Gradient opacity in Uyghur backness harmony

Connor Mayer
University of California, Los Angeles
AMP 2020
connormayer@ucla.edu | www.connormayer.com

1. Background

Phonological opacity (Kiparsky 1973)
Assume rule of the form A → B / C_D. This rule is *opaque* if there are surface forms with either:

- A in environment C_D (underapplication)
- A → B in environments other than C_D (overapplication)

Opacity may pose learning difficulties (Kiparsky 1973)

Many theoretical battles have been fought over opacity

- Baković (2011) shows no theory really captures all its forms

This project:

- Examines an opaque pattern in Uyghur (Turkic: China) based on a large-scale corpus study
- Suggests this opaque pattern is learned based on lexically-specific constraints as function of frequency

2. Uyghur Backness Harmony

Basic pattern: Sounds in suffixes must agree in backness with final vowel in stem.

tʏr-dæ/*-dɑ ‘type-LOC’ pʊl-ʁɑ/*-gæ ‘money-DAT’
mʊnbær-gæ/*-ʁɑ ‘podium-DAT’ ætrɑp-tɑ/*-tæ ‘surroundings-LOC’

The vowels /i e/ are *transparent*.

mæstʃit-tæ /*-tɑ ‘mosque-LOC’ mømin-gæ/*-ʁɑ ‘believer-DAT’
student-lɑr/*-lær ‘student-PL’ amil-ʁɑ/*-gæ ‘element-DAT’

If a stem contains no harmonizing elements, it is lexically specified for backness, with a tendency towards back suffixes.

biz-gæ/*-ʁɑ ‘us-DAT’ welisipit-lær/*-lɑr ‘bicycle-PL’
sir-lɑr/-*lær ‘secret-PL’ hejt-tɑ/*-tæ ‘festival-LOC’

3. Vowel raising

The vowels /æ ɑ/ raise to [i] in medial, open syllables.

bɑlɑ	‘child’	bɑli-lar	‘child-PL’
qɑrɑ-ʃ	‘look-GER’	qɑri-di	‘look-3.SG.PAST’
mewæ	‘fruit’	mewi-si	‘fruit-3.SG.POS’
søzlæ-ʃ	‘talk-GER’	søzli-di	‘talk-3.SG.PAST’

Certain words and morphological constructions resist this raising.

hawɑ	‘weather’	hawɑ-si	‘weather-3.POS’
sæwæb	‘reason’	sæwæb-i	‘reason-3.POS’

/æ/ is more likely to raise than /ɑ/

Vowel raising produces potential opacity in backness harmony because it neutralizes /æ/ and /ɑ/ to transparent [i].

Two possible outcomes for underlying /ɑʁinæ-GA/ ‘friend-DAT”
Opaque harmony: [ɑʁini-gæ] **Surface harmony**: [ɑʁini-ʁɑ]

What do we see?

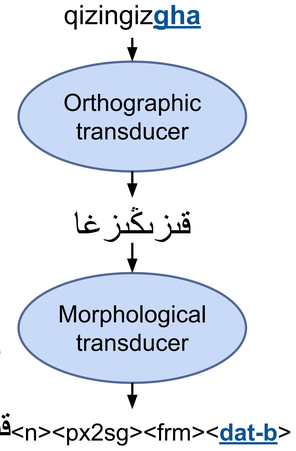
4. Corpus methodology

Examined raising patterns from corpora constructed from two online Uyghur newspapers (total of about 15 million words).

- Uyghur orthography transparently reflects harmony and raising

An automated morphological parser was used to parse words (Washington et al. to appear)

- Maps from orthographic forms to stem + morphological tags.
- Modified to detect suffix backness as well as stem and morpheme identity.
- Text processing done on stem and surface forms to identify raising and opaque harmonization



5. Corpus results

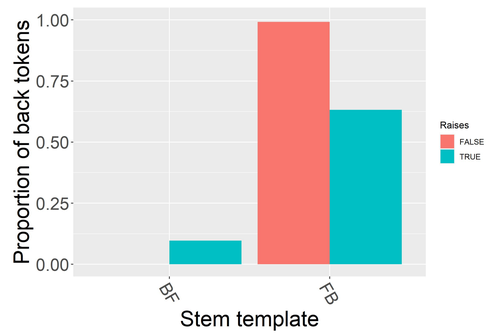
215 stems had the necessary structure to produce opacity.

- BF stems (n=181): e.g. /ɑdæt/ ‘custom’, /sijɑsæt/ ‘politics’
- FB stems: (n=34): e.g. /ætrɑp/ ‘area’, /æhwɑl/ ‘condition’

183 display raising:

- BF stems (n=177)
- FB stems (n=6)

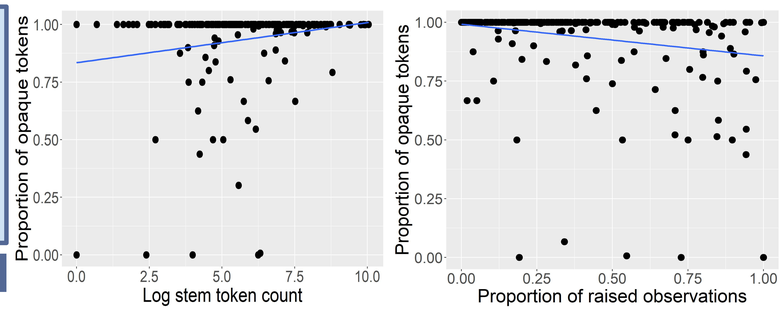
Raised forms are generally opaque, but a portion of these (n=87) vary in whether they display surface harmony or opaque harmony.



e.g., /ɑhɑlæ/ ‘population’

- Opaque harmony in 80% of cases: [ɑhɑli-lær-gæ]
- Surface harmony in 20% of cases: [ɑhɑli-lɑr-ʁɑ]

Linear regression shows that log word frequency and proportion of unraised tokens are significant predictors of opaque harmony.



6. Phonological analysis

Model using maximum entropy optimality theory with paradigm uniformity constraints (e.g. Steriade 2000).

- PU constraint: harmonize as the unraised form does
- Weights for PU learned as function of exposure to base

/ɑhɑlæ-lAr/	Predicted freq.	VAgree-[back] w=3.45	PU(back, ɑhɑlæ) w=4.96	PU(back, ɑʁinæ) w=15.27	DoRaising w=16.98
ɑhɑli-lær	**0.8**	1			
ɑhɑli-lɑr	**0.2**		1		
ɑhɑlæ-lær	**0**				1
ɑhɑlæ-lɑr	**0**	1	1		1
/ɑʁinæ-lAr/					
ɑʁini-lær	**1**	1			
ɑʁini-lɑr	**0**			1	
ɑʁinæ-lær	**0**			1	1
ɑʁinæ-lɑr	**0**	1			1

7. Discussion/Future Directions

Gradient opacity indicates that opacity is not learned as a completely productive process.

- Gradience captured by lexically-specific PU constraints that are weighted as function of base exposure

Is this indicative of language change?

- Raising is relatively new. Could surface harmony become the norm?

Corpora and computational methods provide greater empirical support to phonological theory.

Future work: Explicit learning model!

Selected References

Baković, E. (2011). Opacity and ordering. In J.A. Goldsmith, J. Riggle, and A.C. Yu (Eds.), *The Handbook of Phonological Theory,* 2e. London: Wiley-Blackwell, 40-67. **Kiparsky, P. (1973).** Abstractness, opacity, and global rules. In O. Fujimura (Ed.), *Three dimensions of linguistic theory*. Tokyo: TEC, 57-86. **Steriade, D. (2000).** Paradigm uniformity and the phonetics-phonology boundary. In M.B. Broe and J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V*. Cambridge: Cambridge University Press, 313-334. **Washington, J., Salimzianov, I., Tyers, F.M., Gökırmak, M., Ivanova, S., Kuyrukçu, O. (to appear)**. Free/open-source technologies for Turkic languages developed in the Apertium project. In *Proceedings of the International Conference on Turkic Language Processing (TURKLANG 2019)*.