La géographie des phénomènes phonologiques en Europe
The geography of phonological phenomena in Europe

Thomas Stolz & Aina Urdze
Université de Brême/Allemagne

Pays et paysage linguistiques

28-29 janvier 2011
Paris

Un mot en français

Cette conférence sera présentée en anglais.

Il y a troisraisons pour notre choix de langue de présentation:

1) Les dégrés de compétence en français varient beaucoup (y compris compétence zéro) dans notre groupe de recherche.
2) Le projet dans lequel cette recherche s‘encadre utilise l‘anglais comme langue de travail.
3) Nous avons l‘intention de publier les actes de ce colloque chez Akademie Verlag à Berlin où on n‘accepte que des textes en anglais.

   Nous espérons néanmoins que vous trouverez intéressant ce que nous irons vous présenter.

Table of contents

Aims and goals

This presentation is meant to adduce evidence of the areality of phonological phenomena.

Phonemes are not randomly distributed across the languages of Europe. To the contrary, they follow a geolinguistic logic (just as linguistic categories of the morphosyntactic kind).

To prove our point, we discuss the isoglosses which result from a check of the system of sibilants and affricates in 75 languages of Europe (complemented by additional data from 26 languages of Asia).

Motivation

In their search for structural traits which are typical of the languages of Europe, Haarmann (1976), Ternes (1998, 2010) and Haspelmath (2001) depict the phonological level as being of no interest for geolinguistic studies.

- Haarmann claims that phonological features are randomly distributed in Europe such that no areal clusters are discernible on the map.
- Ternes insists on European phonologies being nothing special in global perspective and thus are not distinctive.
- Haspelmath is of the opinion that there is no single phonological feature which is useful for the identification of the SAE-sprachbund.

Claims

We take issue with our colleagues‘ negative stance because they are too preoccupied with pan-Europeanisms.

We claim that

- areal linguistics is not (only) about identifying continent-wide common traits,
- areal linguistics cannot be reduced to deal exclusively with the spectacular/unique/distinctive,
- areal linguistics has to determine the geolinguistic diffusion of all kinds of phenomena within a given region.

Inspiration

- Studies by Jakobson (1931), Wagner (1964) and Stadnik (2002) demonstrate that it makes sense linguistically to study phonological phenomena in geolinguistic perspective.
- EUROTYP has avoided largely to address phonological issues (exception van der Hulst 1999) such that there is a huge gap to be filled (especially in the light of the failure of Décsy [1973, 2000] who gives a absolutely distorted picture of the European phonological landscape).
- Haspelmath (2001) leaves open a backdoor for areal phonology when he wonders whether perhaps phonologists simply have not looked closely enough at the data to come up with suitable isoglosses.

Project

Our dissatisfaction with the neglect of phonology in the framework of EUROTYP has induced us to accept the challenge ourselves.

This is how the idea of the project EUROPHONOLOGY developed at the University of Bremen.

[Despite the terminological similarity, our project is by no means associated to EUROLINGUISTICS!]

This project aims at providing a full-scale evaluation of all matters phonological connected to the areal-linguistic makeup of Europe.

Previous studies

- Stolz (2004): determines the distribution of phonemic /x/ and /h/ and thereby disproves Décsy‘s „Euroversal“ which assumes that the two phonemes cannot co-exist.
- Stolz (2006): demonstrates that the diffusion of rounded front vowels and phonemic quantities of vowels and consonants is not random in Europe.
- Stolz (2007): shows that the share of monosyllables in the long Swadesh list yields an areally significant pattern with a center of high monosyllabicity in the European Northwest.
- Stolz/Urdze/Otsuka (2010): focuses on the class of liquids which is shown to grow in size on the margins of the European continent.
- Stolz/Urdze/Otsuka (in press a): (post-)velar phonemes distribute over the European map such that various sub-areas can be distinguished.
- Stolz/Urdze/Otsuka (in press b): rare phonemes in Europe tend to accumulate in two niches on the European continent.

Starting-point

- In a sideways remark, Ternes (1998) observes that sibilants and affricates fail to be attested in equal numbers throughout the European area.
- In Stolz (2010), this hypothesis is put to the test – the empirical data corroborate Ternes‘s „intuition“ and, at the same time, suggest that phonological phenomena may be distributed unevenly but in a geolinguistically interesting way across the European languages.

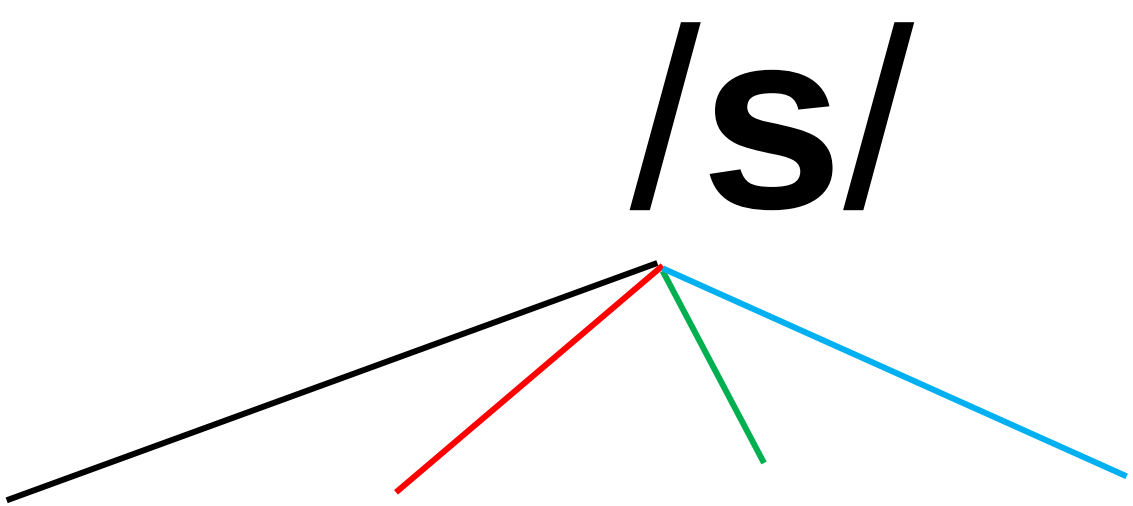
Today

- We elaborate on the earlier findings.
- The original sample of 75 European languages is complemented by a sample of 26 Asian languages.
- Europe includes Trans-Caucasia, Anatolia, Cyprus, Malta, Greenland.
- We highlight the skewed geolinguistic distribution of phonemic sibilants and affricates by way of identifying isoglosses.
- More often than not, we refer to standard varieties.
- We draw our data from the extant descriptive grammars of our sample languages.
- Our approach considers phonemes as discreet countable elements.
- Affricates are complex phonemic units.

Candidates

- We look at eight phonemes for which the International Phonetic Alphabet provides distinct symbols.
- These phonemes are produced apically/laminally/predorsally by friction in the (denti-)alveolar and postalveolar region.

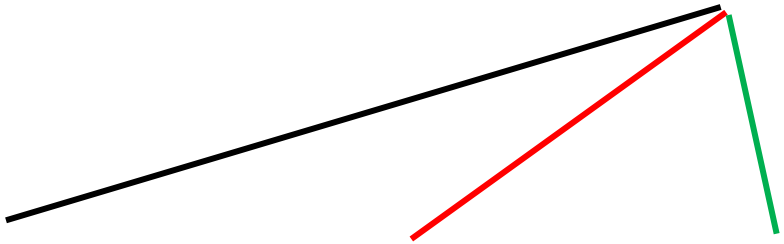
Voiceless alveolar fricative



plain	**palatalized**	**labialized**	**glottalized**	**aspirated**
French	Lithuanian	Lezgian	Chamalal	-----
sous /su/	*siusti* /sʲùsti/	*swax* /sʷaχ/	/s‘:a:/	
‚under‘	‚to carouse‘	‚molar‘	‚rain‘	

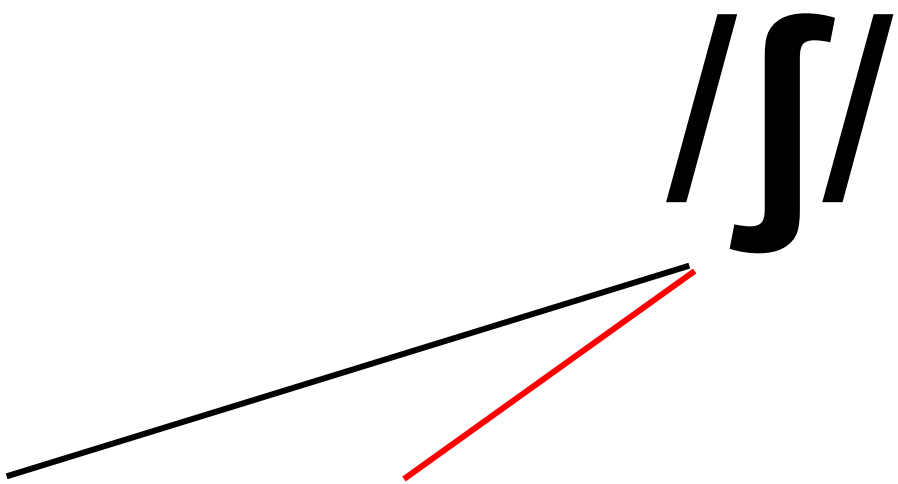
Voiced alveolar fricative

/z/



plain	**palatalized**	**labialized**	**glottalized**	**aspirated**
French	Ukrainian	Lezgian	-----	-----
zéle /zɛl/	*вʻязь* /vjazʲ/	*zwal* /zʷal/		
‚zeal'	‚bundle'	‚boiling'		

Voiceless postalveolar fricative



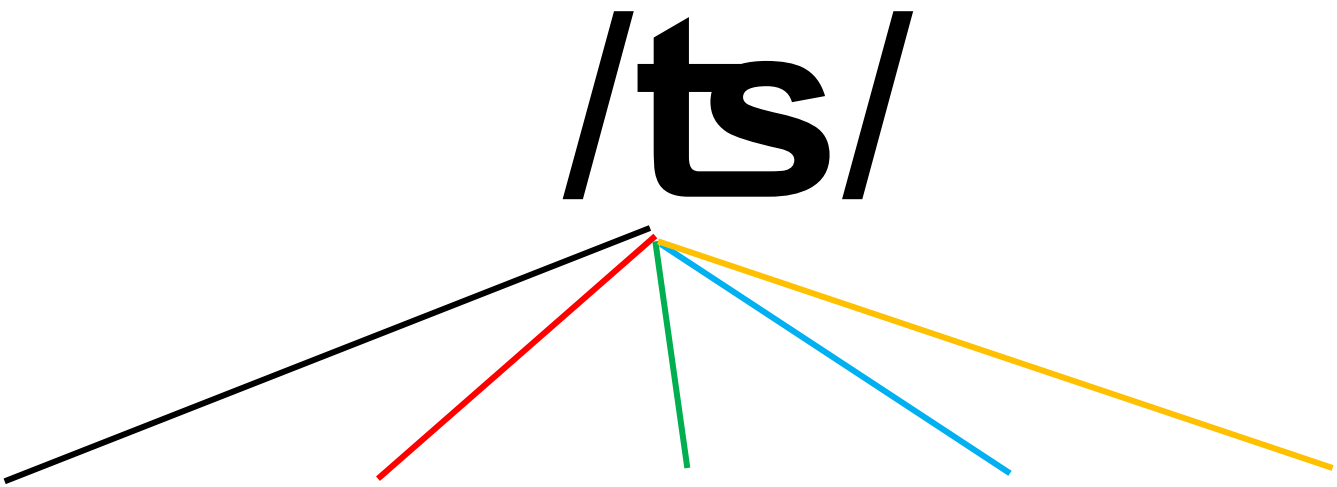
plain	**palatalized**	**labialized**	**glottalized**	**aspirated**
French	Russian	(Lezgian	-----	-----
chou /ʃu/	*щель* /ʃʲelʲ/	dialects)		
‚cabbage'	‚crack'			

Voiced postalveolar fricative

/ʒ/

plain	**palatalized**	**labialized**	**glottalized**	**aspirated**
French	Russian	(Lezgian	-----	-----
joue /ʒu/	*дожеǔ* /daʒʲej/	dialects)		
‚cheek'	‚rains (genitive)'			

Voiceless alveolar affricate



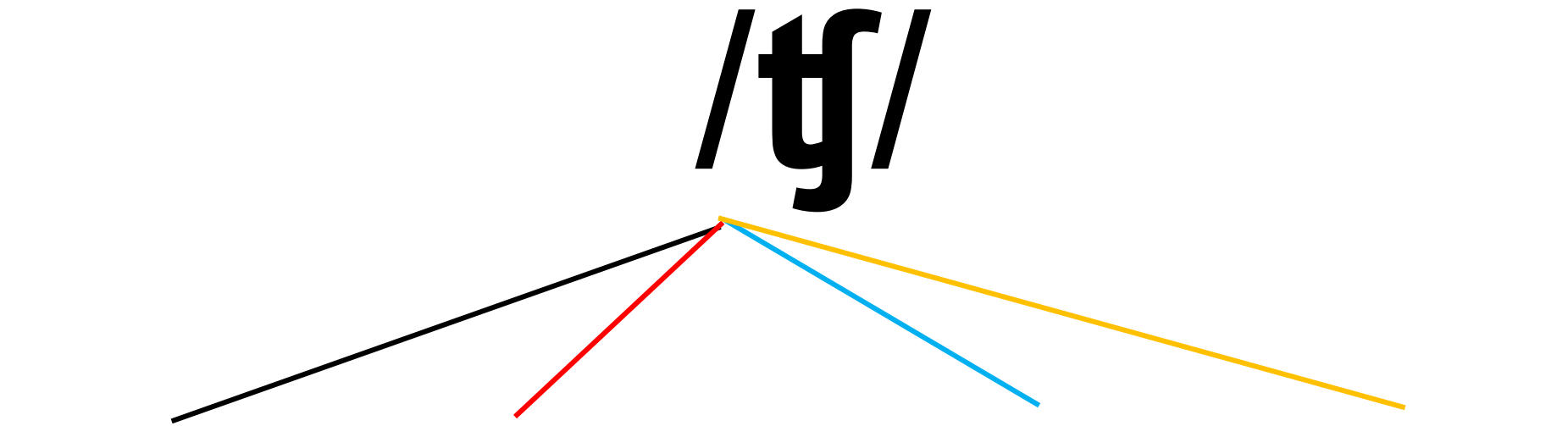
plain	**palatalized**	**labialized**	**glottalized**	**aspirated**
German	Belarusian	Lezgian	Georgian	Armenian
Zahn /ʦa:n/	*быць*/bitsʲ/	*cwar* /ʦʷar/	*c‘igni* /ʦ‘igni/	*c‘ic‘* /ʦʰitsʰ/
‚tooth‘	‚to be‘	‚urine‘	‚book‘	‚stake‘

Voiced alveolar affricate

/dz/

plain	**palatalized**	**labialized**	**glottalized**	**aspirated**
Latvian	Belarusian	-----	-----	-----
dzert /dzɛrt/	*людзи* /lʲùdzʲi/			
‚to drink‘	‚people‘			

Voiceless postalveolar affricate



plain	**palatalized**	**labialized**	**glottalized**	**aspirated**
Hungarian	Ukrainian	(Lezgian	Tsez	Armenian
csak /ʧɑk/	*ніччю* /nʲiʧʲu/	dialects)	*meč‘* /meʧ‘/	*č‘ał* /ʧʰɑʁ/
‚only‘	‚by night‘		‚neck‘	‚fat‘

Voiced postalveolar affricate

/dʒ/

plain	**palatalized**	**labialized**	**glottalized**	**aspirated**
Italian	Ukrainian	----	----	----
giorno /dʒɔrno/	*ніччю* /bdʒilonʲka/			
‚day'	‚little bee'			

Specials

- Accumulated secondary articulations:
  Lezgian: aspiration + labialization
  /ʧʰʷ/: *čuk‘ul* ‚knife‘ /ʧʰʷkʻul/
- Unusual phonemic distinctions:
  Basque: apical vs. laminal

apical:	/ʂ/ *soka* ‚rope‘ /ʂoka/
	/ʦ̻/ *atso* ‚old woman‘ /aʦ̻o/
laminal:	/s̻/ *zopa* ‚soup‘ /s̻opa/
	/ʦ̺/ *atzo* ‚yesterday‘ /aʦ̺o/

Outcasts

- alveolo-palatal fricatives:
  Irish: /ɕ/ *sí* „she‘ /ɕi/

- alveolo-palatal affricates:
  Croatian:
  /tɕ/ *leća* ‚lentils‘ /le:tɕa/
  /dʑ/ leđa ‚back‘ /le:dʑa/

General observations

a) In every sample languages, there exists at least one of the eight phonemes under scrutiny.

b) In each language of the sample, /s/ is a phoneme.

c) Owing to the differences in the phonemicity of secondary articulations, the total number of phonemes in the classes of sibilants and affricates ranges from the minimum of one to the maximum of sixteen phonemes.

Statistics: phoneme frequency

Rank	phoneme	languages	share	comment
1	/s/	75	100%	fricatives
2	/ʃ/	62	83%	
3	/z/	58	77%	
4	/ʧ/	50	67%	affricates
5	/ʦ/	43	57%	
6	/ʤ/	35	47%	voiced
7	/ʒ/	33	44%	
8	/ʣ/	21	28%	

Implications

- Trivial

$$C_{\left\{\begin{array}{l}\text{[+sibilant]}\\ \text{[+affricate]}\end{array}\right\}} \supset$$ */s/*

unmarked

Details

/dʒ/ ⊃ /tʃ/
/dz/ ⊃ /ts/
/ʒ/ ⊃ /ʃ/
/z/

} ⊃ /s/

+voice ⊃ ¬voice

Below the average of **5.96** phonemes per language

Pho-nemes	West		East	
	North	South	North	South
5				Sephardic
4	Faeroese, German, Letzebuergesh	Corsican, French, Occitan, Portuguese	**Tatar**	Greek, Kurmancî
3		Galician, Sardinian	**Chuvash**, **Votic**	
2	Breton, Dutch, Frisian, Irish, Low German, Norwegian, Scots-Gaelic, Swedish, Welsh	Spanish		
1	Danish, **Greenlandic**, Icelandic		**Finnish**	
	Northwest = 15	Southwest = 7	Northeast = 4	Southeast = 3
	West = 22		East = 7	
	29 (North = 19, South = 10)			

High score

Pho-nemes	West		East	
	North	South	North	South
16			Lithuanian	Ukrainian
15				**Lezgian**
12			Belarusian, **Saami**	
10			**Mordvin**, Russian, **Udmurt**	Armenian, **Georgian**, **Laz**, Romanian
9			**Ingrian**, **Veps**	
8			Latvian, **Livonian**, Polish	Albanian, Arumanian, Bulgarian, **Hungarian**, Macedonian, Serbian, Slovak
7		Italian, Rumansh	**Karelian**	Croatian, **Kalmyk**, Slovenian
6	English	**Basque**, Catalan, Friulan, **Maltese**	**Bashkir**, **Estonian**, **Mari**, Sorbian, **Yurak**	**Aramaic**, Czech, **Kazakh,** Ossetic, Romani, **Turkish**
	NW = 1	Southwest = 6	Northeast = 17	Southeast = 22
	West = 7		East = 39	
	46 (North = 18, South = 28)			

Regional asymmetries

	West = 29 languages / 39%		**East = 46 languages / 61%**		
North = 37 languages / 49%	High 1	Low 15	High 17	Low 4	High 18 ~ 49% Low 19 ~ 51%
South = 38 languages / 51%	High 6	Low 7	High 22	Low 3	High 28 ~ 74% Low 10 ~ 26%
	High 7 ~ 24%	Low 22 ~ 76%	High 39 ~ 85%	Low 7 ~ 15%	

Spellout

- While northern languages are equally divided between phonemes inventories above and below the average,
- southern and eastern languages display a clear preference for systems which exceed the average,
- whereas in the West, languages equally clearly tend to be content with small numbers of phonemes of the classes at issue.

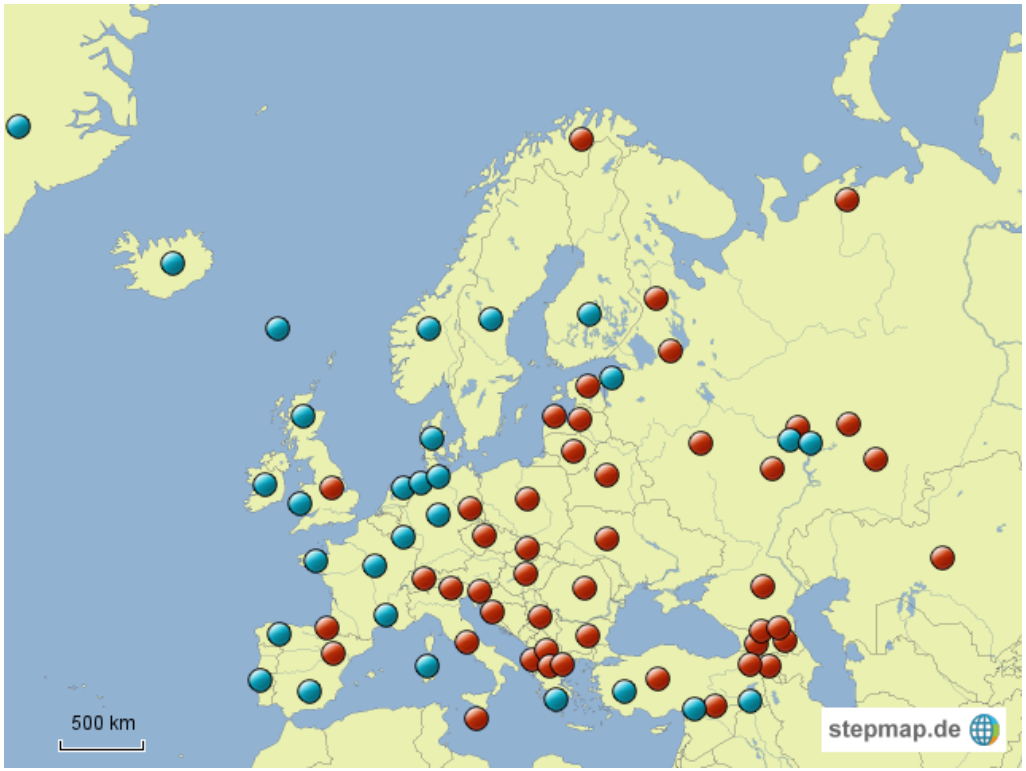
Genetic bias

	High	Low	
Indo-European	25 = 51% of phylum 54% of score	24 = 49% of phylum 83% of score	49 = 65% of sample
Non-Indo-European	21 = 81% of phylum 46% of score	5 = 19% of phylum 17% of score	26 = 35% of sample
	46 = 61% of sample	29 = 39% of sample	75 = 100%

In simple words…

- While Indo-European languages are divided in two groups of almost equal size as to high and low scores,
- Non-Indo-European languages count overwhelmingly among the high scorers
- whereas they are clearly underrepresented in the class of languages which fail to reach the average
- a class which is dominated by Indo-European languages.

High score vs low score



Red dots: above average

Blue dots: below average

Hinterm Horizont geht‘s weiter…

- To complement our European data, we extend our view beyond the Ural Mountains in order to see whether or not the languages in the neighbouring regions of Asia replicate the geolinguistic patterns identified in Europe.
- **Hypothesis**:

  The East-European properties are continued in Siberia and the Middle East.

Below the average of **4.03** phonemes per language

		/s/	/z/	/ʃ/	/ʒ/	/ts/	/dz/	/tʃ/	/dʒ/
4	Buriat Itelmen	/s/ /s/	/z/ /z/	/ʃ/	/ʒ/			/tʃ/, /tʃʲ/	
3	Khanty Even Nanai Oroch Orok	/s/, /sʲ/ /sʲ/ /s/ /s/ /s/		/ʃ/				/tʃʲ/ /tʃ/ /tʃ/ /tʃ/	/dʒʲ/ /dʒ/ /dʒ/ /dʒ/
2	Ket Mansi Nganasan Nivkh	/s/, /sʲ/ /s/, /sʲ/ /s/ /s/	 /z/		/ʒ/				
1	Yakut Evenki Yukaghir Koryak Chukchi	/s/ /s/ /s/		/ʃ/				/tʃ/	
16 languages		/s/ = 13 /sʲ/ = 4	/z/ = 3	/ʃ/ = 3	/ʒ/ = 2	0	0	/tʃ/ = 5 /tʃʲ/ = 2	/dʒ/ = 3 /dʒʲ/ = 2

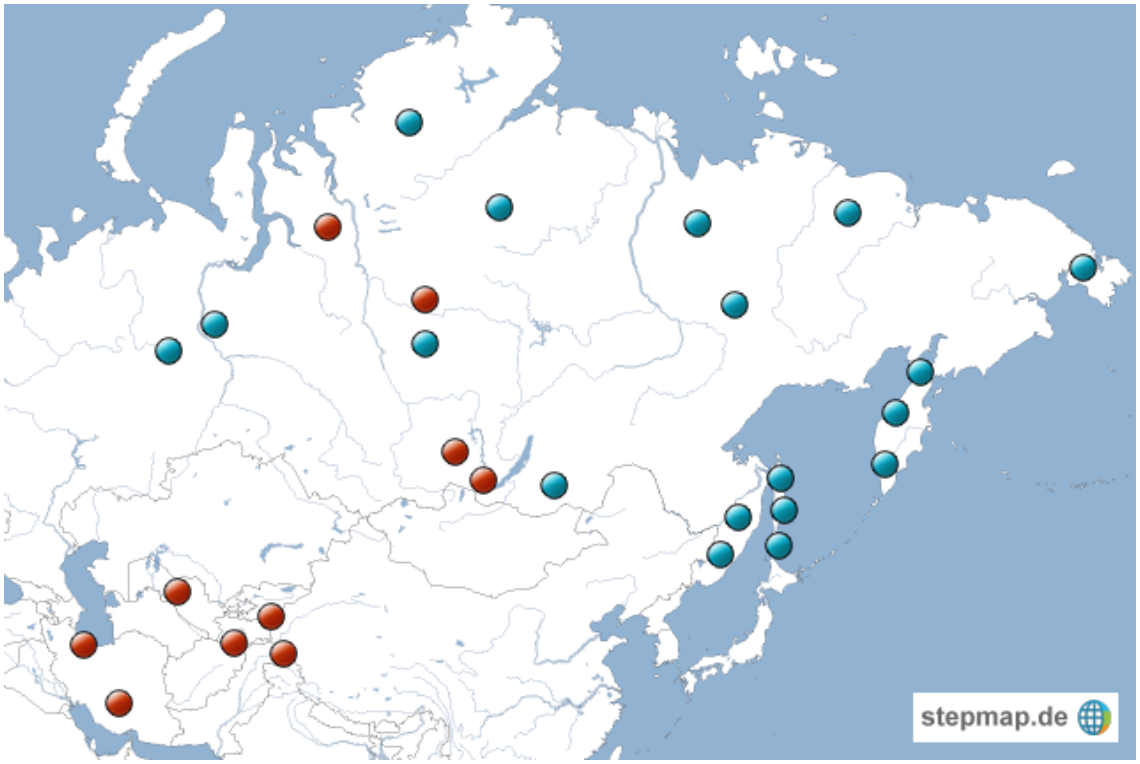
Asian top scorers

		/s/	/z/	/ʃ/	/ʒ/	/ʦ/	/ʣ/	/ʧ/	/ʤ/
12	Wakhi	/s/, /ʂ/	/z/, /ʐ/	/ʃ/	/ʒ/	/ʦ/, /ʈʂ/	/ʣ/, /ʤʐ/	/ʧ/	/ʤ/
8	Rushan	/s/	/z/	/ʃ/	/ʒ/	/ʦ/	/ʣ/	/ʧ/	/ʤ/
8	Kamass	/s/, /sʲ/	/z/, /zʲ/	/ʃ/	/ʒ/	/ʦʲ/	/ʣʲ/		
7	Uzbek	/s/	/z/	/ʃ/	/ʒ/	/ʦ/		/ʧ/	/ʤ/
6	Gilaki	/s/	/z/	/ʃ/	/ʒ/			/ʧ/	/ʤ/
6	Persian	/s/	/z/	/ʃ/	/ʒ/			/ʧ/	/ʤ/
6	Tajik	/s/	/z/	/ʃ/	/ʒ/			/ʧ/	/ʤ/
6	Nenets	/s/, /sʲ/				/ʦ/, /ʦʲ/	/ʣ/, /ʣʲ/		
5	Tuvan	/s/	/z/	/ʃ/	/ʒ/			/ʧ/	
5	Selkup	/s/, /sʷ/		/ʃʲ/				/ʧ/, /ʧʷ/	
10 languages		/s/ = 10 /s+/ = 4	/z/ = 8 /z+/ = 2	/ʃ/ = 8 /ʃ+/ = 1	/ʒ/ = 8	/ʦ/ = 4 /ʦ+/ = 3	/ʣ/ = 3 /ʣ+/ = 3	/ʧ/ = 8 /ʧ+/ = 1	/ʤ/ = 6

Results

- In Asia, low scores are typical features of the North and the East.
- Among the languages which do not reach the average, none is located in the more westerly/southerly regions.
- Middle Asian languages tend to score high.
- The isoglosses which connect high scorers in both continent do not cross the Ural, but take the detour via the Caucasus.
- By and large, the implications identified in Europe are replicated in Asia.

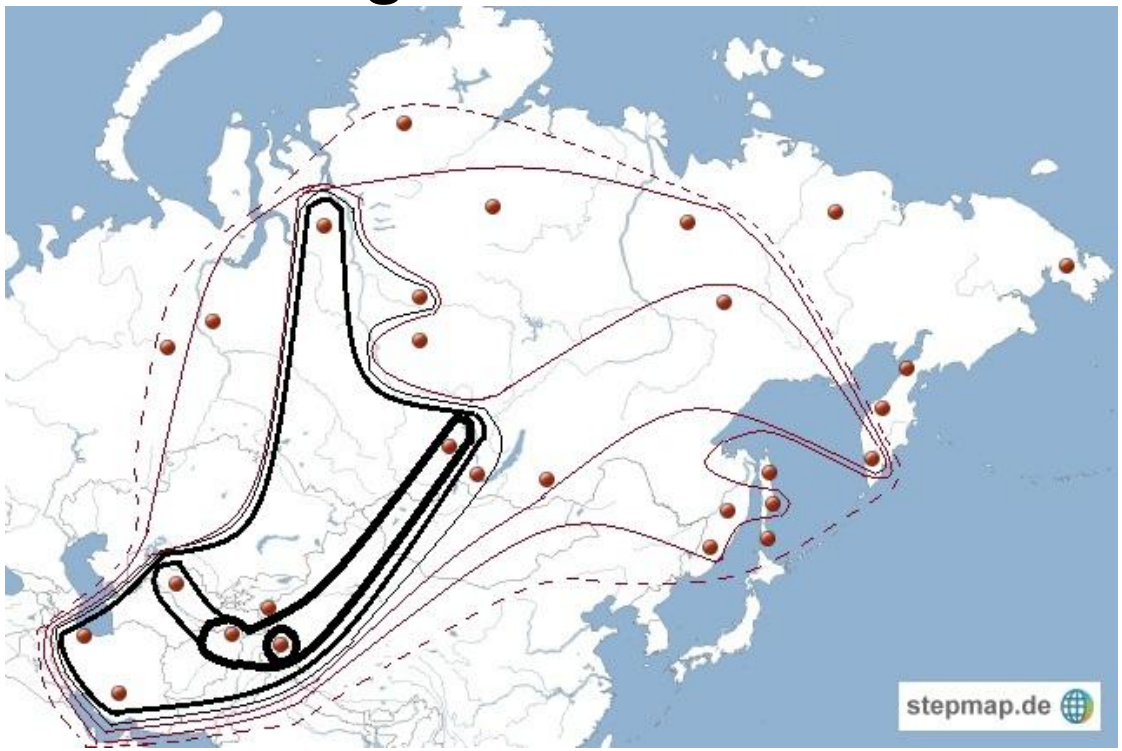
Isoglosses in Asia



Red dots: above average

Blue dots: below average

Isoglosses in Asia



Detailed distribution of languages with 12, 8, 7, 6, 5, 4, 3, 2, and less phonemes

Conclusions

We have shown that areal phonology is not only feasible but also highly interesting.

The combination of quantitative and qualitative methods allows us to draw isoglosses which join languages of divers genetic and typological background to form areas of varying size.

Encouraged by these findings, we will continue to work on a phonological atlas of Europe and its immediate neighbours.

Thank you for your kind attention!

Merci beaucoup de votre attention!