

Berkeley Slavic Conference, February 2010

Alan J. Redd (Anthropology), Marc L. Greenberg (Slavic)

University of Kansas

1. Classification – G & A claim statistical improvements on analyzing lexicostatistics based on PIE and daughter languages. – better absolute chronologies. (a) How do lexicostat analyses compare with phono/morph analysis? Most would say lexicostat and phonology different
2. How well does it work for Slavic?
3. Look at data: (a) Dyen + G&A recognize tree structures for Slavic are not well supported. (b) Therefore Dyen claims 2-dimensional pseudomaps may improve situation.
4. Redd + Green: (a) quantify similarities or differences b/w different sets of data (Dyen vs. Maniczak); (b) quantify similarities or differences b/w lexical vs phono/morphological; and (c) to quantify the correlation between geography and the lexical and phon/morphological data sets.

Family tree and or map-like approaches to Slavic languages?

Abstract

Lexicostatistics is decades old, but newer techniques for computational approaches to historical linguistics have gained attention with the rise of more sophisticated methods of data handling. Thus, for example, Gray and Atkinson (2003, Figure 1) claim to have established, using cognates and a Bayesian tree analysis, an authoritative *Stammbaum* for the Indo-European (IE) language family, including absolute chronologies of its branching.

The present paper examines a smaller subset of IE languages—Slavic—using Bayesian methods and map-like methods in attempt to compare the computational results and model assumptions with received analyses that are closer to the present. We assume that examining a group of languages closer in time to the present, where the splits are more easily verifiable, allows a more fine-grained comparison of different analysis methods. If a close fit can be found between Bayesian trees and maps and traditional analysis in Slavic, it should allow extension to greater time depths and larger families such as Indo-European.

The present paper applies Bayesian trees and map methods to two corpora: the Slavic subset of Indo-European in Gray and Atkinson (2003); and the Slavic text-token set in Mańczak (2004).

•

Gray and Atkinson 2003 have claimed that new models of analysis may be applied to glottochronology that answer previous criticism of the method and overcome the shortcomings. The outcome of their glottochronological experiment demonstrated impressive results in establishing absolute chronologies for Indo-European which correlate with archaeological (Renfrew's out-of-Anatolia and Gimbutas' Kurgan expansion) and genetic evidence (Near-Eastern contribution to the IE gene-pool during the Neolithic) (438). This establishes a root of IE at 8700 BP (Hittite), with Tocharian splitting off at 7900, Greek and Armenian at 7300, Indo-Aryan at 6900, Celto-Germano-Romance at 6100, and Balto-Slavic at 3400.

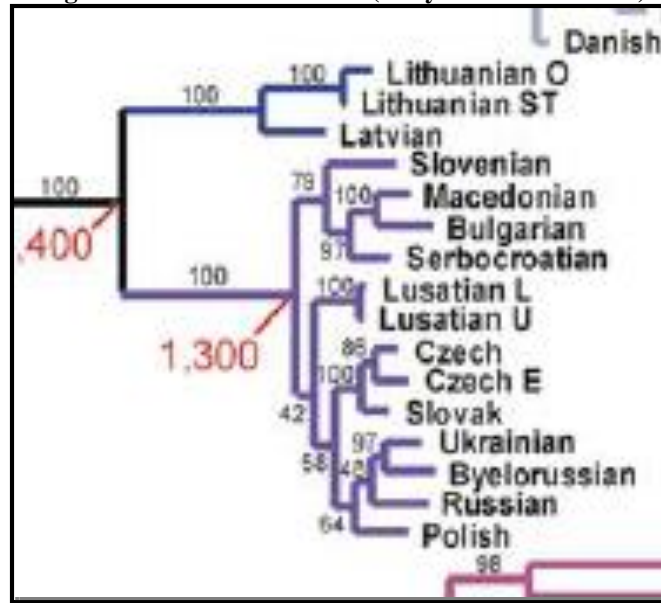
Slide 1: Slavic languages map and Gray & Atkinson Slavic results

Need Dyen et al Quote about inadequacy of family-tree model for Slavic & Celtic b/c of continued contact. This correlates with low posterior probabilities in Slavic splits vs. higher posterior probabilities in other branches. However, G & A find that Slavic has the lowest PP

whereas Celtic and other branches have high PPs among the well-accepted daughter families. (There are other weak points at deeper time depths, e.g., Indo-Iranian + Albanian.)

In G & A Slavic is rooted at 1300 BP, assuming a date of 700 AD for a terminus post-quem for the dissolution of Proto-Slavic, thus roughly corresponding to the traditional date of 500 AD for the beginning of Slavic migrations from Ukraine. Both the low PP & apparent incorrect clustering of Polish with ESl mean that the tree model does not allow absolute dating for Slavic splits. As Dyen suggests, Slavic requires the use of 2-dimensional maps.

Figure 1: Balto-Slavic Detail (Gray & Atkinson 2003)



SLIDE: SCAN OF DYEN's MDS plot

Dyen et al had run the data but claimed that because of contact after the languages had split, Slavic is better represented as a “psuedomap” (add in page).



FIGURE 2. Pseudomap of Slavic Languages.

SLIDE: REDD plot of Dyen

Dyen's data, which is also used by G & A, is a Swadesh-style list (200 semantics items for all IE) with 2449 realizations in form (i.e., tokens possible to match) among 84? languages. Dyen's distance matrix is the lexicostatistical percentage of shared cognates.

There is some support for classical groups: E, W, S. Polish again approaches East. Slovene is an outlier. Find commentary in Dyen why they think this is the case.

Mańczak 2004 – distances expressed as raw N of correspondences between pairs

To look at another sample of lexical correspondence Slavic data we looked at Mańczak 2004, which is not a Swadesh list. Rather, it is a set of correspondences in parallel translations of a Gospel text. Each match between pairs is registered for each time that same form (root, where applicable) is used for the same meaning, thus, POL $w =$ UKR v , but POL $w \neq$ UKR do .

Mańczak expressed these as raw numbers of correspondences between pairs with 1816 total realizations.

total tokens

101

Pod względem statystycznym podobieństwa leksykalne między językami słowiańskimi przedstawiają się następująco:

Polski	Czeski	Słowacki	Górno- łużycki	Dolno- łużycki	Rosyjski	Białoruski	Ukraiński	Słoweński	Serbo- chorywacki	Bułgarski
Cz 1158	Sla 1345	Cz 1345	Dl 1155	Gł 1155	Br 1103	Ukr 1160	Br 1160	Sch 1041	Sle 1041	Sch 1007
Sla 1118	Pol 1158	Pol 1118	Sla 1011	Cz 970	Sch 986	Ros 1103	Pol 900	Pol 956	Bg 1007	Ros 887
Sle 956	Gł 1002	Gł 1011	Cz 1002	Sla 950	Pol 902	Pol 951	Ros 900	Ros 894	Ros 986	Sle 860
Br 951	Dl 970	Dl 950	Pol 835	Pol 786	Ukr 900	Sch 836	Sla 731	Bg 860	Pol 887	Br 813
Ros 902	Sle 835	Sle 799	Sle 815	Sle 763	Sle 894	Bg 813	Sch 719	Cz 835	Br 836	Pol 735
Ukr 900	Ros 738	Ukr 731	Ukr 629	Br 599	Bg 887	Sle 777	Cz 698	Gł 815	Sla 724	Ukr 687
Sch 887	Sch 717	Sch 724	Sch 612	Sch 597	Cz 738	Sla 707	Sle 694	Sla 799	Ukr 719	Sla 589
Gł 835	Ukr 698	Ros 721	Br 612	Ros 596	Sla 721	Cz 693	Bg 687	Br 777	Cz 717	Cz 572
Dl 786	Br 693	Br 707	Ros 601	Ukr 592	Gł 601	Gł 612	Gł 629	Dl 763	Gł 612	Gł 530
Bg 735	Bg 572	Bg 589	Bg 530	Bg 506	Dl 596	Dl 599	Dl 592	Ukr 694	Dl 597	Dl 506
9228	8728	8695	7802	7514	8328	8251	7710	8434	8126	7186
avg: 50.81	48.06	47.88	42.96	41.38	45.86	45.44	42.46	46.44	44.75	39.57

Slide: MDS-ML plot 11 Slavic languages (Mańczak's data)

We converted Mańczak's raw numbers to a distance matrix and created an MDS plot. We found a better fit for the traditional three groups than Dyen et al. had found. The groups could be oriented geographically, as shown, but while the branches were oriented correctly, their situation within the geography was less straightforward. Slovene was no longer an outlier. Polish was found to be near equidistant from all branches.

Slide 3: MDS-ML plot 11 Slavic languages (Dyen 1992)

In order to compare w/ Manczak's data we threw out Macedonian and E-Cz. It still supports clustering and doesn't significantly change the big picture. Also puts Polish to ESL and closest to Ukrainian.

Alan: what is the difference between the Dyen slides you made that are currently in positions 6 and 9 in the slide order?

Slide 4: MDS-ML plot 11 Slavic languages; 315 cognates Atkinson-Gray Jaccard distance

A & G shared their data set with us (thanks) and Redd converted the 1's and 0's to a distance matrix using the Jaccard similarity coefficient {EXPLANATION TO FOLLOW}. This distance matrix was used as input for an MDS plot (using maximum likelihood).

This moves Slovene closer to South Slavic (in contrast to its outlier status in the Dyen MDS). And W Slavic has moved from the center to a more westerly orientation. I.e., closer fit to geography. Polish is again intermediate b/w W & E, but now closer to Russian rather than Ukrainian.

Mańczak data showing differences in lexical matching.

POL tended to match RUS more often in this corpus than POL matched UKR and BEL (yellow highlights), though this was not always the case.

88						85					
Polski	Czeski	Slowacki	Górnoluzycy	Dolnoluzycy	Białoruski	Rosyjski	Ukraiński	Serbochorwacki	Słoweński	Bulgarski	
1 uradować się A	zradovať sa A	zradovať sa A	zwyżelic so B	wjasoly B	uzradavacca A	obradovať sja A	zraditi A	obradovati se A	razveseliti se B	zaradvam se A	
1 ustąpić A	posadzić se A	uložiť sa B	syndryć so A	sedmá A	sjedac' A	saditi sja A	sidiaty A	sjeti A	seniti A	sedna A	
1 ustąpić A	posadzić se A	uložiť sa B	zwyżelic so A	sedmá A	sjedac' A	wozić' D	posidaty A	posjedati A	posedeti A	sednja A	
1 ustąpić A	sedzić B	zajmúť B	sedzić A	sedzić A	sedzić A	zanjać B	zajmaly B	zajmaly B	sedeti A	zajemn B	
4 uszyścić A	uszyścić A	uszyścić A	zastyść A	stysć A	stysć A	uszyścić A	uszyścić A	uszyścić A	uszyścić A	uszyścić A	
1 uszupić A	nechat B	nechat B	dopasćić C	daśi D	pakimac' E	ostawić F	ostawić C	ostaviti F	ostaviti F	ustavja F	
1 ustąpić A	nechat B	nechat B	zwoićić D	plowić E	pakimac' E	ostawić F	dopomyty C	ostaviti F	ostaviti F	ustavja F	
1 uszedł A	obrócić se B	utiahnuť sa C	podat' so D	wustupić A	pałolj X	pośleć E	dopomak' C	dopomak' C	ostaviti F	ostavja F	
1 uwielbić A	odsielić X	odsielić X	čėknyć B	wustupić C	pajolj X	pośleć A	pośleć E	ostaviti F	ostaviti F	ostavja F	
1 uzdravić A	ostawić B	ostawić B	wostawić B	wucencić C	ustawić B	proslawić B	proslawiti B	proslawiti B	ostaviti B	ostavja F	
1 uzdravić A	uzdravować A	uzdravować A	wustrowić A	gój B	azdaraljac' A	wraczać C	udobrowoljaty A	udobrowoljaty A	iscjeliti B	iscjeliti B	
1 uzdrówić A	uzdraviti A	uzdraviti A	wustrowić A	wustrowić A	azdaraljac' A	iscjelić B	udobrowoljaty A	udobrowoljaty A	iscjeliti B	iscjeliti B	
1 użalić się A	milosrednistv B	milosrednistv B	wustrowić A	wustrowić A	azdaraljac' A	iscjelić B	odulaty C	odulaty C	iscjeliti B	iscjeliti B	
8 w A	lutować D	lutować D	mlonóć B	lutować D	zlitavacca D	szalić sja A	żal' A	sažaliti se A	sažaliti se A	smilija se B	
3 w A	do B	do B	do B	do B	u A	w A	do B	u A	u A	v A	
1 w A	do B	do B	do B	do B	u A	w A	u A	u A	u A	v A	
1 w A	na B	do C	do C	do C	u A	w A	do B	u A	u A	v A	
1 w A	na B	do C	do C	do C	u A	w A	do C	u A	u A	v A	
1 w A	na B	do C	do C	do C	u A	w A	do C	u A	u A	v A	
2 w A	na B	do C	do C	do C	u A	w A	u A	u A	u A	v A	
1 w A	na B	do C	do C	do C	u A	w A	do C	u A	u A	v A	
2 w A	s B	s B	z B	z B	z B	s B	z B	s B	z B	s B	
2 w A	v A	do B	do B	do B	u A	w A	do B	u A	u A	v A	
2 w A	v A	do B	do B	do B	u A	w A	v A	u A	u A	v A	
1 w A	v A	do B	w A	w A	u A	v A	u A	u A	u A	v A	
1 w A	v A	na B	na B	na B	u A	v A	u A	u A	u A	v A	
1 w A	v A	na B	na B	na B	u A	v A	na B	na B	na B	na B	
1 w A	v A	w A	w A	w A	u A	v A	u A	u A	u A	v A	
1 w A	v A	w A	w A	w A	u A	v A	u A	u A	u A	v A	
4 w A	v A	w A	w A	w A	u A	v A	u A	u A	u A	v A	
1 w A	v A	w A	w A	w A	u A	v A	u A	u A	u A	v A	
1 wchodzić A	choditi A	vojtić B	chodzić A	hys B	zachodzić A	vchoditi A	zachoditi A	ulaziti C	hoditi A	vlizam C	
1 wchodzić A	choditi A	vchoditi A	wrócić B	hys C	uvachodzić A	vchoditi A	zachoditi A	ulaziti C	hoditi A	vlizam C	
1 wędug A	podle X	podľa X	podle X	po B	vodle A	po B	uwojty A	uwojty A	po B	spored D	
1 wejść A	wejść A	vojtić A	dónić A	hys A	uwojty A	vojtić A	uwojty A	uwojty A	uwojty A	vlizam B	
3 wejść A	wejść A	vojtić A	zastąpić B	hys A	uwojty A	vojtić A	uwojty A	uwojty A	uwojty A	vlizam C	
1 weselić się A	svadba B	svadba B	hrodina C	zogniła B	vjadle A	brak D	veselija A	svadba B	svatočejna B	svadba B	
1 weselić się A	veseliti se A	hodoval' B	kwasaować C	wjasetlić se A	cefycca D	veselija A	veselitysja A	razveseliti se A	veseliti se A	razveselija se A	
1 weselić się A	vesely A	hodoval' B	kwasaować C	wjasetlić se A	cefycca D	veselitysja A	veselitysja A	veseliti se A	veseliti se A	veselija se A	
1 weselić się A	vesely A	hodoval' B	kwasaować C	wjasetlić se A	cefycca D	veselitysja A	veselitysja A	veseliti se A	veseliti se A	veselija se A	
1 wescelnić A	vzdelchnout' A	zastionat' B	ponjeselici so A	wjasetlić se A	pavesjalica A	prizvat' A	prizvat' A	prizvat' A	prizvat' A	vzdělchna A	
2 wezwąć A	povolat' B	povolat' B	powołat' B	wzwołat' B	prykliac' C	prizvat' A	prizvat' A	prizvat' A	prizvat' A	prizvat' A	
1 wezwąć A	povolat' B	povolat' B	wzwołat' B	wzwołat' B	zaklikac' D	prizvat' A	prizvat' A	prizvat' A	prizvat' A	prizvat' A	
1 wezwąć A	porzwać A	porzwać A	hosićić C	prysłój D	pratic' D	prizvat' A	prizvat' A	prizvat' A	prizvat' A	prizvat' A	
2 wezwąć A	porzwać A	porzwać A	prysłój C	prysłój C	zapracić' C	prizvat' A	prizvat' A	prizvat' A	prizvat' A	prizvat' A	
1 wezwąć A	porzwać A	porzwać A	prysłój B	prysłój B	zapracić' C	prizvat' A	prizvat' A	prizvat' A	prizvat' A	prizvat' A	
1 wezwąć A	porzwać A	porzwać A	prysłój B	prysłój B	zapracić' C	prizvat' A	prizvat' A	prizvat' A	prizvat' A	prizvat' A	
1 wezwąć A	porzwać A	porzwać A	prysłój B	prysłój B	zapracić' C	prizvat' A	prizvat' A	prizvat' A	prizvat' A	prizvat' A	

SLIDE: Birnbaum. Traditional schematic isogloss map for phonological isoglosses.

SLIDE: BIRNBAUM PHONOLOGY MDS PLOT

Converted into 0s (archaisms) and 1s (shared innovations), the MDS plot yielded a similar pseudomap to previous, though with three distinct branches. Again, Polish is an outlier with higher number of innovations distinct from others.

SLIDE: CORRELATION W GEOGRAPHY & 3 data sets

Shows best fit overall with geography with G & A data, least good with Dyen. Maniczak and Birnbaum were also close fits with geography.

Conclusions

References

Atkinson, Quentin D. 2009. Review of *Language Classification by Numbers*. By April McMahon and Robert McMahon. Oxford: Oxford University Press, 2005. Pp xvii, 265. *Diachonica* 26/1: 125–133.

Birnbaum, Henrik. 1966. The Dialects of Common Slavic. H. Birnbaum and Jaan Puhvel. *Ancient Indo-European Dialects*: 153—197. Berkeley and Los Angeles: Univ. of California Press.

Dyen, Isidore, Joseph B. Kruskal, and Paul Black. 1992. *An Indoeuropean Classification: A Lexicostatistical Experiment*. Philadelphia: American Philosophical Society.

Gray, Russell D. and Quentin D. Atkinson. 2003. Language-Tree Divergence Times Support the Anatolian Theory of Indo-European Origin. *Nature* 426: 435–439.

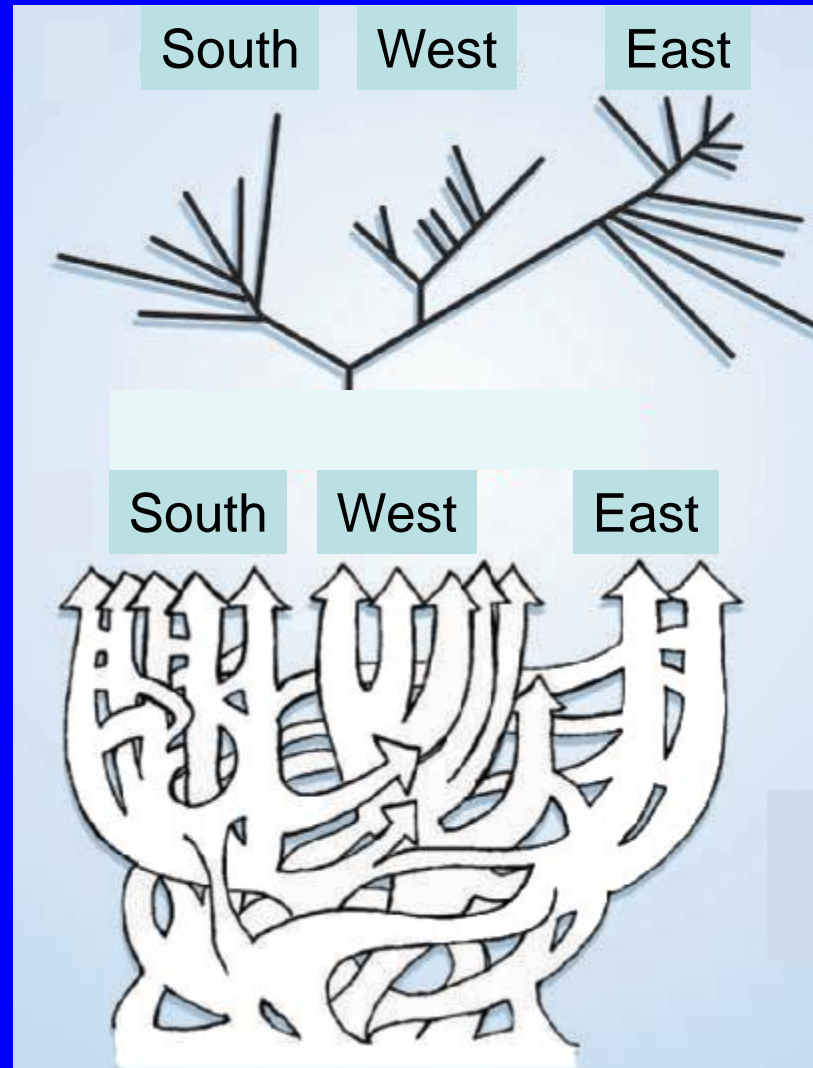
Mańczak, Witold. 2004. *Przedhistoryczne migracje słowian i pochodzenie języka staro-cerkiewno-słowiańskiego*. Cracow: PAU.

Family tree and or map-like approaches to Slavic languages?

Alan J. Redd (Anthropology) &
Marc L. Greenberg (Slavic)
University of Kansas

“Slavic Languages: Time and Contingency”, UC
Berkeley 12–13 Feb. 2010

Slavic language evolution: tree model or exchange model?

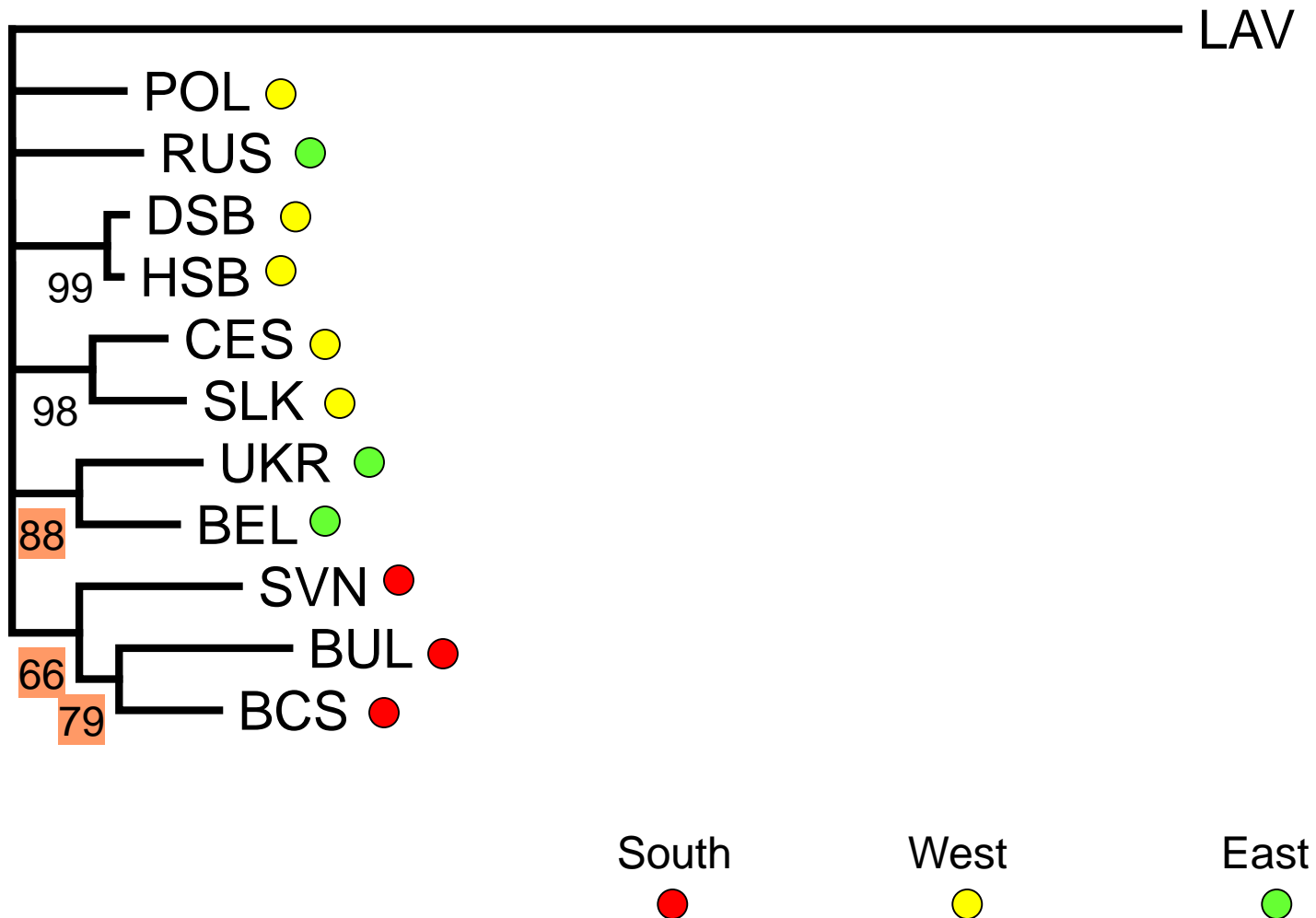


Slavic language map: West, South, and East.



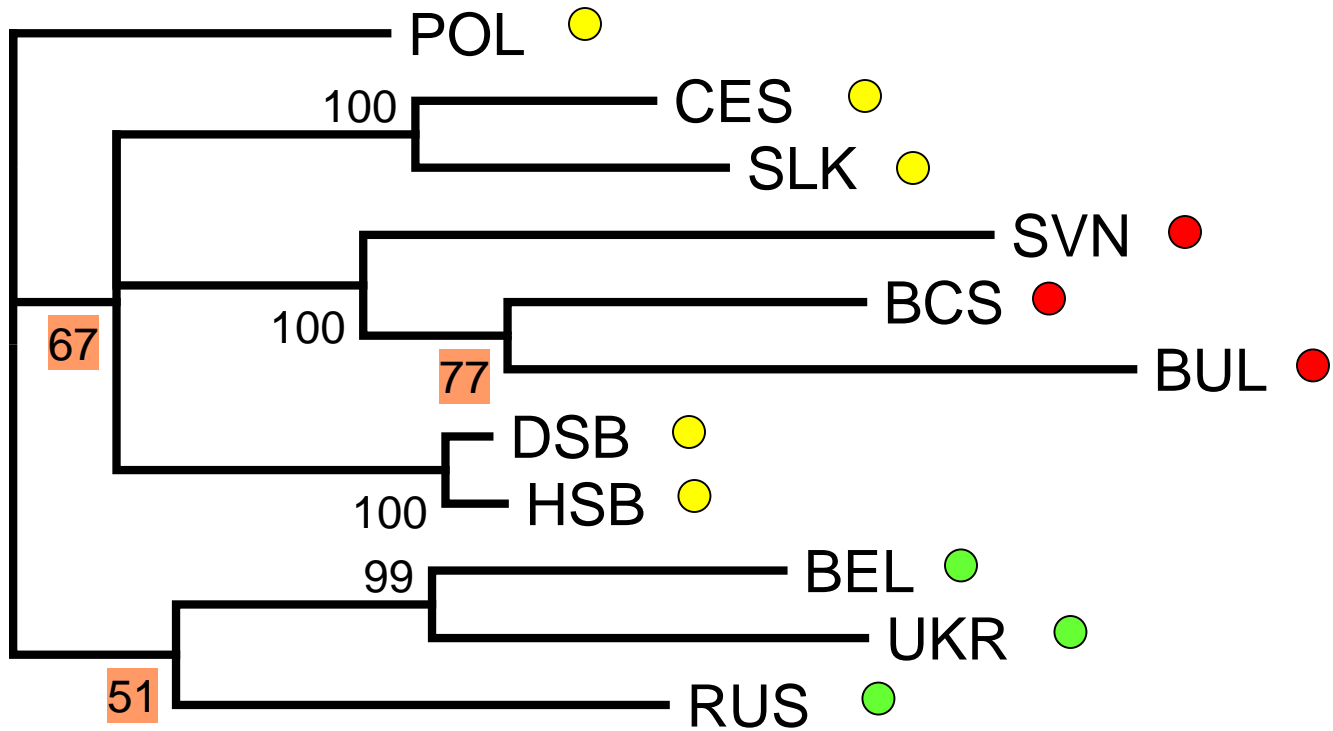
Tree model: Bayesian analysis

418 lexical items, 12 languages



Tree model: Bayesian analysis

314 lexical items, 11 languages



South



West

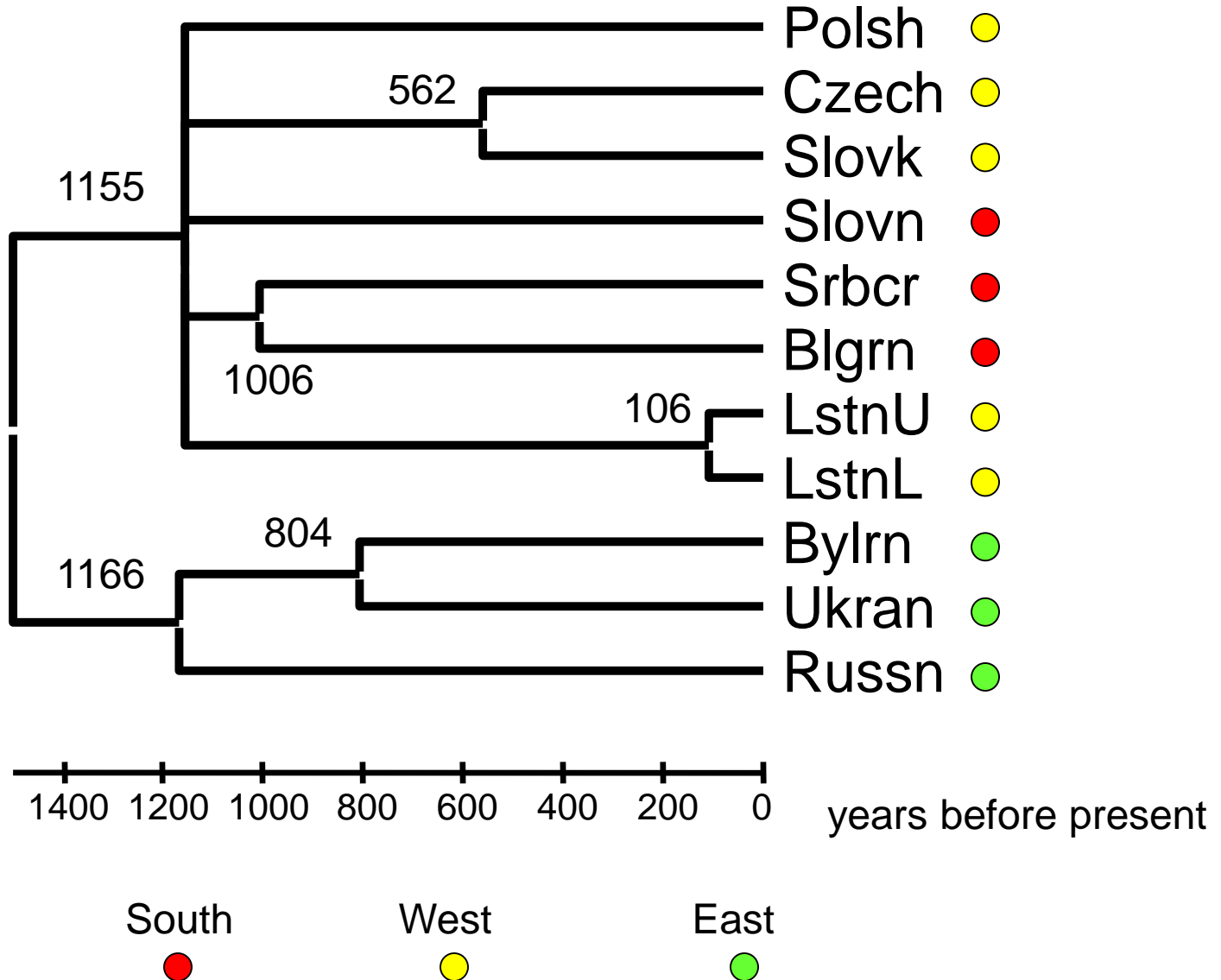


East



Tree model: Bayesian analysis

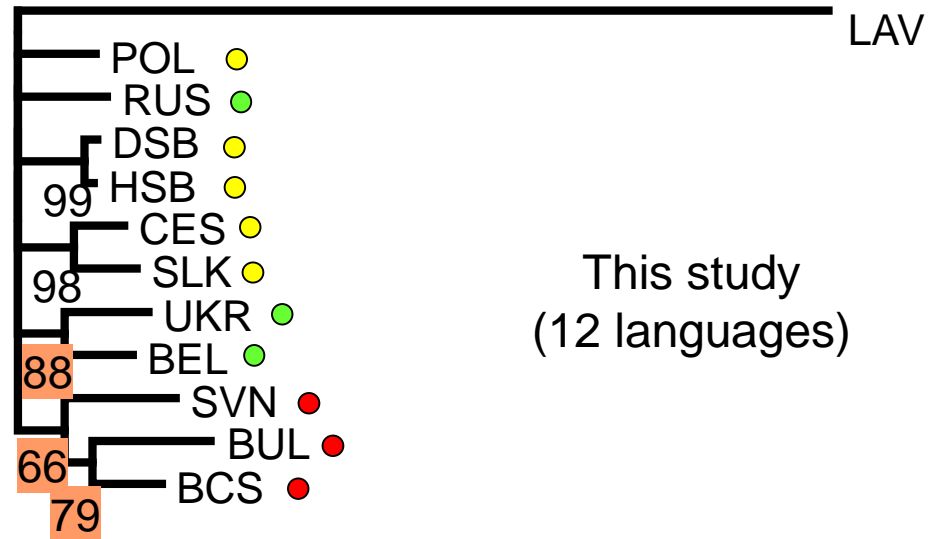
314 lexical items, 11 languages; linearized tree



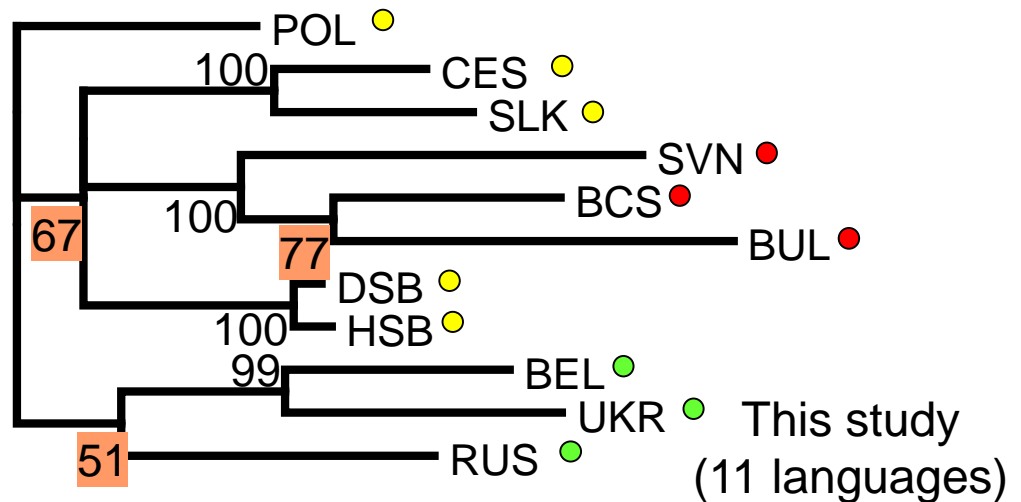
Summary slide of Tree model: Bayesian analysis; lexical items



G&A-2003
(87 languages)



This study
(12 languages)



This study
(11 languages)

MDS plot: Figure 2 Dyen, Kruskal & Black (1992)

200 cognates; 13 languages; % of shared cognates for Swadesh list

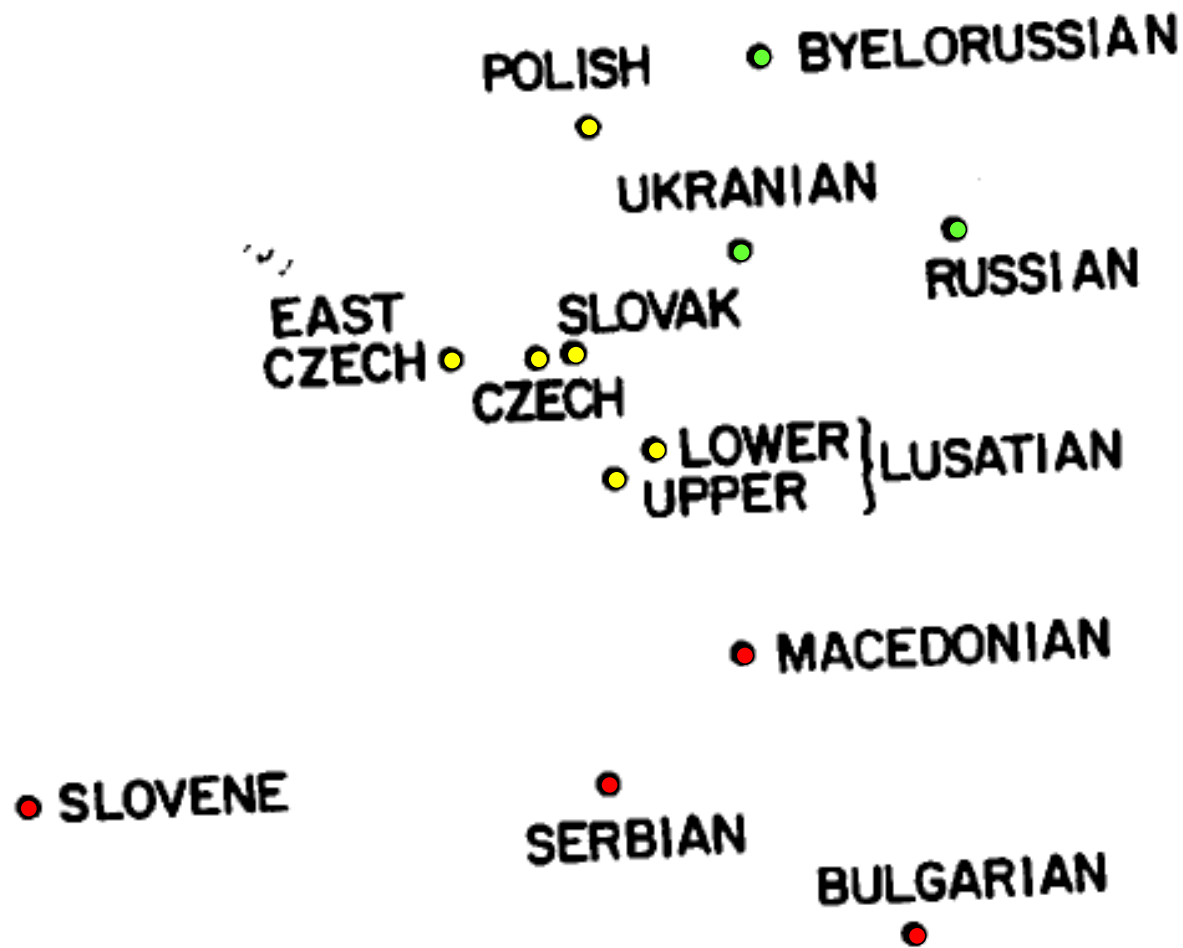
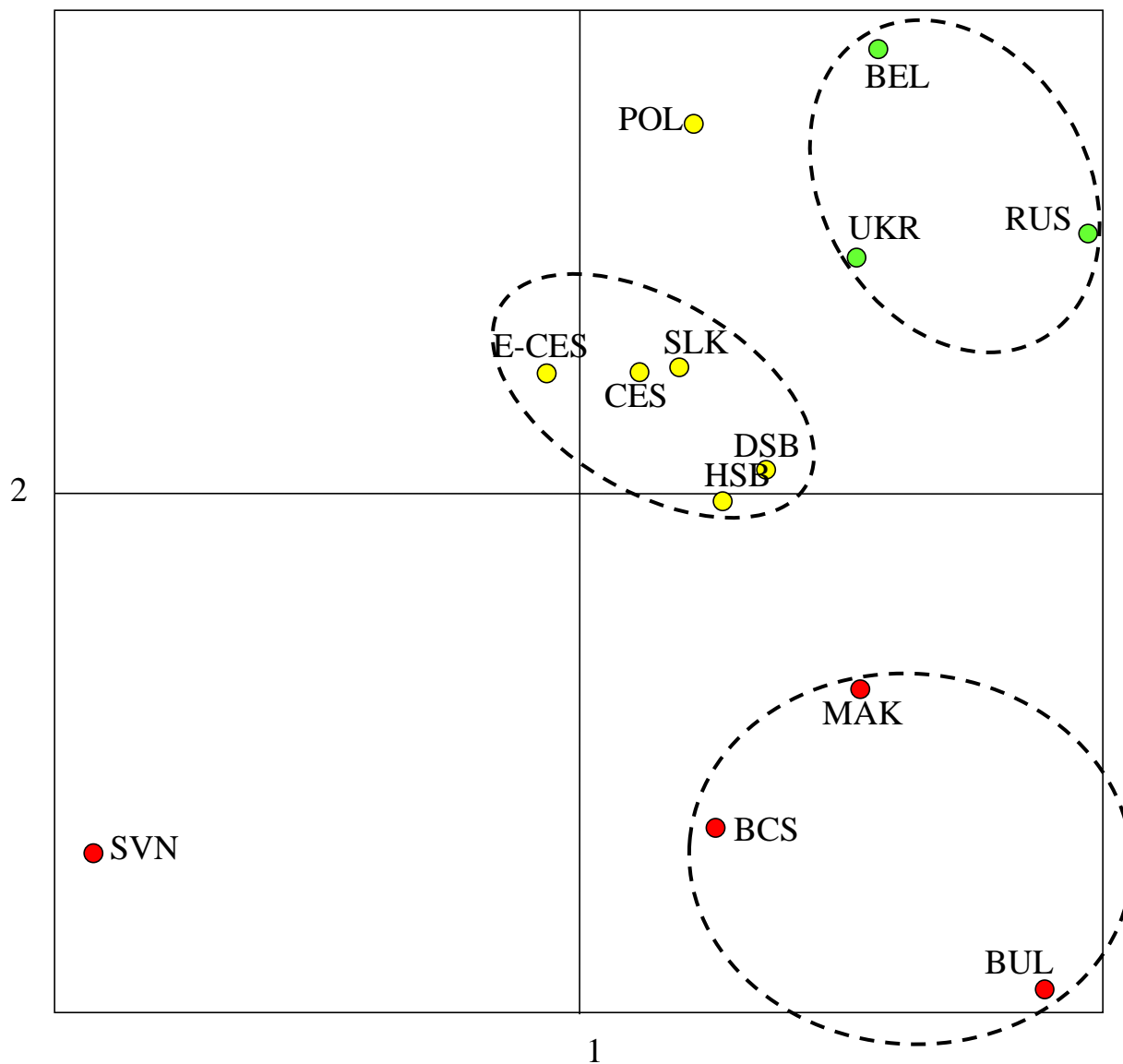


FIGURE 2. Pseudomap of Slavic Languages.

MDS plot: after Figure 2 Dyen, Kruskal & Black (1992)

200 cognates; 13 languages; % of shared cognates for Swadesh list



Mańczak 2004 – distances expressed as raw N of correspondences between pairs

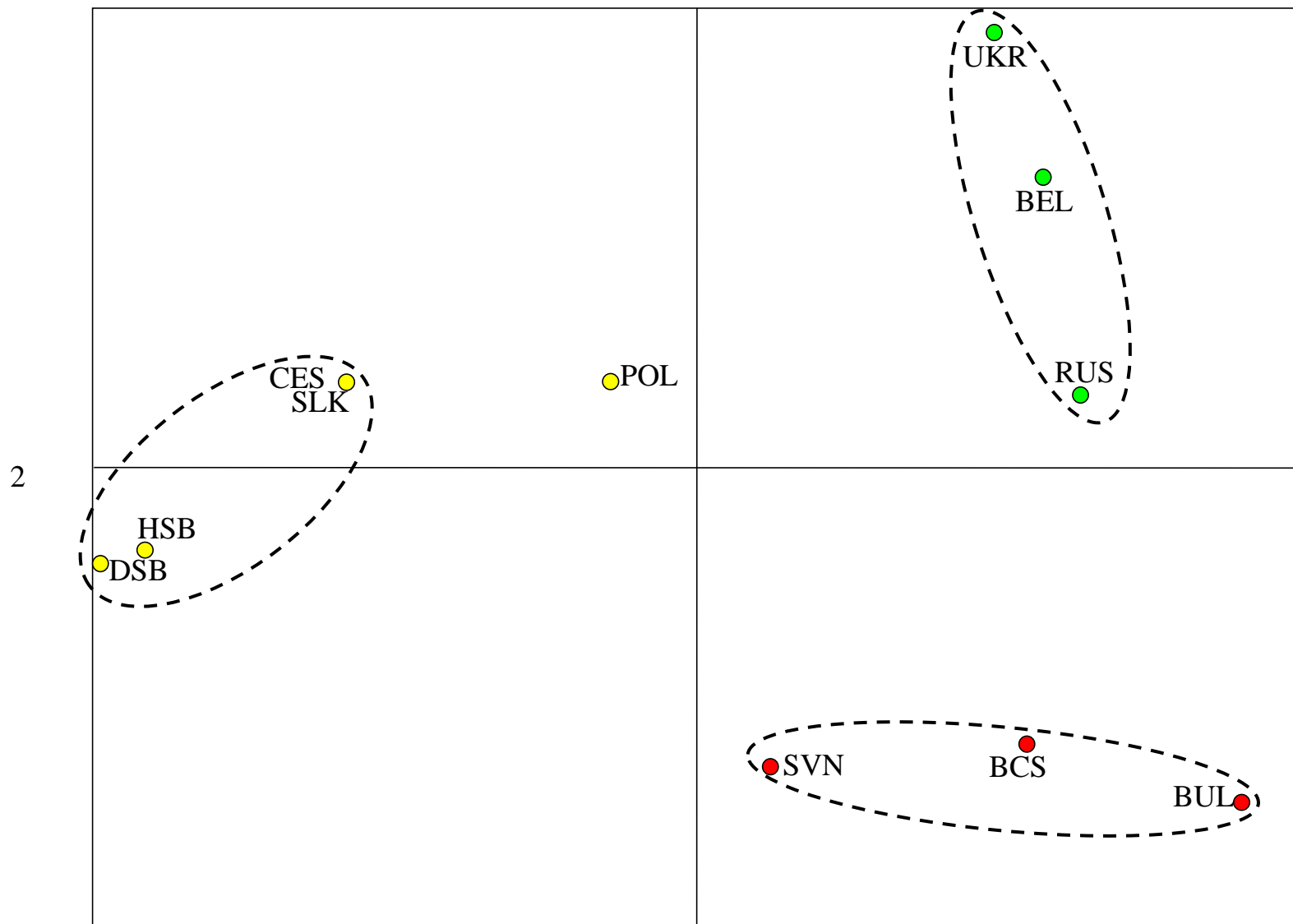
$k=1$ $\left(\frac{100}{6} \right)$ 10 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 10
 total tokens

Pod względem statystycznym podobieństwa leksykalne między językami słowiańskimi przedstawiają się następująco:

Polski	Czeski	Słowacki	Górno- łużycki	Dolno- łużycki	Rosyjski	Białoruski	Ukraiński	Słoweński	Serbo- chorwacki	Bułgarski
Cz 63.77 1158	Sła 74.06 1345	Cz 74.06 1345	Dł 63.6 1155	Gł 63.6 1155	Br 60.74 1103	Ukr 63.88 1160	Br 63.88 1160	Sch 57.32 1041	Sle 57.32 1041	Sch 55.45 1007
Sła 56.1 1118	Pol 63.33 1158	Pol 63.33 1158	Sła 55.67 1011	Cz 53.41 970	Sch 54.3 986	Ros 60.74 1103	Pol 49.56 900	Pol 52.4 956	Bg 55.45 1007	Ros 48.24 887
Sle 52.3 956	Gł 55.48 1002	Gł 55.48 1002	Cz 55.48 1002	Sła 53.21 950	Pol 49.67 902	Pol 52.32 951	Ros 49.56 900	Ros 49.23 894	Ros 49.23 894	Sle 47.36 860
Br 52.3 951	Dł 53.21 970	Dł 52.95 950	Pol 45.98 835	Pol 43.28 786	Ukr 49.56 900	Sch 46.04 836	Sła 40.25 731	Bg 47.36 860	Pol 48.81 887	Br 44.97 813
Ros 49.902	Sle 45.98 835	Sle 44.00 799	Sle 44.88 815	Sle 42.02 763	Sle 49.23 894	Bg 44.71 813	Sch 39.46 719	Cz 45.98 835	Br 46.04 836	Pol 40.47 735
Ukr 49.900	Ros 40.84 738	Ukr 40.25 731	Ukr 34.64 629	Br 32.99 599	Bg 48.24 887	Sle 42.77 777	Cz 38.94 698	Gł 44.88 815	Sła 39.87 724	Ukr 37.83 687
Sch 48.887	Sch 39.46 717	Sch 39.46 724	Sch 33.2 612	Sch 32.82 597	Cz 40.64 738	Sła 38.7 707	Sle 38.72 694	Sła 44.40 799	Ukr 39.87 719	Sła 32.43 589
Gł 45.835	Ukr 38.44 698	Ros 49.2 721	Br 33.2 612	Ros 32.82 596	Sła 39.7 721	Cz 38.16 693	Bg 37.82 687	Br 42.79 777	Cz 39.48 717	Cz 31.50 572
Dł 43.786	Br 38.16 693	Br 38.93 707	Ros 33.04 601	Ukr 32.592	Gł 33.09 601	Gł 32.7 612	Gł 34.64 629	Dł 42.02 763	Gł 33.2 612	Gł 29.19 530
Bg 40.735	Bg 31.50 572	Bg 32.43 589	Bg 29.19 530	Bg 27.86 506	Dł 32.82 596	Dł 32.98 599	Dł 32.6 592	Ukr 38.22 694	Dł 32.87 597	Dł 27.86 506
9228	8728	8695	7802	7514	8328	8251	7710	8434	8126	7186

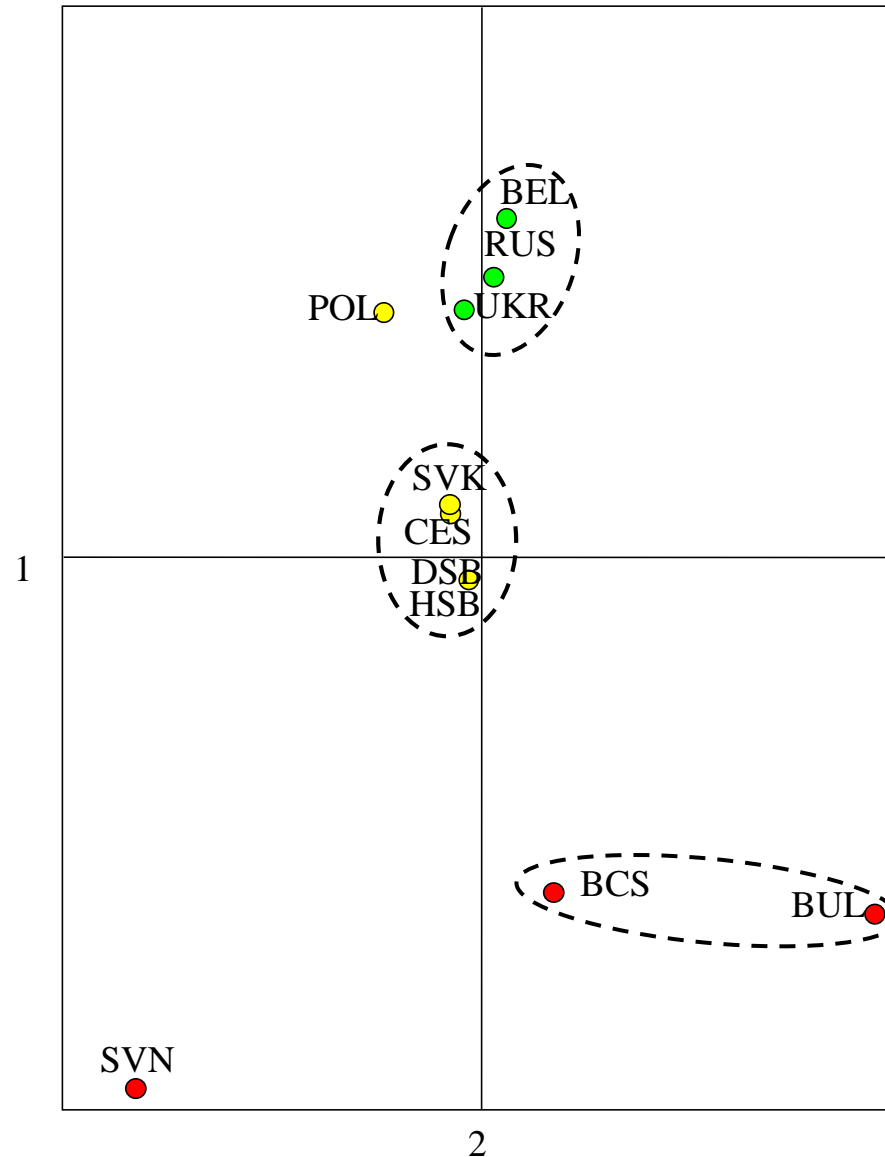
avg: 50.81 48.06 47.88 42.96 41.38 45.86 45.44 42.46 46.44 44.75 39.57

MDS-ML plot: 11 languages lexical items; this study
data from: Mańczak (2004), 1816 tokens from Gospel texts; % shared



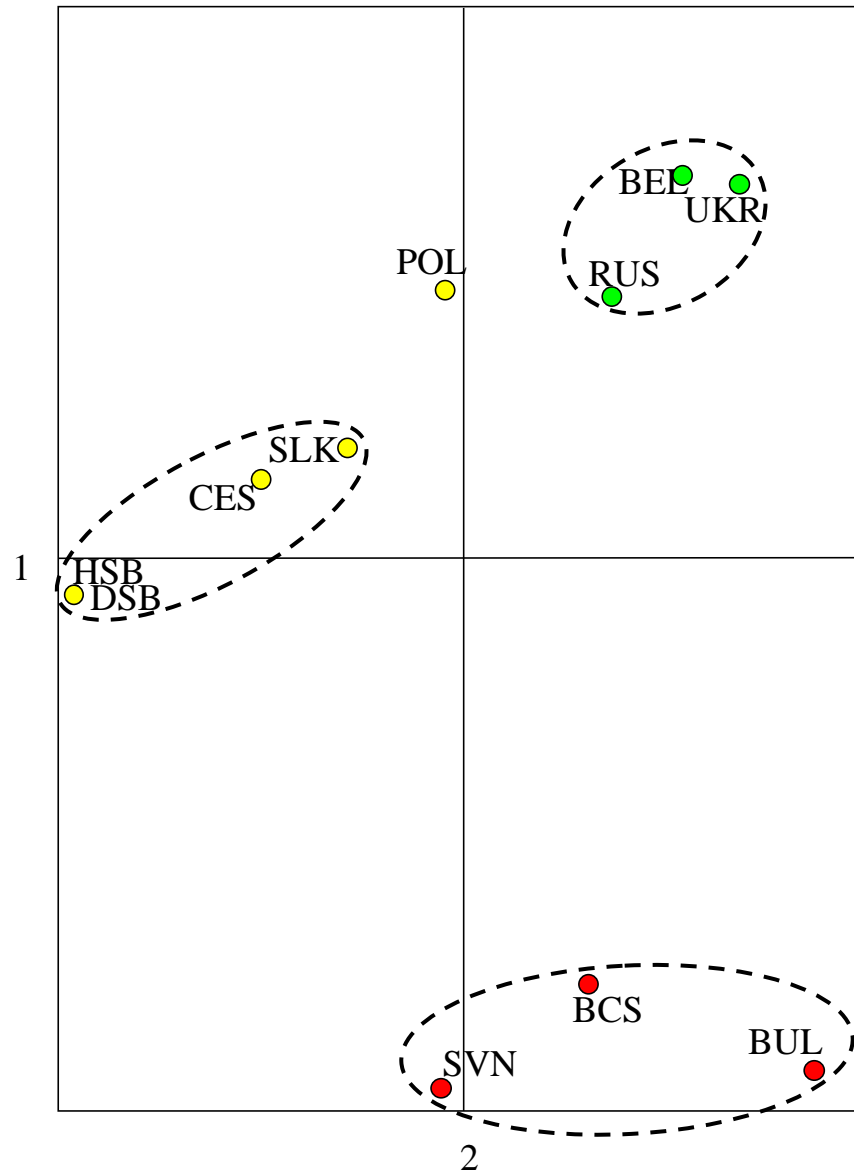
MDS-ML plot: 11 languages; this study

data from: Dyen, Kruskal & Black (1992), 200 cognates



MDS-ML plot: 11 Slavic languages; this study

Data from: Atkinson & Gray (2003); 315 cognates, Jaccard distance



Slide of lexical patterns with POL towards RUS (Mańczak data); POL = RUS ≠ UKR

88

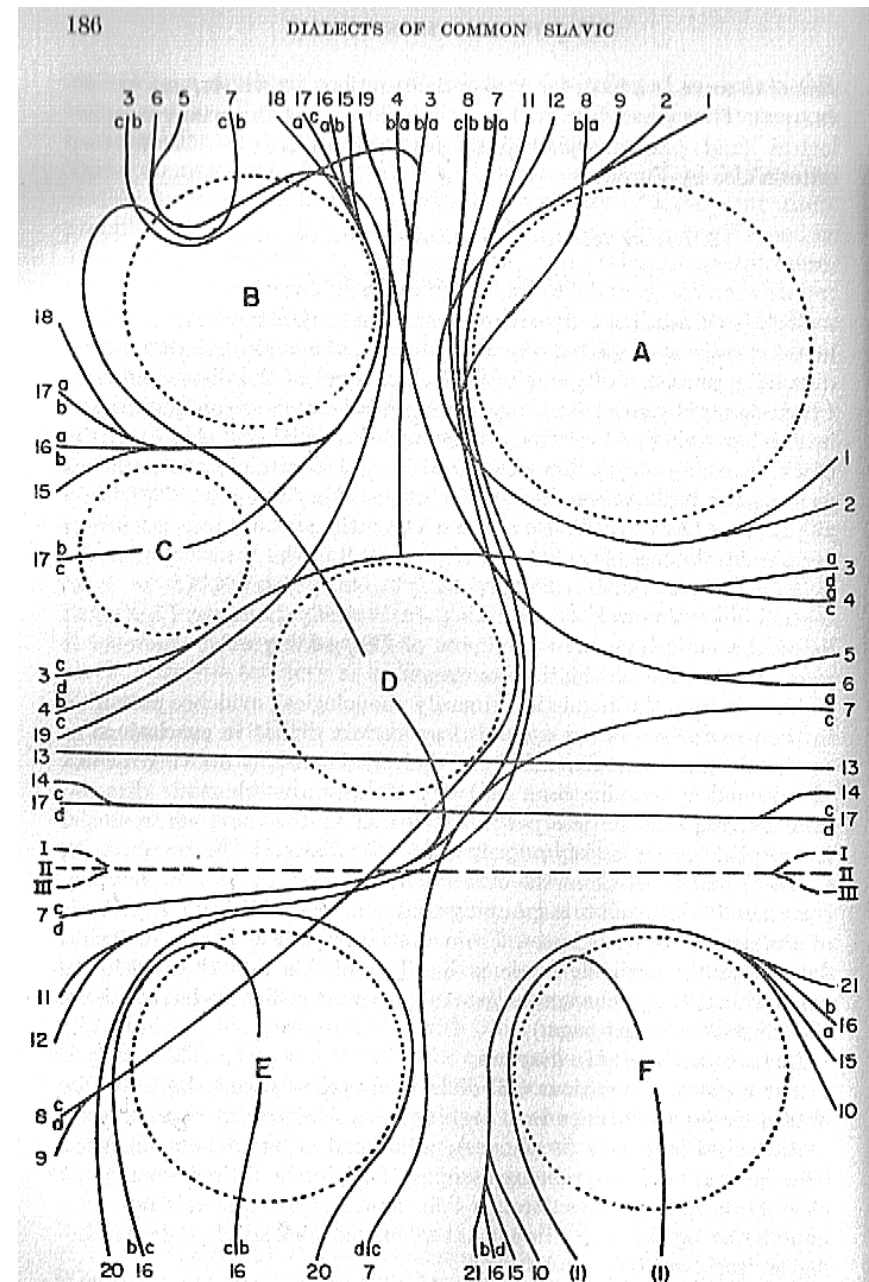
	Polski	Czeski	Slowacki	Górnolужицки	Dołnolужицки	Białoruski
1	uradować się A	zradovat se A	zaradovat' sa A	zwjeselić so B	wjasoły B	uzradavacca A
1	usiąść A	posadit se A	uložit' sa B	sydnyć so A	sednuć A	sjadac' A
1	usiąść A	posadit se A	uložit' sa B	zesydać so A	zasedać se A	sesci A
1	usiąść A	sedět A	sedět A	zaujmat' B	sedžeć A	sesci A
4	usłyszeć A	uslyšet A	počut' B	zasłyšeć A	stýšaš A	pačuc' B
1	ustąpić A	nechat B	nechat' B	dopuścić C	daši D	pakinue' E
1	ustąpić A	nechat B	nechat' B	zwolić D	psiwdaš E	dapuskac' C
1	ustąpić A	obrátiť se B	utiahnut' sa C	pođać so D	wustupiš A	pajšoŭ X
1	uszedł A	odšel A	odišiel X	ėknyć B	wustupiš C	pajšoŭ X
1	uwielbić A	oslawit B	oslawiť B	woslawić B	wucesiś C	uslawi'ć B
1	uzdrawiać A	uzdravovat A	uzdravovat' A	wustrowjeć A	gójš B	azdaraŭljac' A
1	uzdrowić A	uzdravit A	uzdravit' A	wustrowić A	wugójš B	azdaravic' A
1	uzdrowić A	uzdravit A	uzdravit' A	wustrowiś A	wustrowiś A	pazdaravec' A
1	użalić się A	milosrdenstvf B	útrpnost' C	miłosć B	lutowaś D	zlitavacca D
8	w A	do B	do B	do B	do B	u A
3	w A	do B	do B	do B	do B	u A
1	w A	do B	na C	do B	do B	u A
1	w A	na B	do C	do C	do C	u A
1	w A	na B	do C	do C	do C	u A
1	w A	na B	do C	do C	do C	u A
2	w A	na B	na B	do C	do C	u A
1	w A	na B	na B	w A	w A	u A
2	w A	s B	s B	z B	z B	z B
2	w A	v A	do B	do B	do B	u A
2	w A	v A	do B	do B	do B	u A
1	w A	v A	do B	w A	we A	u A
1	w A	v A	v A	na B	na B	u A
1	w A	v A	v A	w A	na B	na B
1	w A	v A	v A	w A	po B	u A
1	w A	v A	v A	w A	w A	u A
4	w A	v A	v A	w A	w A	u A
1	w A	v A	v A	w A	w A	u A
1	wchodzić A	chodit A	vojsť B	chodźić A	hyś B	zachodzić A
1	wchodzić A	veházet A	veházat' A	wrócić B	hyś C	uwachodzić A
1	według A	podle X	podl'a X	po B	za C	vodle A
1	wejść A	vejít A	vojsť A	dónć A	hyś A	uvajsci A
3	wejść A	vejít A	vojsť A	zastupić B	hyś A	uvajsci A
1	wesele A	svadba B	svadba B	hosćina C	vjasella A	brak D
1	weselić się A	veselit se A	hodovat' B	kwasaować C	wjaseliś se A	cešycca D
1	weselić się A	veselý A	hodovat' B	kwasaować C	wjaseliś se A	vesjalicca A
1	weselić się A	veselý A	hodovat' B	wjeselić so A	wjeseliś se A	pavesjalicca A
1	weselić się A	veselý A	hodovat' B	wjeselić so A	wjeseliś se A	vesjalicca A
1	westchnąć A	vzdechnout A	zastonat' B	pozdychnyć A	zdychnuś A	udzychnuć A
2	wezwać A	povolat B	povolat' B	zwolać B	pówolaś B	pryklikac' C
1	wezwać A	povolat B	zwolat' B	zwolać B	pšosyś C	zaklikac' D
1	wezwać A	pozvat A	povolat' B	hosćićel C	pšosyś D	prasic' D
1	wezwać A	pozvat A	povolat' B	prosyć C	pšosyś C	zaprasic' C
2	wezwać A	pozvat A	pozvat' A	hosćićel B	pšosyś C	prasic' C
1	wezwać A	pozvat A	pozvat' A	prosyć B	pšosyś B	zaprasic' B
1	wezwać A	pozvat A	pozvat' A	prosyć B	pšosyś B	zaprasic' B

89

	Rosyjski	Ukraiński	Serbochorwacki	Słoweński	Bułgarski
	obradovat'sja A	zradity A	obradovati se A	razveseliti se B	zaradvam se A
	sadit'sja A	sidaty A	sjesti A	sesti A	sedna A
	vozleć' B	posidaty A	posjedati A	posedeti A	nasedja A
	zanjať' B	zajmaty B	sjesti A	sedeti A	zaemam B
	uslyšať' A	začuty B	čuti B	zaslišati A	čuja B
	ostavit' F	dopustyty C	ostaviti F	pustiti C	ostavja F
	dopuskat' C	dopustyty C	ostaviti F	pustiti C	ostavja F
	pošel E	vidijšov X	otići F	kreniti G	otteglja se C
	pošel A	pišov X	otići D	pobegniti E	otida D
	proslavit' B	proslavyty B	proslaviti B	oslawiti B	proslavja B
	vračevat' C	uzdorovljaty A	iscjelivati D	zdraviti A	celja D
	isceelit' C	uzdorovyty A	iscjeliti C	ozdraviti A	izcelja C
	iscelet' B	odružaty C	iscjeliti B	ozdraveti A	ozdravja A
	sžaliti'sja A	žal' A	sažaliti se A	smiliti se B	smilja se B
	v A	do B	u A	v A	v A
	v A	do B	u A	v A	v A
	v A	do B	u A	v A	v A
	v A	do C	u A	v A	na B
	v A	do C	u A	v A	v A
	v A	v A	u A	v A	v A
	v A	do C	u A	v A	v A
	v A	v A	u A	v A	v A
	v A	z B	s B	z B	s B
	v A	do B	u A	v A	v A
	v A	v A	u A	v A	v A
	v A	u A	u A	v A	v A
	v A	na B	na B	v A	na B
	v A	v A	na B	v A	v A
	v A	u A	u A	v A	v A
	v A	u A	po B	po B	v A
	v A	u A	u A	v A	na B
	v A	u A	u A	v A	nad B
	zachodit' A	zachodyty A	ulaziti C	hoditi A	vlizam C
	vchodit' A	vchodyty A	ulaziti D	priti C	vlizam D
	po B	vodle A	po B	priti C	spored D
	vojtji A	uvijty A	učii A	priti A	vlizam B
	vojtji A	uvijty A	učii A	priti A	vlizam C
	brak D	vesillja A	svadba B	svatovščina B	svatba B
	radovat'sja E	veselytsja A	veselytsja A	razveseliti se A	razveselja se A
	veselit'sja A	veselytsja A	veselitysja A	veseliti se A	veselja se A
	poveselit'sja A	potišytsja C	radity C	proveseliti se A	poveselja se A
	veselit'sja A	radity C	radity C	veseliti se A	veselja se A
	vzdechnut' A	zidchnuty A	uzdahnuti A	vzdihnuti A	váždžchna A
	prizvat' A	pryklykaty C	prizvat' A	pryklykaty C	povikam D
	zvat' A	klykaty D	zvat' A	klykaty D	pokanja F
	zvat' A	poklykaty E	zvat' A	poklykaty E	pokanja G
	zvat' A	zaprasyty C	zvat' A	zaprasyty C	pokanja E
	zvat' A	poklykaty D	zvat' A	poklykaty D	pokanja F
	pozvat' A	poklykaty C	pozvati A	poklykaty C	pokanja E
	pozvat' A	zaprosyty B	pozvati A	zaprosyty B	pokanja D

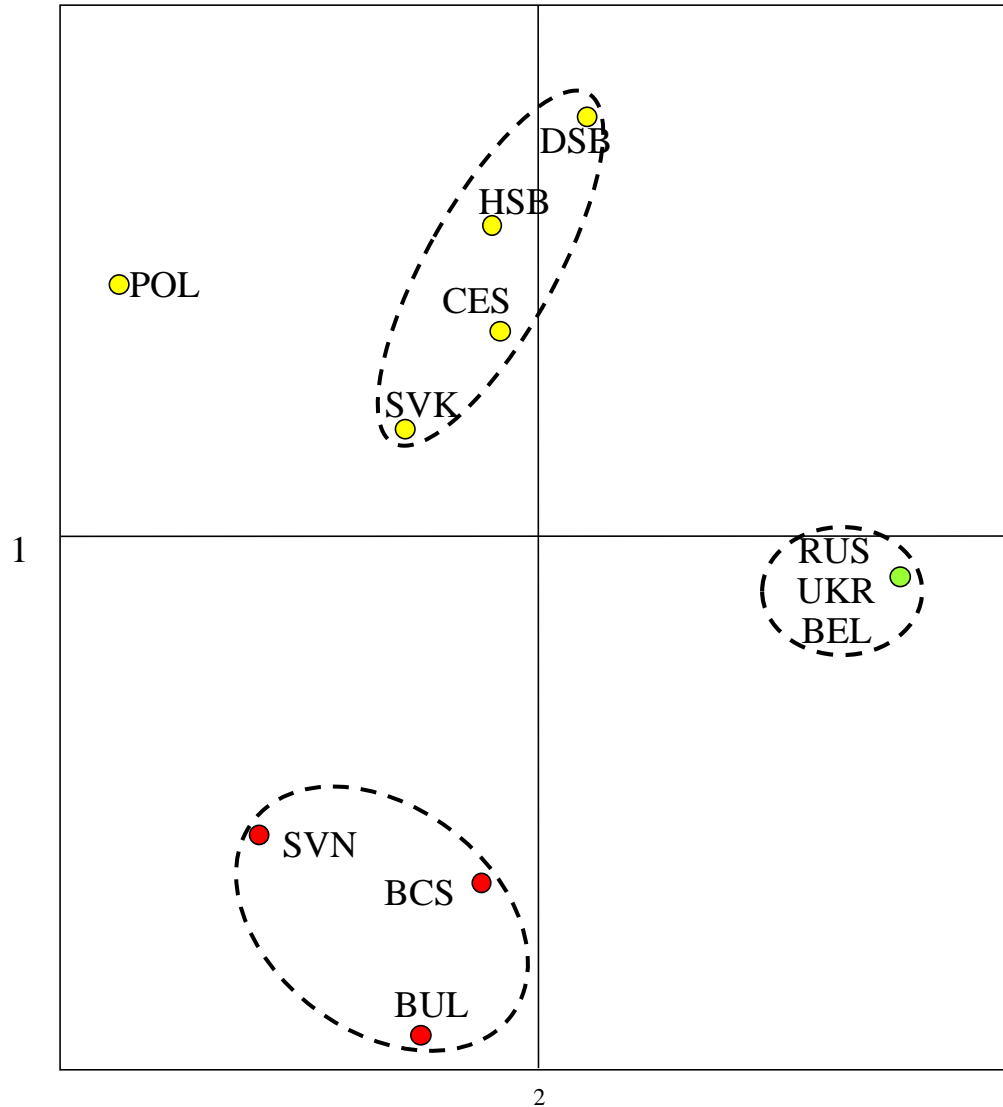
Birnbaum 1966: Phono- and morphological isoglosses

- A = East Slavic
- B = Lekhitic
- C = Sorbian
- D = Czecho-Slovak
- E = Slovene/BCS
- F = Macedo-Bulg.



MDS plot 11: Slavic phonological innovations; this study

data from: Birnbaum (1966); 40 isoglosses; Jaccard distance

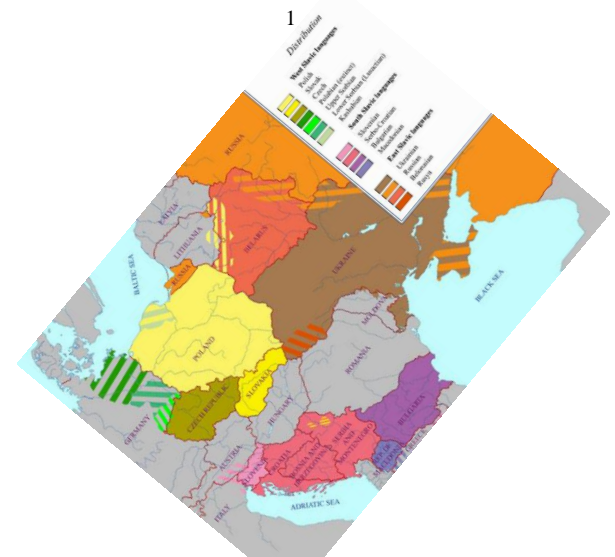
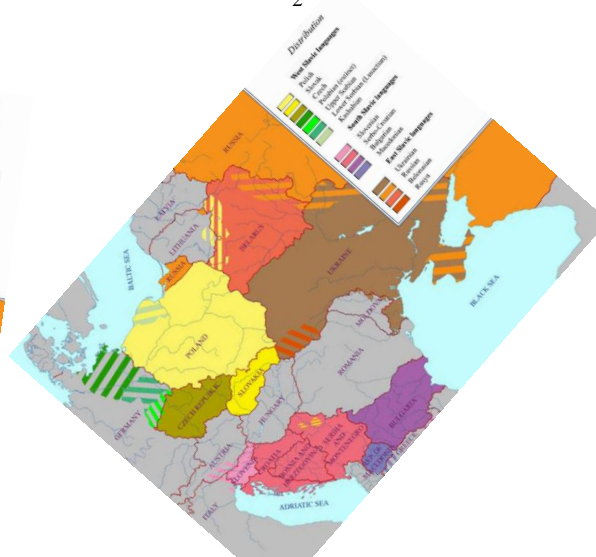
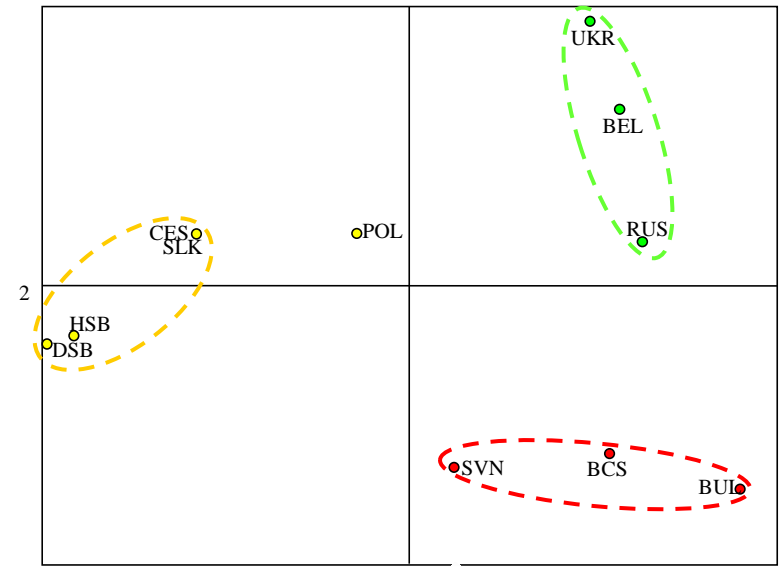
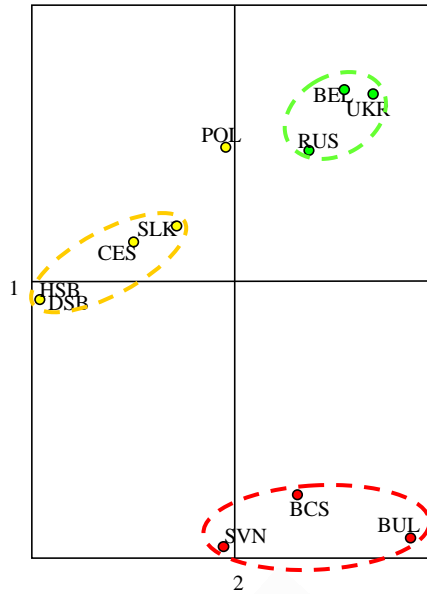
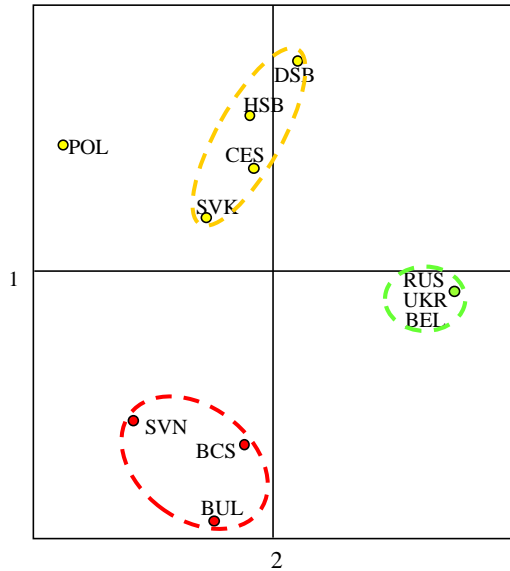


Summary of MDS plots; this study

Birnbaum-1966

G&A-2003

Mańczak-2004



Correlations with geography and MDS plots

Data set	Geography correlation ¹	p-value
Dyen-1992	0.381	ns
G&A-2003	0.587	p < 0.05
Manczak-2004	0.531	p < 0.05
Birnbaum-1966	0.516	p < 0.05

¹Mantel Test

Correlations among MDS plots—data sets

	Dyen-1992	G&A-2003	Manczak-2004
G&A-2003	0.758		
Manczak-2004	0.319	0.728	
Birnbaum-1966	0.501	0.698	0.672

Mantel test; all comparisons $p < 0.05$