehlen. Füllt die Lücken aus. Gebt für jeden der Bäume an, ob er manuell oder automatisch generiert ist, sowie ob er von Algorithmus A oder B generiert ist.





A	B	Γ	Δ	E
?	?	?	?	?

⚠ In dieser Teilaufgabe wurden die manuellen Annotierungen und Stabilitätsindizes absichtlich ausgelassen.

(M) (10 Punkte) Automatisierte Verfahren, die auf Dolgopolsky-Klassen basieren, können ungenaue Ergebnisse erzielen. In diesem Beispiel identifiziert das automatische Verfahren mehr Ähnlichkeiten zwischen dem Siriono und einer bestimmten Macro-Jê-Sprache (Khĩsêtjê) als zwischen dem Siriono und anderen Tupí-Sprachen. Schlagt ein geändertes automatisiertes Verfahren vor, das eine richtige Klassifikation erzielen würde, wenn es auf die Macro-Jê- und Tupí-Wortlisten oben angewendet würde, und beschreibt es *kurz*.

⚠ Diese Teilaufgabe wird nur bei Punktgleichheit zwischen den besten Teams bewertet werden.

Die Autoren danken Alejandra Vidal, Maria Konoshenko, Ilya Gruntov und Jamthô Suyá für die Beantwortung ihrer Fragen zu einzelnen Sprachen. —*Andrey Nikulin, Milena Veneva*

Viel Erfolg!