

Theoretical linguistics meets Turkic languages
A unification of inter- and cross-linguistic patterns

Jonathan North Washington

12 May 2007

1 *Introduction*

2 *Background*

- Turkic languages
- Study of linguistics
- Turkic languages in light of linguistics

3 *The data*

- Allomorphy
- Environment?
- Consistency?

4 *The theory*

- Overview
- Application to the data
- Typology

5 *Conclusion*

What this talk is about

- My thesis work? Not really
- Turkic languages
- Linguistics
- What they do for each other
- Why this is useful

What you should get out of this talk

- What are Turkic languages? (And what are they like?)
- What role does (theoretical) linguistics play?
- Why are Turkic languages interesting to linguists?
- What am I writing my thesis about?
- **Why do I enjoy working with all of this?**

Turkic languages

Facts

Turkic languages are:

- Spoken by about 1/4 billion people today
- In use all across Eurasia — especially Central Asia
- Official national languages of 6 countries
- Official status in numerous regions of Russia and China
- “Medium density” languages
 - ignored by linguists
 - ...and everyone else

What's linguistics all about?

- What?
 - “Scientific study of language”
- Why?
 - “What comprises the human capacity for language?”
- How?
 - Collect data
 - Pick out patterns in a given language
 - Pick out patterns across languages
 - Develop formalisms to unify the patterns
 - Develop a theoretical framework
 - ... which explains cognitive/psychological reality

How about phonology?

- Phonology
 - “Scientific study of sound systems and patterns”
 - Sub-branch of linguistics
 - Same goals and basic principles
 - Current dominant theory: “Optimality Theory”

Introduction to Phonology

Using Turkic-language data

English		Tatar
'hands'	[z]	qull <u>ar</u>
'apples'	[z]	ålm <u>ar</u>
'girls'	[z]	qız <u>lar</u>
'lakes'	[s]	kü <u>ll</u> är
'languages'	[t̪z]	tell <u>är</u>
'hearts'	[s]	yör <u>ä</u> kl <u>är</u>

Why are Turkic languages interesting?

- Why are Turkic languages [potentially] interesting [to linguists]?
 - Allomorphy (e.g., vowel harmony) ...and lots of it
 - **Cross-language variation in allomorphy**

Turkic Plural Morphemes

How many forms?

- Uzbek: -lar
- Turkish: -lar, -ler
- Tatar: -lar, -lär, -nar, -när
- Khakas: -lar, -ler, -nar, -ner, -tar, -ter
- Kazakh: -lar, -ler, -dar, -der, -tar, -ter
- Tuvan: -lar, -ler, -nar, -ner, -dar, -der, -tar, -ter
- Kyrgyz:
-lar, -ler, -lor, -lör, -dar, -der, -dor, -dör, -tar, -ter, -tor, -tör
- Yakut:
-lar, -ler, -lor, -lör, -nar, -ner, -nor, -nör, -dar, -der, -dor, -dör,
-tar, -ter, -tor, -tör

Turkic Plural Morphemes

Are there patterns?

	V	y/w	r	l	m/n/ŋ	z(etc)	s(etc)
Uzbek	V.lar	y.lar	r.lar	l.lar	n.lar	z.lar	s.lar
Tatar	V.lAr	y.lAr	r.lAr	l.lAr	n.nAr	z.lAr	s.lAr
Kumyk	V.lAr	y.lAr	r.lAr	l.lAr	l.lAr	z.lAr	s.lAr
Turkish	V.lAr	y.lAr	l.lAr	l.lAr	n.nAr	z.lAr	s.lAr
Khakas	V.lAr	y.lAr	r.lAr	l.lAr	n.nAr	—	s.tAr
Tuvan	V.lAr	y.lAr	r.lAr	l.dAr	n.nAr	—	s.tAr
Kazakh	V.lAr	y.lAr	r.lAr	l.dAr	n.dAr	z.dAr	s.tAr
Yakut	V.lAr	y.dAr	r.dAr	l.lAr	n.nAr	—	s.tAr
Bashkir	V.lAr	y.ðAr	r.ðAr	l.dAr	n.dAr	ð.ðAr	s.tAr

Optimality Theory

- What is it?
 - A theory of phonology
 - i.e., attempts to explain patterns found in sound systems
- What it strives for:
 - Tries to explain everything that languages do
 - (i.e., all existing sound systems)
 - Tries to make it impossible to describe things that languages don't do

Optimality Theory

The basic principles

- One input form (stored in brain)
- Many possible output forms (infinite)
- “Constraints” limit output forms
 - Faithfulness constraints: “be like input form”
 - Markedness constraints: “don’t break universal principles”
- All constraints may be violated
- Constraints are ranked
- One “optimal” output form is selected (what you say)
- Chosen based on not violating higher-ranked constraints

Constraints

- ICC[X] — Two adjacent sounds must be identical in way X
- SyllCon — Don't allow rising sonority across syllable boundaries
- SCS (Syllable Contact Slope) — Have falling sonority across syllable boundaries
- FaithOnset — Sounds at the beginning of syllables don't change
- FaithStem — Sounds in a word's stem don't change
- ID[son] — The sonority of a segment must not be lost

Example

Turkish

- ID[son], ICC[nas] » SyllCon, SCS, ICC[lat] » FaithStem » FaithOnset

	/n+l/	ID[son]	ICC[nas]	SyllCon	SCS	ICC[lat]	FStem	FOnset
☞ a.	n.n			*				*
b.	l.l			*			*!	
c.	n.l		*!	*	*	*		
d.	n.d	*!	*					*

Typology

- Different ranking per language
- Minute differences in ranking shown minute differences in systems
- For example:

	V	y/w	r	l	m/n/ŋ	z(etc)	s(etc)
Tatar	V.lAr	y.lAr	r.lAr	l.lAr	n.nAr	z.lAr	s.lAr
Kumyk	V.lAr	y.lAr	r.lAr	l.lAr	l.lAr	z.lAr	s.lAr

- Tatar: ICC[nas] » ICC[lat] » Faith
- Kumyk: ICC[lat] » ICC[nas] » Faith
- The rest?

What you've learned

A taste of:

- The structure of Turkic languages
- Linguistic methodology
- Phonology
- How phonology accounts for differences in related languages
- Why what I do is cool

Conclusion

Thank you :)