



# Mongolian

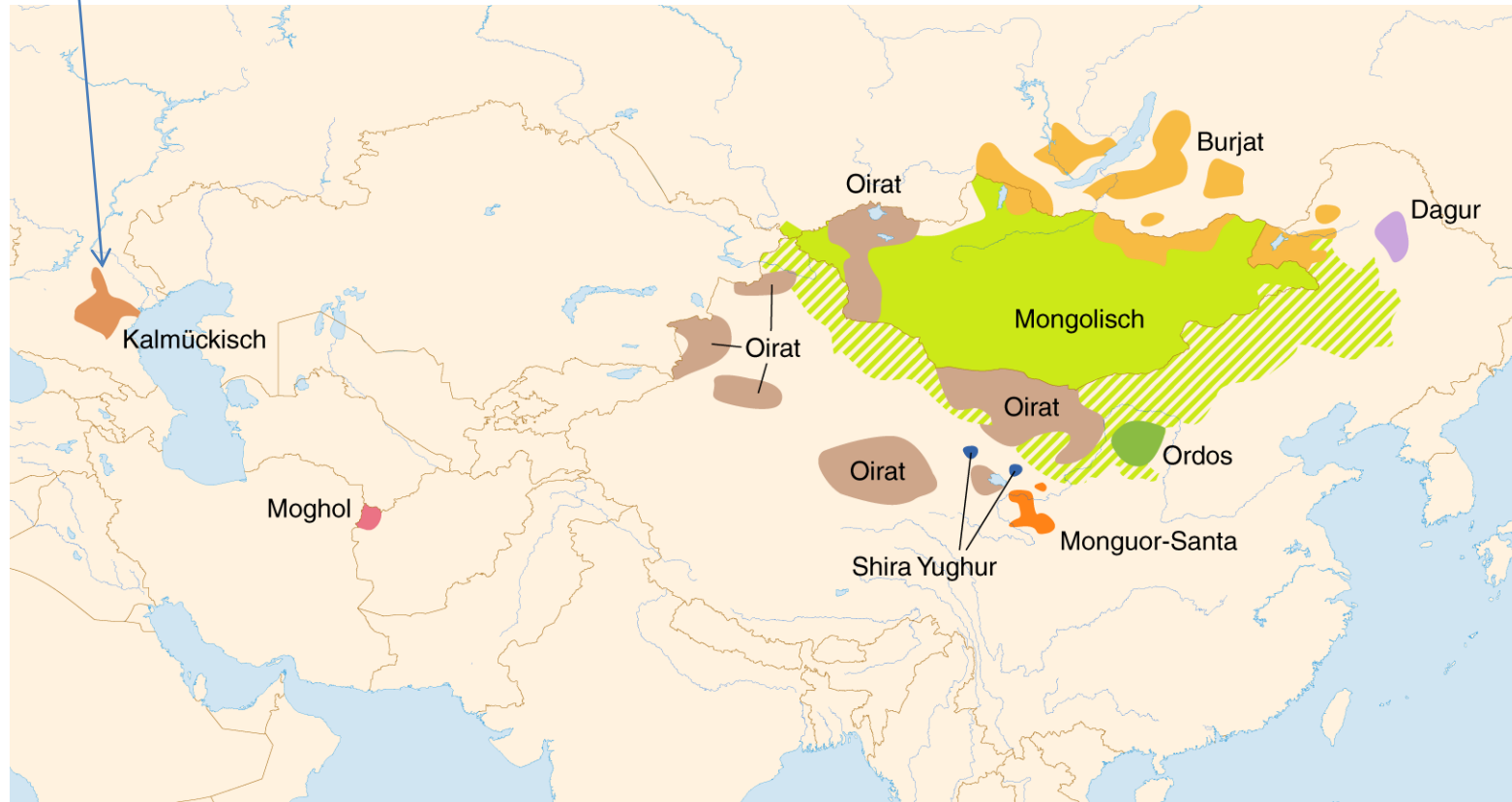
Монгол хэл

ᠮᠣᠩᠭᠣᠯ ᠬᠡᠯ

Lara Martin

Kalmykia

# Mongolic Languages



Altaic family

**Mongolisch/ Khalkha:** ~5 million speakers

(Numbers decreasing in China)

# Orthography

- Traditional script (Galik)
- Feb. 1, 1941 – March 25, 1941: Latin
- 1941 – 1994: Cyrillic
- 1994 – Now: readopt traditional, but still Cyrillic



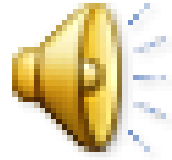
First Bogd Gegen - Zanabazar

# Linguistic Information

- 8-9 cases
  - Lots of suffixes
- Max syllable: CVVCCC
- No short vowels at the end of a syllable
- Plural-marking is always optional
- SOV
- 7 vowels, short & long

		Labial		Dental		Palatal	Velar		Uvular
		plain	pal.	plain	pal.		pal.	plain	
<b>Nasal</b>		m	mi	n	ni			ŋ	
<b>Plosive</b>	voiceless aspirated	(p <sup>h</sup> )	(pi <sup>h</sup> )	t <sup>h</sup>	ti <sup>h</sup>		(k <sup>h</sup> )	(k <sup>h</sup> )	
	voiceless	p	pi	t	ti				
	voiced						gi	g	g
<b>Affricate</b>	voiceless aspirated			tsh		tj <sup>h</sup>			
	voiceless			ts		tʃ			
<b>Fricative</b>	central	(f)		s		ʃ	xi	x	
	lateral			ʃ	ʃi				
<b>Trill</b>				r	ri				
<b>Approximant</b>		w	wi			j			

# A Taste



Хүн бүр төрж мэндлэхдээ эрх чөлөөтэй, адилхан нэр төртэй, ижил эрхтэй байдаг. Оюун ухаан нандин чанар заяасан хүн гэгч өөр хоорондоо ахан дүүгийн үзэл санаагаар харьцах учиртай.

Khün бүр törzh mendlekhee erkh čölöötei, adilkhan ner törtei, izhil erkhtei baidag. Oyuun ukhaan nandin čanar zayaasan khün gegč öör khoodondoo akhan düügiin üzel sanaagaar khar'tsakh učirtai.

*All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.*

# Machine Translation

ter hel-eh-güj-geer üün-ijg bič-sen šüü  
s/he without:saying it-accusative write-perfect particle  
*'s/he wrote it without saying [so] [i.e. without saying that s/he  
would do so, or that s/he had done so], I can assure you.'*

So many cases! → problem with POS tagging and  
lemmatization (Chimeddorj & Fujii, '12)

# (Limited) Resources

- Google has it, but no synthesis
- Automatically created Chinese-Mongolia Corpora (Zhu et al., '11)
- 5-million-word corpus using TEI XML tagging (Jaimai & Chimeddorj, '08)
- Mongolian-English dictionary made from limited resources (Sornlertlamvanich et al., '08)
- Hand-crafted rules that extracted loan words from Japanese (Khaltar et al., '06)

**БАЯРЛАЛАА!**



# Sources

- Mongolia map from [http://en.wikipedia.org/wiki/Mongolic\\_languages](http://en.wikipedia.org/wiki/Mongolic_languages)
- Flag image from [http://3.bp.blogspot.com/-xnK4IWu\\_xnE/UnfdiwD3acl/AAAAAAAAAJfM/u8ZDietTGl/s1600/three+flags.jpg](http://3.bp.blogspot.com/-xnK4IWu_xnE/UnfdiwD3acl/AAAAAAAAAJfM/u8ZDietTGl/s1600/three+flags.jpg)
- Population fact & reading (*Article 1 of the Universal Declaration of Human Rights*) from <http://www.omniglot.com/writing/mongolian.htm>
- Consonant chart from [http://en.wikipedia.org/wiki/Mongolian\\_language](http://en.wikipedia.org/wiki/Mongolian_language)
- <http://www.omniglot.com/writing/mongolian.htm>
- <http://www.languagesgulper.com/eng/Mongolian.html>