A large-scale corpus study of phonological opacity in Uyghur

Connor Mayer

Abstract

This paper examines a case of phonological opacity in Uyghur that results from an interaction between backness harmony and a vowel reduction process that converts harmonic vowels into transparent vowels. A large-scale written corpus study shows that although opaque harmony with the underlying form of a reduced vowel is the dominant pattern, cases of surface-true harmony also exist. The rate of surface- true harmony varies across roots and is correlated with a number of factors, including root frequency and morphological composition. These data pose problems for standard accounts of opacity, which do not predict such variation. I propose an analysis where variation emerges as the result of conflict between lexical knowledge of the harmonizing class of a root and sensitivity to surface phonotactic constraints. This analysis is implemented in a strictly parallel model using lexically-indexed constraints, underscoring theoretical connections between opacity and exceptionality.

1 Introduction

This paper presents an exploratory written corpus study of a phonological pattern in Uyghur (Turkic: China) whereby a vowel reduction process converts harmonic vowels into transparent vowels, rendering the vowel harmony pattern opaque. Opaque patterns are of interest to phonological theory because of the challenges they pose for learning and for certain classes of phonological models. In particular, opacity has been a perpetual difficulty for strictly parallel phonological models such as classical Optimality Theory (Prince & Smolensky, 1993/2004), which do not straightforwardly predict its existence.

The paper has two primary goals. The first is to present new empirical data on a case of phonological opacity in an understudied language. Uyghur provides a valuable opportunity to study an opaque pattern at scale using text data, because both vowel reduction and harmony are reflected orthographically. The data show that although opaque harmony is the majority pattern, there is variability in rates of opacity both within and between roots: roots sometimes display surface-true harmony. Rates of opacity are correlated with factors like root frequency and the presence of certain derivational suffixes.

The second is to highlight challenges these data pose for standard theories of opacity, which do not typically predict such variability. I will outline an analysis that is compatible with the data, where observed variation emerges as the result of conflict between lexical knowledge of the harmonizing class of a root, and sensitivity to surface phonotactic constraints. In addition to accounting for the various correlates of opacity observed in the corpus, this analysis can be implemented in a strictly parallel model using lexically-indexed constraints. This aligns with previous proposals that lexical factors play a fundamental role in opaque phenomena, and underscores the theoretical connections between opacity and exceptionality (Łubowicz, 2003; Sanders, 2003; Mielke et al., 2003; Pater, 2014; Nazarov, 2019).

I thank Bruce Hayes, Kie Zuraw, Adam McCollum, Tim Hunter, Eric Bakovic, Travis Major, Canaan Breiss, and the attendees of AMP 2020, mfm 2022, UCLA Phonology Seminar, and the Berkeley Phorum for their feedback at various stages of this project; Jonathan Washington for his help with the Apertium transducer; Tyler Carson, Daniela Zokaeim, and Rutvik Gandhasri for their work on the webscrapers; and Gülnar Eziz, Mahire Yakup, Gulnisa Nazarova, and other members of the Uyghur community for sharing their language and culture with me. This research was supported by the Social Sciences and Humanities Research Council of Canada.

Section 2 provides background on phonological opacity. Section 3 describes the processes of backness harmony and vowel reduction in Uyghur, and how they interact to produce opacity. Section 4 presents the results of a large written corpus study looking at rates of opacity. Section 5 discusses why these results are challenging for standard theories of opacity. Section 6 presents an analysis of the corpus data. Finally, Section 7 discusses implications for theories of opacity in general, and how future research might proceed.

2 Phonological opacity

Opacity is a type of structured phonological exceptionality. Kiparsky (1971, 1973) defines it as follows:

(1) Assume a phonological rule $\mathbb{R}$: A → B / C_D. $\mathbb{R}$ is *opaque* iff there are surface forms with either:

a. *A* in the environment *C_ D* (underapplication opacity)

b. *B* derived from *A* in environments other than *C_ D* (overapplication opacity)

Opacity arises when either a conditioned alternation appears not to occur despite its conditions being met, or appears to occur when its conditions have not been met.

Kiparsky (1973) associated opacity of types (1a) and (1b) with *counterfeeding* and *counterbleeding* rule orders respectively. In counterfeeding opacity, the structural conditions for rule $\mathbb{R}$ to apply are created by a different rule $\mathbb{P}$ that applies after $\mathbb{R}$: hence the necessary conditions are not met when $\mathbb{R}$ applies. Changing the rule ordering such that $\mathbb{P}$ applied before $\mathbb{R}$ would produce a *feeding* order where $\mathbb{R}$ applies transparently to the conditioning environment produced by $\mathbb{P}$.

In counterbleeding opacity, the conditions for $\mathbb{R}$ are met when it applies, but are subsequently altered by a different rule $\mathbb{Q}$ that applies after $\mathbb{R}$. Changing the rule ordering such that $\mathbb{Q}$ applies before $\mathbb{R}$ would produce a *bleeding* order where $\mathbb{R}$ transparently fails to apply because $\mathbb{Q}$ removes its conditioning environment.

More recently, interest in opacity has stemmed from debates on the merits of serial models such as *SPE*-style rules (e.g., Chomsky & Halle, 1968) vs. parallel models such as Optimality Theory (e.g., Prince & Smolensky, 1993/2004). Parallel models have difficulty correctly predicting cases of counterbleeding opacity, which generally produce faithfulness violations with no corresponding markedness repairs to motivate them. They also have difficulty with most types of counterfeeding opacity, which fail to repair a markedness violation whose repair is evident elsewhere.

A number of theoretical mechanisms have been proposed to handle these cases, including mechanisms that incorporate some degree of serialism into OT, such as sympathy (McCarthy, 1999), stratal OT (Kiparsky, 2000; Bermúdez-Otero, 2003; Nazarov & Pater, 2017; Bermúdez-Otero, 2018), candidate chain theory (McCarthy, 2007), and serial markedness reduction (Jarosz, 2014), as well as purely parallel mechanisms, such as constraint conjunction (Kirchner, 1996), paradigm uniformity (Steriade, 2000), language-specific constraints (Pater, 2014), or indexed constraints (Nazarov, 2019, 2020, 2021). The need for such bespoke mechanisms has been seen as a point in favor of serial models, which handle these cases of opacity without issue (e.g., Vaux, 2008).

Although counterfeeding and counterbleeding orderings are the best known configurations that result in opacity, the typologies of opacity enumerated in Baković (2007, 2011) and Baković and Blumenfeld (2019) show that these orderings are neither sufficient nor necessary conditions for opacity. They identify a number of cases of overapplication opacity that are not predicted by SPE-style rule ordering, and some which are *only* able to be described by parallel models. Thus the characterization of opacity as a unique challenge for parallel models is a simplification, though accurate in broad strokes.

In light of the lack of a unified account of opacity from either serial or parallel theories, Baković (2007) suggests that the field focus on Kiparsky's claim that opaque patterns are more difficult to learn than transparent ones. The basic motivation for this claim is that phonological processes that interact in an opaque fashion make generalization about those processes difficult: opaque forms constitute exceptions to otherwise robust generalizations. Kiparsky (1971) supports this claim by presenting a number of cases of historical change where an opaque process is reanalyzed as a transparent one.

Subsequent research has presented evidence that opaque processes are learned as phonemic contrasts or lexicalized patterns rather than productive rules (e.g., Hooper/Bybee, 1976; Mielke et al., 2003; Sanders, 2003; Sumner, 2003; Zhang, 2019; Bowers, 2019), though evidence also exists that some opaque processes are applied productively in language games and other contexts (e.g., Donegan & Stampe, 1979; Al-Mozainy, 1981; Vaux, 2011).

3 Opacity in Uyghur backness harmony

Uyghur is a southeastern Turkic language spoken by over 12 million people in the Xinjiang Uyghur Autonomous Region in the People's Republic of China, neighboring countries such as Kazakhstan and Kyrgyzstan, and various diasporic communities (Engesæth et al., 2009/2010; Nazarova & Niyaz, 2013). It has SOV word order with highly agglutinative morphology that is almost exclusively suffixing.

The opaque phenomenon under consideration arises from the interaction of two independent processes: backness harmony and vowel reduction. I will introduce these processes separately before demonstrating how their interaction leads to opacity.

The reader is referred to Mayer et al. (accepted) for a more detailed description of Uyghur phonology.

3.1 Segments involved in backness harmony

Like most Turkic languages, Uyghur has backness harmony. Native Turkic roots tend to be harmonic, but extensive borrowing has resulted in a high degree of root-internal disharmony. As a consequence, backness harmony is most evident as a morphophonological process, where, broadly speaking, segments in many suffixes must agree in backness with the roots they attach to (e.g., Lindblad, 1990; Hahn, 1991a, 1991b; Engesæth et al., 2009/2010; Abdulla et al., 2010).

	Front		Back	
	Unrounded	Round	Unrounded	Round
High	i	$\underline{\text{y}}$		$\underline{\text{u}}$
Mid	e	$\underline{\text{ø}}$		$\underline{\text{o}}$
Low	$\underline{\text{æ}}$		$\underline{\text{ɑ}}$	

Table 1: The Uyghur vowel system. Harmonizing vowels are underlined.

	Front	Back
Voiceless	k	q
Voiced	g	ʁ

Table 2: Harmonizing Uyghur consonants

Segments that participate in backness harmony are shown in Tables 1 and 2. The underlined vowels in Table 1 serve as harmony triggers (that is, they determine the backness of suffixes attached to roots containing them), while the non-bolded vowels are transparent to harmony. The harmonizing consonants may also serve as harmony triggers, though they tend to be weaker than vowels. This paper will focus primarily on harmony driven by vowels. In addition to serving as triggers of harmony, the harmonizing vowels and consonants both emerge as the outcome of harmony in harmonizing suffixes.

3.2 A description of Uyghur backness harmony

The examples of harmony below include the locative suffix /-DA/ (surface forms: [-tɑ], [-dɑ], [-tæ], [-dæ]), the plural suffix /-lAr/ (surface forms: [-lɑr], [-lær]), or the dative suffix /-GA/ (surface forms: [-qɑ], [-ʁɑ], [-kæ], [-gæ]). I assume that /A/ is unspecified for the feature [back], /D/ for [voice], and /G/ for either (Archangeli, 1988). Voicing alternations in the initial segment are caused by voice assimilation, and are orthogonal to harmony.

The basic characterization of backness harmony is that suffixes must agree in backness with the final front /y ø æ/ or back /u o ɑ/ harmonizing root vowel.

(2) *Simple front harmonizing forms*
tyr-dæ 'type-LOC'
pæn-lær 'science-PL'
munbær-gæ 'podium-DAT'

(3) *Simple back harmonizing forms*
pul-ʁɑ 'money-DAT'
top-qɑ 'ball-DAT'
ætrɑp-tɑ 'surroundings-LOC'

The vowels /i e/ are *transparent* to harmony. They do not serve as harmony triggers, but allow the harmonic value of preceding segments to "pass through" them.[1]

(4) *Front roots with transparent vowels*
mæstʃit-tæ 'mosque-LOC'
ymid-lær 'hope-PL'
mømin-gæ 'believer-DAT'

(5) *Back roots with transparent vowels*
student-lɑr 'student-PL'
uniwersitet-tɑ 'university-LOC'
ɑmil-ʁɑ 'element-DAT'

Roots without any harmonizing segments typically take back suffixes, but some take front suffixes (see McCollum, 2021; Mayer et al., 2022).

(6) *Neutral roots that take back suffixes*
sir-lɑr 'secret-PL'
din-ʁɑ 'religion-DAT'
hejt-tɑ 'festival-LOC'
peʔil-lɑr 'verb-PL'
tip-qɑ 'type-DAT'

(7) *Neutral roots that take front suffixes*
biz-gæ 'us-DAT'
bilim-gæ 'knowledge-DAT'
welisipit-lær 'bicycle-PL'

[1] Jonathan Washington (p.c.) notes that there is uncertainty about the status of /e/ as a transparent vowel; it typically shows up in relevant environments only in loanwords, which often have idiosyncratic harmonizing behavior. The apparent transparency may be due to lexical rather than phonological factors.

Back suffixes appear to be the unmarked class in Uyghur. There has been a general diachronic shift in the population of neutral roots towards back suffixes (Lindblad, 1990), and recent loanwords that lack strong harmonizing segments typically take back suffixes.

3.3 Vowel reduction

The second process that contributes to opacity in the Uyghur harmony system is *vowel reduction* or *raising*, which raises the low vowels /ɑ æ/ to [i] in medial open syllables in derived environments.

(8) /ɑ/ *vowel reduction*

bɑlɑ	'child'	bɑli-ni	'child-ACC'
ɑpɑ	'mom'	ɑpi-si	'mom-3.POSS'
ɑŋlɑ-ʃ	'listen-GER'	ɑŋli-di	'listen-3.SG.PST'
qɑrɑ-ʃ	'look-GER'	qɑri-di	'look-3.SG.PST'

(9) /æ/ *vowel reduction*

ɑpæt	'disaster'	ɑpit-i	'disaster-3.POSS'
mewæ	'fruit'	mewi-si	'fruit-3.POSS'
søzlæ-ʃ	'talk-GER'	søzli-di	'talk-3.SG.PST'
kyʧæ-ʃ	'strive-GER'	kyʧi-di	'strive-3.SG.PST'

The underlying form cannot in general be predicted from forms where vowel reduction could have applied, as many words have underlying /i/ in these positions, as in /tɑksi/ 'taxi' or /æsli/ 'origin'. Certain roots resist raising categorically; in the current paper we consider only roots that undergo raising.

3.4 Opaque interactions between backness harmony and vowel reduction

Vowel reduction has the potential to introduce opaque behavior into the vowel harmony system. Consider, for example, the root /ɑʁinæ/ 'friend'. The final vowel raises without exception:

(10) /ɑʁinæ-ni/ → [ɑʁini-ni] 'friend-ACC'

What happens for forms such as /ɑʁinæ-DA/ 'friend-LOC', where the vowel in the suffix must harmonize with the final vowel in the root? There are two possibilities: *opaque harmony* according to the underlying form of the root and *surface-true harmony* according to the raised form of the root.

A rule ordering where harmony precedes raising predicts the opaque form [ɑʁinidæ]:

(11) *Harmony precedes raising*

UR	/ɑʁinæ-DA/
Harmony	ɑʁinæ-dæ
Raising	ɑʁini-dæ
SR	[ɑʁini-dæ]

This opacity is precisely the kind that classical OT has difficulty accounting for: there is an explicit markedness violation (failure to harmonize), with no apparent motivation (cf. forms like /tɑksi-DA/ → [tɑksidɑ] 'taxi-LOC').

If raising instead precedes backness harmony, we would expect surface-true form [ɑʁini-dɑ]:

(12) *Raising precedes harmony*

UR	/ɑʁinæ-DA/
Raising	ɑʁini-DA
Harmony	ɑʁini-dɑ
SR	[ɑʁini-dɑ]

This surface-true harmony is straightforward to represent in both serial and parallel models.

3.5 Past work on opacity in Uyghur

Which of these patterns do we observe in Uyghur? Pedagogical materials do not generally discuss these cases in any detail, since roots that can generate opaque harmony are a relatively small slice of the lexicon. Those that do suggest that opacity is the correct outcome (e.g., Hahn, 1991b, Section 4.3.5). Hahn describes this in terms of roots falling into a particular 'harmonic category', with vowel reduction processes 'disguising' otherwise obvious clues to this category.

As is typical for opaque phenemona (Kiparsky, 1971, 1973; Mielke et al., 2003), the rule ordering that produces opaque harmony reflects the relative diachronic development of each process. Backness harmony is an ancient property of Turkic languages (e.g., Clauson, 1972), while raising appears to be a newer phenomenon in Uyghur: Chagatay, the closest direct ancestor to Uyghur, had no such raising process (Bodrogligeti, 2001). Opaque harmony thus maintains historical patterns of root backness at the cost of surface disharmony.

The exception to this is that certain suffixes in Uyghur, such as the diminutive /-ʧæ/ and the adjectival suffix /-ɑnæ/ have been described as displaying surface-true harmony as in raised forms (Hahn, 1991b; Vaux, 2000; Halle et al., 2000; Hall & Ozburn, 2018).[2]

(13) *Transparent harmony in raised* /-ʧæ/

	/bɑʁ-ʧæ-DA/	→	[b<u>ɑ</u>ʁʧid<u>ɑ</u>]	'park-DIM-LOC'
cf.	/bɑʁ-ʧæ-m-DA/	→	[bɑʁʧ<u>æ</u>md<u>æ</u>]	'park-DIM-1.SG.POSS-LOC'

Uyghur speakers I have worked with agree that opaque harmony is the prescriptively correct form. However, in addition to the suffixes above, speakers have identified certain forms where surface-true harmony is mandatory (e.g. /ærzɑn-i-GA/ → [ærzinigæ] 'cheap (ones)-3.POSS-DAT') or both surface and opaque harmony are acceptable (/æzɑn-i-GA/ → [æziniʁɑ]/[æzinigæ] 'call to prayer-3.POSS-DAT'). These observations of variability were one of the motivations for the corpus study presented below.

Theoretical work on the interaction between vowel reduction and harmony has claimed that there is an asymmetry between vowels (Vaux, 2000; Halle et al., 2000; Hall & Ozburn, 2018): raised /æ/ is opaque and continues to behave as a front vowel trigger, while raised /ɑ/ is transparent, behaving identically to underlying /i/. However, these claims have been based off only eight data points collected from a single speaker, and the empirical validity of these data is unclear (see Mayer, 2021, Section 3.3.7). One of the main goals of this paper is to put claims about opacity in Uyghur on stronger empirical footing.

In addition to being an important empirical question, a better understanding of this pattern is valuable from a theoretical perspective: opaque patterns such as the one in (11) are not predicted to exist by many strictly parallel phonological models, and indeed, the Uyghur pattern has been used to argue in favor of serial models (Vaux, 2000). Before we turn to the empirical data, the next section describes the problems opaque patterns like (11) pose for parallel models.

3.6 Modeling opacity in serial and parallel models

The kind of opacity shown in (11) is straightforward to represent in serial rule-based models: the rule or rules driving harmony are simply ordered before the rule that drives raising. An analysis under such a model is fundamentally identical to the derivation in (11).

[2] /-ʧæ/-final stems are sometimes repaired to be internally harmonic: e.g., the root /bɑʁʧæ/ 'park' (lit. 'orchard-DIM') is frequently produced as [bɑʁʧɑ].

This pattern poses challenges for an analysis in classical OT. I will assume a simple constraint that motivates vowel harmony, which are roughly a combination of the local and non-local AGREE constraints used by Hayes et al. (2009):

(14) VAGREEBACK: Assess one violation if the harmonizing vowel in a suffix does not match the backness of the final harmonizing vowel in the root.

The following constraints will drive raising:

(15) *UNREDUCED: Don't have low vowels in medial open syllables.

(16) ID[HEIGHT]: Don't change the height of segments in the input.

*UNREDUCED is shorthand for a more detailed analysis of the pressures that drive vowel reduction.

When relevant I will employ a constraint that prevents specified [back] values from being altered.

(17) ID[BACK]: Don't change the backness of segments in the input.

ID[BACK] prevents underlyingly specified vowels in roots and certain harmony-blocking suffixes from being altered. This constraint is not violated when a segment underspecified for backness in the input is assigned a backness value in the output, nor is it violated when /æ ɑ/ are raised to [i]: assuming that [i] is unspecified for backness, these processes violate DEP and MAX constraints respectively, which are low-ranked and omitted from the tableaux below.

/ɑʔilæ-lAr/	*UNREDUCED	VAGREE	ID[HEIGHT]
☹ a. ɑʔili-lær		*!	*
💣 b. ɑʔili-lɑr			*
c. ɑʔilæ-lær	*!		
d. ɑʔilæ-lɑr	*!	*	

Table 3: Failed tableau for [ɑʔili-lær] 'family-PL'. The sad face indicates the candidate that should have won, and the bomb indicates that candidate that did win.

These constraints allow Classical OT to derive only surface harmony, as shown in Table 3. The desired candidate [ɑʔili-lær] is harmonically bounded by the winning *[ɑʔili-lɑr], and so will never be the optimal candidate under any ranking.

An analysis using Stratal OT succeeds in capturing this opacity (Kiparsky, 2000; Bermúdez-Otero, 2003, 2018). Stratal OT divides the grammar into several strata (e.g., the stem, the word, the phrase) and assigns each of these levels a separate OT grammar, with differing constraint rankings. The outputs of lower strata serve as the inputs to higher strata. We can capture the opaque pattern here by proposing that vowel harmony occurs at a lower stratum than vowel reduction, as in Table 4.

I employ this formalism here because it is widely used in the contemporary literature, and because there is some evidence in Uyghur that raising applies across a larger domain than backness harmony, which suggests it occurs in a different stratum. For example, in phrases like *Adil Hesenge berdi* 'Adil gave it to Hesen', the dative *-ge* [-gæ] may raise to *-gi* [-gi] in rapid speech (Hahn, 1991b, p. 53).

Thus serial rule-based analyses predict opaque harmony straightforwardly, while strictly parallel analyses predict only surface-true harmony. Modifications to strictly parallel models that incorporate some degree of serialism also predict opaque harmony, though they differ in their attribution of the particular mechanism responsible for it.

Word stratum

/ɑʔilæ-lAr/	ID[BACK]	ID[HEIGHT]	VAGREE	*UNREDUCED
a. ɑʔili-lær		*!	*	
b. ɑʔili-lɑr		*!		
☞ c. ɑʔilæ-lær				*
d. ɑʔilæ-lɑr			*!	*

Phrase stratum

/ɑʔilæ-lær/	ID[BACK]	*UNREDUCED	VAGREE	ID[HEIGHT]
☞ a. ɑʔili-lær			*	*
b. ɑʔili-lɑr	*!			*
c. ɑʔilæ-lær		*!		
d. ɑʔilæ-lɑr	*!	*	*	

Table 4: Tableaux for the derivation of [aʔili-lær] 'family-PL' at the word stratum (top) and phrase stratum (bottom). At the word stratum, the constraint driving raising is ranked below its corresponding faithfulness constraint, meaning harmony applies but raising does not. The output from the word stratum serves as the input to the tableau for the phrase stratum. At this stratum, the constraint driving raising is now ranked above its corresponding faithfulness constraint, meaning raising can apply.

4 A corpus study of opacity in Uyghur backness harmony

In order to investigate the interaction of vowel reduction and backness harmony, I performed a corpus study using three large text corpora.[3].

Uyghur uses a number of different orthographies depending on where it is written (Perso-Arabic, Cyrillic, or Latin). In each of these, the alternations conditioned by the raising and harmony processes are represented orthographically. Hence text corpora allow us to gather large-scale empirical data on their interaction.

The first corpus was generated from the Radio Free Asia (RFA) Uyghur language website.[4] RFA is a US-sponsored non-profit news organization. The second was generated from the website of *Uyghur Awazi* (Uyghur Voice), an Uyghur-language newspaper published in Almaty, Kazakhstan.[5] The third was generated from *Uyghur Akadémiyisi* (Uyghur Academy), a legal research organization that publishes articles on Uyghur culture and politics.[6]

Corpora were generated from the websites using *web scrapers*: software that, given a starting URL, instructions for how to navigate between pages, and instructions for which information to retrieve from each page, can download content from all pages on a site, or multiple sites. Such programs are useful for generating corpora from publicly available internet resources, in formats that are useful to researchers.

There are separate web scrapers for the Uyghur Awazi,[7] RFA,[8] and Uyghur Academy[9] websites. These scrapers were written by undergraduate research assistants at UCLA and UCI.[10] in collaboration with the author.

A summary of the contents of each corpus is shown in Table 5.

[3]The code used in this paper can be found at `https://github.com/connormayer/uyghur`$_{c}orpora$

[4]`https://www.rfa.org/uyghur/`

[5]`http://uyguravazi.kazgazeta.kz/`

[6]`https://www.akademiye.org/ug/`

[7]`https://github.com/connormayer/uyghur`$_{t}ools/tree/master/uyghur_{a}wazi_{s}craper$

[8]`https://github.com/yzgncx/RFA-Scraper`

[9]`https://github.com/rgandhasri87/UyghurAcademyWebsiteSpider-UG`

[10]Thanks to Daniela Zokaeim, Tyler Carson, and Rutvik Gandhasri for their work on these programs.

Corpus	Article count	Word count	Date retrieved	Orthography
Uyghur Awazi	24,080	4,197,550	January 2020	Latin/Cyrillic
RFA	45,777	7,986,484	July 2022	Perso-Arabic
Uyghur Academy	2,635	2,423,519	July 2022	Perso-Arabic

Table 5: Summary of corpora.[11]

In addition to the contents of each article, the scrapers retrieved the author, the date, and the URL.

4.1 Parsing the corpora

In order to extract information about the interaction between backness harmony and vowel reduction from the corpus, I modified an existing Uyghur morphological transducer to detect the backness of suffix forms (`https://github.com/apertium/apertium-uig`; Littell et al., 2018; Washington et al., 2019). This transducer is part of Apertium, a free and open-source rule-based machine translation platform (`https://www.apertium.org`).

The transducer maps from surface forms to underlying analyses that consist of roots plus morphological tags indicating the backness of any harmonizing suffixes. For example, if the input is the surface form *qizingizgha* "to your daughter" the output analysis will be *qiz<n><px2sg><frm><dat-b>*. This indicates that the root is *qiz*, a noun <n>, and is suffixed with the 2nd person singular possessive marker in its formal form <px2sg><frm> followed by the dative suffix in its back form <dat-b>.

The output of the transducer was used to count the frequency of front or back suffixes for each root. Simple text processing comparing the parsed root and surface forms was used to detect whether vowel reduction occurred and to extract phonotactic properties of the roots and tokens.

Additional details of the transducer and data processing are presented in Appendix B.

4.2 Quantitative results

4.2.1 Comparing harmonic and disharmonic roots

Here I consider only tokens where (a) the final two harmonizing elements of the root are vowels; (b) the underlying final vowel in the root is either /æ/ or /ɑ/; (c) the final vowel undergoes raising; and (d) the raised vowel is followed by at least one harmonizing suffix.

Fig. 1 breaks down parsed roots into four classes according to their final two harmonizing vowels: two back vowels (BB; 774 roots; e.g., /bɑlɑ/ 'child'), a back vowel followed by a front vowel (BF; 311 roots; e.g., /ɑdæt/ 'custom'), a front vowel followed by a back vowel (FB; 7 roots; e.g., /ærzɑn/ 'cheap'), and two front vowels (FF; 528 roots; e.g., /sypæt/ 'quality'). The BF and FB classes have the potential to produce opaque harmony.

The FB class has very few roots and tokens compared to the others. Roots with this shape are relatively uncommon, and those that exist tend not to undergo raising.[12] Fig. 2 breaks down BF and FB roots by their rates of opaque harmony.

Because the final two vowels in the BB and FF roots agree in backness, opaque harmony and surface harmony predict the same results. These roots almost categorically take the expected suffix forms. The

[11] The earlier scrape date for the Uyghur Awazi corpus is due to a recent redesign. This redesign changed the structure of the site and pages, which requires modifications to the scraper. It was not possible to make these modifications in time for the submission of this paper.

[12] Three of these raising FB roots end in the derivational suffix /-dɑʃ/ '-mate', as in [xizmætdɑʃ] 'officemate'. This suffix may display idiosyncractic harmonizing behavior in a similar way to the suffixes discussed in the next section, but there is insufficient data to determine this.

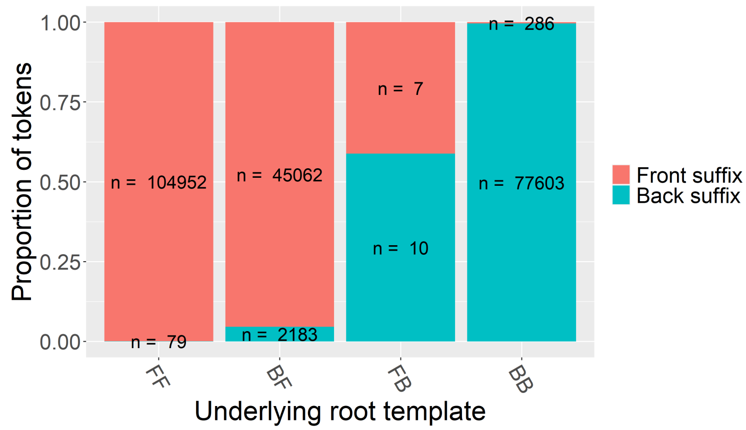

Figure 1: Suffix harmony choice in tokens where the final root vowel raises, broken down by root class. Token counts are overlaid on each category.

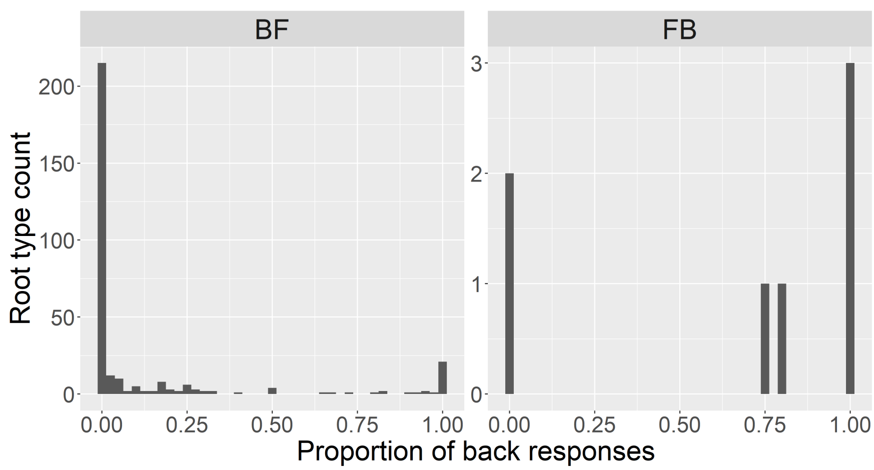

Figure 2: Histograms showing the distribution of rates of back suffix application in BF and FB roots .

disharmonic FB roots and BF roots behave similarly to BB and FF roots, respectively, but both show higher rates of surface-true harmony. Chi-squared tests show show significantly different rates of back suffix choice between BB and FB roots (χ^2 = 650.47; df = 1; $p < 0.0001$) and between FF and BF roots (χ^2 = 4597.6; df = 1; $p < 0.0001$). Thus the quality of the vowel preceding the raised vowel, and not just the underlying quality of the raised vowel, affects suffix choice: when the backness of the preceding vowel conflicts with the backness of the raised vowel, the suffix becomes more likely to agree with the preceding vowel.

4.2.2 Opacity in derivational suffixes

Recall that previous work on opacity in Uyghur has suggested that certain derivational suffixes like the dimunitive /-ʧæ/ and the adjectival suffix /-ɑnæ/ behave idiosyncratically, preferring surface-true harmony. Manual inspection of the corpus data revealed a similar pattern for the suffix /-nɑmæ/, meaning 'writings of' (e.g., *baburname* "The writings of Babur"). Fig. 3 breaks down harmony rates in the set of BF roots according to suffix.[13] The suffix /-ʧæ/ rarely displays opaque harmony, aligning with observations made in the literature. /-ɑnæ/ and /-nɑmæ/ display higher rates of opaque harmony, but not as high as the general population of BF roots.

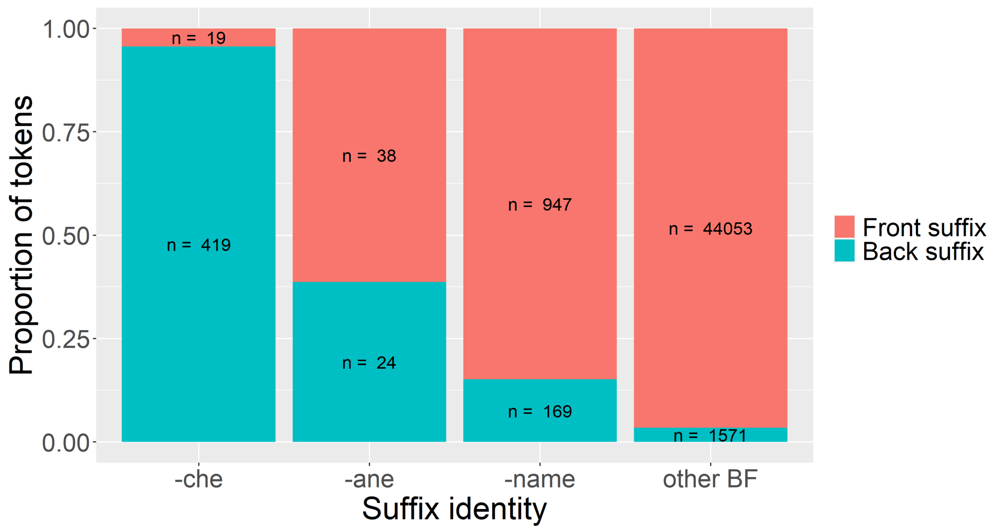

Figure 3: Suffix choice in raised BF roots broken down by root-final derivational suffix. 'Other BF' refers to BF roots that do not end in one of the three derivational suffixes. Token counts are overlaid on each category. The tokens of *-che* included here all have a preceding B vowel, as in /bɑʁʧæ/ 'park'.

4.3 Predicting opacity

In order to identify potential factors that contribute to rates of opacity, I fit a Bayesian mixed-effects logistic regression model to the set of raised BF and FB tokens with at least one harmonizing suffix attached (for discussion of the use of logistic regression in modeling categorical corpus data, see Speelman, 2014). The model was fit in R (R Core Team, 2017) using the `brms` package (Bürkner, 2017).

Bayesian models treat the statistical parameters of the model as random variables and use sampling techniques to estimate the posterior distribution over parameter values given the observed data and prior beliefs about credible parameter values (see, e.g., Kruschke, 2014; Nicenboim & Vasishth, 2016). The model here uses the default, weakly-informative priors.

I report two common summary statistics. The median, or most credible, value of each parameter, and the 95% Credible Interval (95% CI): the range in which the central 95% of the sampled parameter values fall. A 95% CI that does not include 0 is interpreted as meaningful, since it indicates that the directionality of the effect is highly credible. See Appendix C for more detail on the model, as well as an analysis of the same data using a non-Bayesian approach.

The dependent variable was coded as either opaque harmony (1) or surface-true harmony (0). An opaque response was defined as a back suffix attached to a raised FB root, or a front suffix attached to a raised BF root.

The independent variables were selected based on previous work on Uyghur or other languages with similar vowel harmony systems:

- The **log token frequency of the root** in the corpora, normalized to count per million words. This

[13] The transducer typically includes these suffixes as part of the root.

was included as a predictor because frequency is often an important driver of phonological variability (e.g., Coetzee & Kawahara, 2012).

- The **proportion of tokens of the root that are raised**. This is defined as the number of tokens of a root containing the raised allomorph divided by the total number of tokens of that root.

  For example, the root /ɑpæt/ 'disaster' occurs 1719 times in the corpus. Of these tokens, 544 are in forms that exhibit raising (e.g., [ɑpit-i] 'disaster-3.POSS') and 1175 are in unraised forms (e.g., [ɑpæt-lær] 'disaster-PL'). Thus the proportion of raised tokens for this root is 544/1719 = 0.32.

  This variable was included based on the observation by Hahn (1991b) that raised forms obscure the harmonic class of a root.

- The **identity of the underlying raised vowel** (F or B). This allows us to test the proposal that raised /ɑ/ is more likely to display surface-true harmony than raised /æ/ (Vaux, 2000).

- The **distance between the final two harmonizing root vowels**, counted in segments. Previous work on Uyghur (Mayer, 2021) and languages with similar harmony systems (Hayes et al., 2009; Rebrus & Törkenczy, 2017, 2021) suggest that the influence of a vocalic harmony trigger in the root decreases as greater numbers of transparent vowels intervene between it and the following suffix. In a root like /ɑpæt/ 'disaster' this distance is 1 while in /pɑʔɑlijæt/ 'activity' it is 3.

- The **distance between the root and the first harmonizing suffix**. This is the number of tags between the final tag of the root and the tag of the first harmonizing suffix. For example, /ɑpæt-i-GA/ 'disaster-3.POSS-DAT' has a distance of 1 (the non-harmonizing 3.POSS intervenes between the root and the harmonizing DAT) while /ɑpæt-GA/ 'disaster-DAT' has a distance of 0. Rebrus and Törkenczy (2017, 2021) found no influence of this factor on harmonizing behavior in Hungarian.

- Whether the root ends in one of the three **derivational suffixes** discussed above. This was operationalized as three separate boolean variables that indicate whether the root ends in /-ʧæ/, /-ɑnæ/, and /-nɑmæ/, respectively.

Random intercepts were defined for:

- The **author** of the article from which the token was drawn. This controls for different rates of opacity across writers. See Appendix E for considerations of authorship attribution. Because authors within each corpus are unique, a random effect for source corpus is unnecessary.

- **Root identity**. This controls for the idiosyncratic tendencies of roots that are not captured by the dependent variables.

The results shown in Table 6 suggest several frequency-related influences on opacity rates: more frequent roots are more likely to harmonize opaquely, but roots that frequently occur in their raised forms, with the underlying vowel identity obscured, are less likely to harmonize opaquely.

There are also phonological contributions. The model shows that underlying identity of the raised vowel is a significant predictor of opaque harmony, with underlying back vowels being less likely to harmonize opaquely (cf. Vaux, 2000). The distance between the final two harmonizing vowels in the root is also positively correlated with rates of opacity: as the disharmonic vowel in the root becomes further from the suffix, its influence decreases.

Finally, there is also evidence for morphological influences on opacity rates. The three derivational suffixes described above each produce greater rates of surface-true harmony than the general population of roots, and the specific rates vary between suffixes.

Term	Median	95% CI
Intercept	4.79	$[3.16, 6.47]$
Log token count	0.32	$[0.12, 0.51]$
Proportion raised	-2.21	$[-3.72, -0.73]$
Final vowel (reference level BF)	-3.68	$[-6.36, -1.11]$
Root vowel distance	0.79	$[0.40, 1.21]$
Suffix distance	-1.29	$[-1.45, -1.13]$
Has -*ane*	-4.64	$[-6.83, -2.65]$
Has -*name*	-2.10	$[-3.30, -0.89]$
Has -*che*	-9.59	$[-11.55, -7.74]$
Author (standard deviation)	1.01	$[0.76, 1.32]$
Root (standard deviation)	2.35	$[1.98, 2.80]$

Table 6: Statistical results from a mixed-effects logistic regression model whose parameters were estimated using Bayesian inference. The median is the most commonly sampled value, and the 95% Credible Interval shows the central range in which 95% of the sample values occur. A credible interval that does not contain zero is interpreted as a meaningful effect. The coefficients here are the estimated change in the log odds of opaque harmony given a unit increase in the corresponding variable, where the log odds are defined as $\log \frac{P(\text{opaque harmony})}{1-P(\text{opaque harmony})}$. Positive values indicate an increased likelihood of opaque harmony, while negative values indicate a decreased likelihood.

The distance between the root and the first harmonizing suffix is negatively correlated with opacity, which is unexpected given results from Rebrus and Törkenczy (2017, 2021). Because the majority of roots that can generate opacity are BF roots, I suspect this tendency might reflect an increased preference for the use of the default back suffix forms as distance between root and harmonizing suffix increases.

5 Theoretical challenges of these data

Given the quantitative results described above, a descriptively adequate (Chomsky, 1965) model of opacity in Uyghur should be able to account for the following properties:

- The majority of raised tokens harmonize opaquely, but cases of surface-true harmony also exist.
- The rate of opaque harmony varies across roots. This variation is correlated with a number of phonological and morphological factors, as well as frequency.

As shown in Section 3.6, both serial rule ordering and extensions of OT such as Stratal OT can predict either categorical surface-true harmony or categorical opaque harmony, but neither straightforwardly predicts the properties above. In fact, that *any* variability in rates of opacity exists at all is quite surprising under these analyses, since variation in the order of rules/strata is not expected (e.g., if backness harmony occurs at the word stratum and raising at the phrase stratum, then raising should never precede harmony).

Although it is likely possible to produce a descriptively adequate serial analysis that captures these properties by introducing the necessary theoretical machinery, such as probabilistic rule or stratum reordering, such modifications are likely to be ad hoc. A model that is explanatorily adequate will predict these properties rather than simply encode them. The next section briefly outlines such a model.

6 An analysis of opacity in the corpus

Backness harmony is typically treated as a phonological process (e.g., van der Hulst, 2016): whether a root takes front or back suffix allomorphs depends on which variant will minimize resulting surface disharmony (by some criteria), as in the simple model presented in Section 3.6. I will refer to this as the *phonological component* of backness harmony.

There are cases, however, where surface phonological properties are not sufficient to determine harmonizing behavior. In Uyghur, for example, the pair of words /sir/ 'secret' and /bir/ 'one' are nearly identical: however, the former takes back suffixes (e.g., [sirlɑr] 'secret-PL'), while the latter takes front suffixes (e.g., [birgæ] 'one-DAT'). Similarly, while the majority of roots containing no harmonizing vowels and a velar consonant /k g/ take front suffixes (e.g., [kir-gæ] 'dirt-DAT'), a smaller number take back suffixes (e.g., [gips-qɑ] 'plaster-DAT').

Hayes (2016) uses the term *zones of variation* to describe similar roots in Hungarian. Because their harmonic class is at best partially predictable from phonological properties, speakers disproportionately rely on lexical knowledge: that /sir/ takes back suffixes while /bir/ takes front suffixes must simply be memorized as a fact about each root. A consequence is that such roots typically display higher degrees of variability in suffix choice, particularly in wug tests where lexical knowledge is absent.

In this sense, backness harmony is similar to lexical class systems like grammatical gender. Roots can be thought of as falling into one of two harmonic classes (front or back) according to which suffix forms they take. I will refer to this as the *lexical component* of backness harmony.

Speakers simultaneously learn the lexical and phonological components of the harmony system (e.g., Zuraw, 2000). In the vast majority of cases in Uyghur, these components favor the same suffix choices. In a smaller number of cases, such as the zones of variation described above, the phonological component is less informative and lexical knowledge plays a larger role. The opaque forms discussed in this paper are cases where the phonological and lexical components actively *conflict*: in a case like /ɑpæt-i-GA/ 'disaster-3.POSS-DAT', the phonological component favors [ɑpitiʁɑ] because it is harmonic, while the lexical component favors the disharmonic [ɑpitigæ] because the root falls in the class of front harmonizers.

6.1 Modeling variable rates of opacity using lexically-indexed constraints

The model below will use Maximum Entropy Harmonic Grammar (henceforth MaxEnt; Goldwater & Johnson, 2003), a variant of Optimality Theory with numeric constraint weights (Pater, 2009).[14] Higher weights indicate a greater penalty for constraint violation. MaxEnt uses these weights and violation profiles to compute probability distributions over output candidates. See Appendix F for more detail.

The phonological component will be modeled using the simple VAGREE and *UNRAISED constraints introduced in Section 3.6. Because all roots undergo raising categorically, the ID constraints receive weights of zero and are omitted.

Lexical knowledge about the harmonic class of individual roots is modeled using *lexically indexed constraints* (e.g., Kraska-Szlenk, 1997; Ito & Mester, 1999; Pater, 2010; Moore-Cantwell & Pater, 2016). These are constraints that bear an index, and may only be violated by morphemes that bear the same index. The use of these constraints allows general phonological knowledge to be separated from idiosyncratic lexical knowledge in the grammar, and they have previously been applied in studies of opacity (Pater, 2014; Nazarov, 2019, 2020, 2021).

Assume the following pair of constraints:

(18) HARMONIZEBACK$_i$: Assess one violation if a front suffix is attached to a root with index i.

[14] The analysis is in principle compatible with any OT framework with numerically-weighted constraints.

(19) HARMONIZEFRONT$_i$: Assess one violation if a back suffix is attached to a root with index i.

The relative weight of these constraints reflects the certainty of the harmonic class of the root indexed by i. Identical weights indicate maximum uncertainty about class membership.

/ɑʔilæ$_i$-lAr/	Pred. Freq.	Obs. Freq.	H	VAGREEBACK $w = 5$	HARMBACK$_i$ $w = 0$	HARMFRONT$_i$ $w = 14.17$
ɑʔili-lær	2898	2898	5	1	1	0
ɑʔili-lɑr	0	0	14.17	0	0	1

Table 7: Tableau for the consistently opaque form /ɑʔilæ-lAr/ 'family-PL'.

/ɑhɑlæ$_j$-lAr/	Pred. Freq.	Obs. Freq.	H	VAGREEBACK $w = 5$	HARMBACK$_j$ $w = 2.78$	HARMFRONT$_j$ $w = 9.08$
ɑhɑli-lær	1969	1971	7.78	1	1	0
ɑhɑli-lɑr	537	535	9.08	0	0	1

Table 8: Tableau for the variably opaque form /ɑhɑlæ-lAr/ 'resident-PL'.

Tables 7 and 8 demonstrate how the relative weights of the HARMONIZE constraints determine the rate of opaque harmony. In Table 7, the large relative weights indicate total certainty that /ɑʔilæ/ takes front suffixes, and thus opaque harmony is categorical despite the resulting surface harmony violation. In Table 8, the smaller relative weights indicate less certainty in the lexical class membership: thus we see a relatively high rate of predicted surface-true harmony. Note that if the weight of VAGREEBACK were set to 0, near-categorical opaque harmony would be predicted in Tableau 8.[15]

6.2 Weighting the indexed constraints

The previous section showed how varying the relative weights of the indexed HARMONIZEBACK and HARMONIZEFRONT constraints produces different rates of opacity across roots. We now turn to the question of how these weights might be learned from data.

Because the weights of these constraints reflect certainty of class membership, it is natural to think of them in terms of probabilities. Let $P(\text{HC}|x)$ be a probability distribution over harmonic classes $\text{HC} \in \{\text{FRONT}, \text{BACK}\}$ given a root x. We will assume that the weights of these constraints are proportional to $P(\text{HC}|x)$, and hence inversely proportional to one another, as follows:

$$
\begin{aligned}
w_{\text{HarmonizeBack}_i} &\propto P(\text{HC} = \text{BACK}|x_i) \\
w_{\text{HarmonizeFront}_i} &\propto P(\text{HC} = \text{FRONT}|x_i) \\
&\propto 1 - P(\text{HC} = \text{BACK}|x_i)
\end{aligned}
$$

When $P(\text{HC} = \text{BACK}|x_i) \approx P(\text{HC} = \text{FRONT}|x_i)$ and there is high uncertainty about the class, the weights of the constraints will be roughly the same. When the probability of one class is greater than the other, the weights will change proportionally.

How do we determine $P(\text{HC}|x)$? That is, what properties are speakers sensitive to when determining the harmonic class of a root? I speculate here on a number of potentially relevant factors based on previous research on grammatical gender classification and the results from the corpus study:

[15]The reader might find it helpful to calculate the predicted frequencies under this weighting based on the equations in Appendix F.

- The **distribution** of the root (Kupish et al., 2022, refer to this as structural cues): do we typically see this root with front suffixes or back suffixes?

- **Phonotactic properties** of the root. Root phonotactics often provide useful clues to lexical class (e.g., Becker & Dow, 2013; Becker & Gouskova, 2016; Kupish et al., 2022), though the extent to which they are predictive varies between languages.

  Phonotactics in Uyghur are often highly predictive of harmonic class: if a root ends in a back vowel, you can be quite certain that it will belong to the class of back harmonizers, even if you have never encountered a suffixed form.

- **Morphological composition** of the root. Previous work suggests that morphological composition influences gender classification (e.g., Meunier et al., 2008; Becker & Dow, 2013). In Uyghur, certain derivational suffixes are more prone to surface-true harmony than others. This may relate to whether these derived forms are treated as roots in their own right, in which case opaque harmony might be expected, or as roots with disharmonic suffixes, in which case surface-true harmony would be preferred.

- **Prior biases**: Back harmony is the default class in Uyghur, and speakers may encode an overall preference for this class.

The frequency-based effects observed in the corpus study may connect to each of these factors: roots that occur more frequently provide greater evidence of their harmonic class; roots that occur more frequently in their unraised form provide greater evidence of their phonotactic properties; and the relative frequency of root and derived forms has been shown to predict morphological decomposability (e.g., Hay, 2001). Thus frequency plays an important role in this model, similar to other models of phonological variability (e.g., Coetzee & Kawahara, 2012).

6.3 Validating the model

To validate the model, I fit six simple MaxEnt models to the set of raised BB, BF, FB, and FF tokens from the corpora that had at least one harmonizing suffix. For simplicity, tokens for each root were aggregated based on whether they had front or back suffixes. This means the models do not consider the identity or number of suffixes.

The models are:

1. A **surface** model of harmony which contains a VAGREE constraint that mandates surface-true harmony. This model is only sensitive to surface properties of the output.

2. An **opaque** model that contains the HARMONIZEBACK and HARMONIZEFRONT constraints. The weights of the HARMONIZE constraints are shared by all roots, and violations of these constraints are based only on the underlying identity of the final vowel (e.g., attaching a back suffix to a root will violate HARMONIZEFRONT if the underlying final vowel is F, but not if it's B). This model is thus sensitive only to underlying vowel identity when choosing suffixes.

3. A **lexical** model. This uses the same constraints as the opaque model, except that (a) suffixed forms always violate their corresponding HARMONIZE constraint (e.g., forms with a back suffix always violate HARMONIZEFRONT); and (b) the global weights of the HARMONIZE constraints are scaled by $P(\text{HC}|x)$ for each root. This model is sensitive only to $P(\text{HC}|x)$ when determining suffix choice.

4. An **opaque-surface** model that combines the constraints in the opaque and surface models. This model allows independent contributions from both phonological and lexical factors, but lexical factors are entirely restricted to the identity of the final root vowel.

5. A **lexical-surface model** that combines the constraints in the lexical and surface models. This model allows independent contributions from phonological and lexical factors, but the strength of the lexical factors differ between roots. This corresponds to the analysis presented in this paper.

6. An **oracle** model where the probability of each form equal is set to its empirical frequency. This gives the maximum likelihood estimate of the data.

In order to estimate $P(\text{HC}|x)$ for each root, I trained a logistic regression model to predict the likelihood of observing each root with a back suffix based on properties of the root. This model is described in more detail in Appendix G. Crucially, this model was trained on the entire set of BB, BF, FB, and FF roots, not just raised tokens. The generalizations learned by the model thus reflect the general harmonizing behavior of the roots, not just their behavior in raised contexts.

Models were fit using the `maxent.ot` R package (Mayer & Zuraw, in prep). Table 9 shows the number of free parameters, log likelihood, and Bayesian Information Criterion (Burnham & Anderson, 2004) of the fitted models. BIC is a metric used in model selection that considers the number of free parameters and the fit to the data. Lower values are preferred, and a difference of > 10 is considered large.

Model	k	Log Likelihood	BIC
Surface	1	−114,333	228,679
Opaque	2	−13,686	27,397
Lexical	2	−12,039	24,104
Opaque-surface	3	−11,378	22,794
Lexical-surface	**3**	**−10,003**	**20,044**
Oracle	1620	−5,482	30,965

Table 9: Results for each model. k is the number of free parameters fitted to the data.[17] Higher log likelihoods indicate better model fit. BIC is a numeric criterion for comparing models, where lower values are better. $n = 230,161$ in all cases. The number of parameters for the oracle model is equal to the number of roots.

These results show that the variability in the data is better accounted for by models where (a) both phonological and lexical considerations influence suffix choice; and (b) the strength of the lexical component varies across roots according to $P(\text{HC}|x)$.

7 Discussion

Based on the results of the corpus study, this paper suggests that the opacity observed in Uyghur backness harmony is not strictly a phonological phenomenon: it is best analyzed as an interaction between phonological pressures towards surface-true harmony and lexical knowledge of the harmonic class of a root.

The analysis presented here is not only able to correctly model root-specific rates of opacity, but also accounts for the relationship of these rates to factors such as frequency and morphological composition.

[17] A more conservative estimate of k for the lexical models might include the free parameters in the logistic regression model used to calculate $P(\text{HC}|x)$. This would mean $k = 10, \text{BIC} = 24,203$ for the lexical model and $k = 11, \text{BIC} = 20,142$ for the lexical-surface model.

Speakers have been shown to be sensitive to these factors when determining the grammatical gender of a stem. If Uyghur speakers face a similar task when determining the harmonic class of a stem, the influence of these factors is to be expected.

When considering root-specific behavior in the Hungarian vowel harmony system, Hayes and Londe (2006) note that an acceptable analysis must preclude the existence of pathological examples like a simple B root (e.g., the hypothetical /pɑb/) categorically taking front suffixes. The current model allows for this pathology, to the extent that it is possible to weight the indexed constraints in such a way that this behavior would be predicted.

Setting aside the question of how such a pathological root could arise in Uyghur in the first place, even in a case where there is sudden massive and unequivocal distributional evidence for /pɑb/ as a front harmonizer, the model predicts that its long-term survival would be uncertain for two reasons:

1. The phonological component exerts pressure for surface-harmony based on general patterns in the language.

2. A root like /pɑb/ would be highly atypical within the class of Uyghur front roots. This would decrease $P(\textsc{hc} = \textsc{front})$.

Both of these factors predict that a root like /pɑb/ would be produced with back suffixes at least some of the time. Over generations this variation would decrease $P(\textsc{hc} = \textsc{front}|/\text{pɑb}/)$, likely resulting in a gradual drift towards back suffixation.

The model's predictions for the future of opacity in Uyghur are less clear. Although total surface-true harmony would be preferable from this the perspective of the phonology, the zones of variation in the harmony system preclude a total shift away from lexical factors.

There may also be biases towards the maintenance of this kind of opacity in harmony systems. The idea that there is pressure for roots to maintain the same harmonic class across allomorphs is closely related to a principle based on vowel harmony in Hungarian stated by Rebrus and Törkenczy (2017, 2021):

(20) **Harmonic Uniformity**: The harmonic class of a [suffixed] form is identical to that of its root (= monomorphemic base) (Rebrus & Törkenczy, 2017, p. 142).

Lexically-indexed HARMONIZE constraints mandate a similar property: the harmonic class of a root should remain identical across its allomorphs, even when doing so produces surface disharmony.

7.1 Implications for phonological theory

Opacity in the analysis presented above is modeled using lexically-indexed constraints, which were initially proposed to deal with phonological exceptionality. Although opacity has historically been represented using different theoretical mechanisms than exceptionality, the connections between opacity and exceptionality (and hence the lexicon) has been explored in previous literature (Łubowicz, 2003; Mielke et al., 2003; Sanders, 2003), and lexically-indexed constraints have been proposed as a tool for modeling opacity (Pater, 2014; Nazarov, 2019, 2020, 2021). The current paper provides additional evidence for a fundamental connection between exceptionality and (at least some forms of) opacity. A clearer understanding of this relationship might help to unify the variegated types of opaque phenomena observed to date (Baković, 2011; Baković & Blumenfeld, 2019).

An additional consequence of the treatment of opacity as a type of exceptionality is that it allows the Uyghur data to be analyzed using a strictly parallel model: the various serial mechanisms that have been proposed to handle opacity in OT are unnecessary in this case, and, indeed, do not predict the variable rates of opacity observed in the corpora. If speakers do indeed represent opacity as a type of phonological exceptionality, the inability of parallel models to represent opaque phenomenon as a purely phonological process may be seen as a point in their favor, rather than a failure.

7.2 Limitations and future directions

Because this is a corpus study, it is necessarily exploratory and does not constitute hypothesis confirmation (see Roettger, 2019, for more on this distinction). Exploratory analyses are most valuable as a tool for hypothesis generation. It will be important to test the predictions made by the model proposed here in controlled, experimental contexts. Experimental tasks that probe speakers' determination of harmonic class while carefully controlling for the phonological and lexical properties of roots will be informative in identifying the contributing factors and their relative importance (see Mayer, 2021, Ch. 4 for some preliminary work in this direction).

An additional limitation of the current study is that it relies on a relatively restricted genre of text: newspaper and academic articles. It will be valuable to look at opacity in more conversational corpora of Uyghur when such corpora become available.

Finally, because this study uses corpora where authorship cannot be uniquely determined, it does not tell us to what extent the variability we observe happens at the level of the individual or in aggregate across the population. It does, however, tell us where we might expect variation, either within or across speakers, or both. The sources of this variation can be carefully determined using more granular corpus or experimental studies.

8 Closing remarks

In addition to its empirical and theoretical contributions, this paper demonstrates the value of taking a more holistic and comprehensive empirical approach to linguistic data collection and analysis. The internet has allowed for the proliferation of large amounts of textual data, even for relatively small languages such as Uyghur. Computational tools such as morphological transducers can allow us to marshal the complexity inherent in such large data sets, and provide access to new types of empirical data that allow us to supplement other data sources and measure phonological patterns writ large. Such tools will become increasingly important as phonological theory continues to develop.

A Abbreviations

1.SG.POSS = first-person singular possessive, 3.POSS = third-person possessive, 3.SG.PST = third-person singular past, ACC = accusative, DAT = dative, DIM = diminutive, GER = gerund, LOC = locative, PFV = perfective, PL = plural, PROG = progressive.

B The Uyghur morphological transducer

This appendix describes the morphological transducer used to analyze the corpora in this paper.

The transducer is a modified version of the *apertium-uig transducer* (Littell et al., 2018; Washington et al., 2019). This is implemented using finite-state transducers (FST): specifically, within the HFST framework (Helsinki Finite State Technology; Linden et al., 2011). A FST is a finite-state automaton (FSA) that contains two tapes: in this case, one corresponding to underlying analyses and one to surface forms. Each transition or arc in the transducer has a symbol corresponding to each tape. Either tape may be designated as the input. The transducer reads the input and takes the appropriate transitions between states. The symbols on the transitions corresponding to the output tape are written to an output buffer. If the transducer reaches a valid output state after consuming the entire input, then the contents of the output buffer are returned.

Any SPE-style system that uses sequences of rewrite rules to map from underlying analyses to surface forms can be implemented as a finite-state transducer (Johnson, 1972; Kaplan & Kay, 1994). In practice, this poses several problems, the most serious of which is that although the mapping from an underlying analysis to a surface form is deterministic given a set of rules, the inverse is not true in general. In fact, it is possible for a given surface form to correspond to a large, or even infinite, number of underlying analyses under certain rule systems. This quickly becomes intractable for any practical implementation of a morphological transducer. The *two-level morphology* framework (Koskenniemi, 1983, 1984, 1986; Beesley & Karttunen, 2003), which is implemented in HFST, was designed to mitigate these issues.

Two-level morphology divides the mapping between underlying analyses and surface forms into two stages. The first stage maps between a morphological analysis and an abstract morphophonological form, which allows a minimal representation of roots and suffixes. For example, the analysis *qiz<dat>* will map to *qiz>{G}{A}* at this level, where > represents a morpheme boundary and *{G}* and *{A}* are essentially archiphonemes. It is this stage that solves the problem of overgeneration of underlying analyses: every valid underlying root must be explicitly encoded in the transducer.

The output of the first level then serves as input to the next level, which maps abstract morphophonological forms to surface forms. In this case, the phonological rules specified in the transducer will map *{G}* to *gh* and *{A}* to *a*, producing the surface form *qizgha* "to a girl."

In HFST, the first stage is implemented using the LEXC formalism, while the second is implemented using the TWOLC formalism. The rules specified at these levels are compiled into FSTs, which are then compose-intersected to form a single transducer. This transducer will only accept or propose underlying roots that are specified in the lexicon. Unfortunately, this introduces a degree of brittleness, since the transducer will not recognize any forms that are not present in the lexicon, and has no means by which to 'guess' the underlying form from the surface form unless augmented with additional tools.

The transducer can also map in the oppposite direction: between surface forms and underlying analyses that consist of roots plus morphological tags. For example, if the input is the surface form *qizingizgha* "to your daughter" the output analysis will be *qiz<n><px2sg><frm><dat>* (I use Latin orthography throughout this section rather than IPA, since it more closely reflects the input to the transducer). This indicates that the root is *qiz*, a noun <n>, and is suffixed with the 2nd person singular possessive marker in its formal form <px2sg><frm> followed by the dative suffix <dat>.

B.1 Modifying the transducer

I modified this transducer to detect the harmonic quality of suffixes when mapping from surface to underlying forms.[18] While a form like *qizingizgha* maps to *qiz<n><px2sg><frm><dat>* under the original transducer, it maps to *qiz<n><px2sg><frm><dat-b>* under this modified system, indicating that the dative suffix surfaces in one of its back harmonizing forms (*-qa* or *-gha*) rather than one of its front harmonizing forms (*-ke* or *-ge*).

The modified transducer splits each tag corresponding to a harmonizing suffix into three different forms corresponding to front variants (e.g., *<dat-f>*), back variants (e.g., *<dat-b>*), and ambiguous variants (e.g., *<dat>*). These tags are mapped to more restricted, though still abstract, morphological forms in the first stages. For example, *<dat-f>* will map to *{Gf}{Af}*, while *<dat-b>* will map to *{Gb}{Ab}*.

The second stage has been modified to map the newly introduced archiphonemes at the first stage to a restricted set of surface forms with corresponding backness. For example, it maps *{Gf}* and *{Af}* to only front allophones, and *{Gb}* and *{Ab}* to only back allophones. In addition, the restrictions the phonological component of the transducer imposes on harmony have been lifted. The original transducer, for example, would reject a form like **at-ler*, "horses", for being disharmonic. The modified transducer will simply interpret this as an instance of the front form of the plural suffix.

In a few cases, multiple parses that correspond to identical surface parses were removed from the transducer for simplicity. Take the word *doktur* 'doctor' as an example. This noun may be parsed as a nominal (appropriate in cases like *doktur chong* 'the doctor is old'), copular form (appropriate in cases like *Adil doktur* 'Adil is a doctor'), and so on. Because such distinctions are not relevant for the current project, all but the nominal parse was removed.

Finally, the vowel raising processes described in Section 3.3 can obscure the harmonizing quality of suffixes: for example, the surface realization of /dost-lAr-m/ 'friend-PL-1.SG.POSS' is *dostlirim*, which does not allow the backness of the plural morpheme to be determined. In such cases the modified transducer does not attempt to guess the backness of the suffix (i.e., to report either *<pl-f>* or *<pl-b>*), but will instead remain agnostic, simply reporting *<pl>*.

B.2 Interpreting and sanitizing the transducer output

Applying the transducer to the corpus produces one or more possible parses for each word that the transducer is able to recognize. The transducer was able to successfully analyze about 2.6 million of the 4.2 million words in the *Uyghur Awazi* corpus (61%), about 7 million of the 8 million words in the RFA corpus (87%), and about 1.8 million of the 2.4 million words in the Uyghur Academy Corpus (75%). The typical reason the transducer would fail to parse a word is because the root is not included in the transducer's list of valid roots: this is a key limitation of rule-based parsers. Tokens for which the transducer produced no parse were discarded.

Additional filtering was done to sanitize the data produced by the transducer. One challenge that arises is how to deal with multiple analyses in cases where they provide conflicting information about the root. The surface form *orgini*, for example, could correspond to underlying /orgɑn-i/ 'organ-3.POSS' or to /or-GAn-i/ 'harvest-PFV-3.POSS' (though the latter is unlikely because it corresponds to the front form of /-GAn/ attached to an umambiguously back root). I take a maximally conservative approach and discard all tokens for which such ambiguous parses are possible.

In addition, a number of suffixes are harmony blockers. Such suffixes, like the continuous suffix /-wɑt/, block harmony by failing to harmonize, and impose their own harmonic values on following suffixes.

[18] The code for this modified transducer can be found at `https://anonymous.4open.science/r/apertium-uig-EF36`.

(21) *Harmony blocking* /-wɑt/

	/søzlæ-wɑt-GAn/	→	[søzlæwɑtqɑn]	'speak-PROG-PFV'
cf.	/søzlæ-GAn/	→	[søzligæn]	'speak-PFV'

I discard tokens containing such suffixes.

Finally, a number of spurious parses were omitted based on manual inspection of the results. Particular attention was paid to tokens reflecting surface-true harmony to ensure these were not overcounted.

C Additional details on statistical models

C.1 Posterior samples

The Bayesian models were fit using the default parameters of the `brm` function. Fig. 4 shows samples from the posterior for each coefficient.

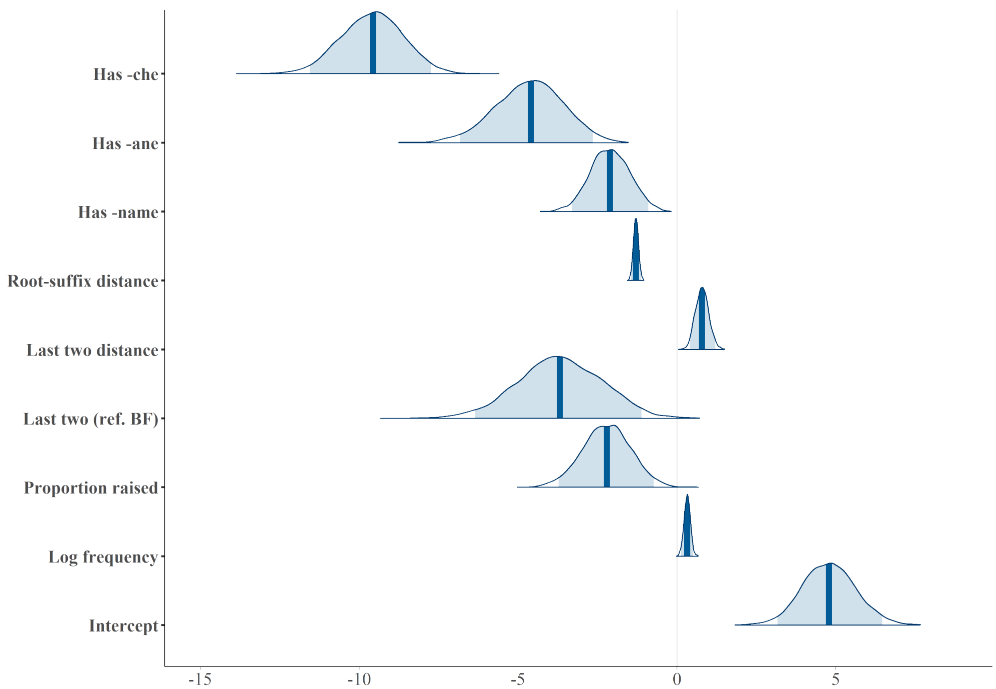

Figure 4: Plot of samples from the posterior for each model parameter. The dark areas are the median values, and the shaded areas are the 95% Credible Interval.

C.2 Frequentist analysis

The details of the model here are identical to those described in Section 4, except that the model was fitted and interpreted using the `lmer` package.

Whether the effect of each predictor was significant was calculated using the likelihood ratio test, which can be used to estimate p-values for mixed-effects models. The coefficients of each predictor and its associated χ^2 and p-value are shown in Table 10.

Term	β	χ^2	p
Intercept	4.83	–	–
Log token count	0.33	9.33	< 0.005
Proportion raised	-2.14	9.26	< 0.005
Final vowel (reference level BF)	-3.71	7.53	< 0.01
Root vowel distance	0.76	14.61	< 0.0005
Suffix distance	-1.29	273.36	< 0.0001
Has *-ane*	-4.70	22.04	< 0.0001
Has *-name*	-2.14	11.89	< 0.001
Has *-che*	-9.59	106.16	< 0.0001
Author	–	1011.5	< 0.0001
Word	–	3827.3	< 0.0001

Table 10: Statistical results from a mixed-effects logistic regression model. Significance testing was done using the likelihood ratio test. The coefficients here are the estimated change in the log odds of opaque harmony given a unit increase in the corresponding variable, where the log odds are defined as $\log \frac{P(\text{opaque harmony})}{1-P(\text{opaque harmony})}$. Positive values indicate an increased likelihood of opaque harmony, while negative values indicate a decreased likelihood.

D Gradient opacity within a single root

This appendix briefly illustrates examples of variation between opaque and surface-true harmony for a single root: *idare* /idɑræ/ 'office, bureau'. When used with the auxiliary verb *qilmaq* 'do', it can also mean 'to rule' or 'to govern'. This root has an overall frequency in the corpora of 1,122/million words and occurs in its raised form *idari* [idɑri] in 74% of tokens. The high frequency of the raised allomorph is likely due to the contexts in which this word tends to be used: since it's common to talk about an office relating to a person or entity, the root is often realized with the 3.POSS suffix [-si], as in *ürümchi sayaset* ***idarisi*** 'Ürümchi Tourism Office'.

When it occurs in its raised form with a harmonizing suffix attached, it displays opaque harmony in about 89% of cases (913 tokens) and surface-true harmony in about 11% of cases (113 tokens).

Below are examples of this root in its unsuffixed form and in its raised form with both opaque and surface-true harmony. All examples come from the RFA corpus. Translations are my own.

(22) *Unsuffixed token of 'idare'*
... döletni qanun arqiliq ***idare*** *qilish...*
'... the rule of law..." (literally 'ruling the country by means of the law')

(23) *Opaque token of raised 'idare'*
1980-yillardin boyan merkiziy axbarat ***idariside*** *ishligen.*
"He has worked at the Central Intelligence Agency since the 1980s."

(24) *Surface-true token of raised 'idare'*
Gülnar xanim saqchi ***idarisida*** *qandaq mu'amilige uchridi?*
"What kind of treatment did Gülnar receive at the police station?"

E Authorship in the corpora

Articles in the *Awazi* corpus are always published published under the anonymous byline *admin*. The *akadémiye* corpus is similar, with a small number of exceptions (resulting in a total of four unique author bylines). The RFA corpus provides greater attribution, but commonly uses partially anonymous bylines, such as *muxbirimiz Erkin* ‘our reporter Erkin’, resulting in a total of 154 unique author bylines.

These factors make consistent authorship attribution difficult. When coding the authorship variable used in the mixed effects logistic regression model described in Section 4, I use the listed author name when available. If no attribution is provided, the author is coded as ‘None_<corpus_name>’ to distinguish between unnamed authors in the different corpora.

F Maximum Entropy Optimality Theory

In a Maximum Entropy Optimality Theory grammar, each constraint is associated with a real-valued weight that represents its strength. In a grammar with N constraints, the weight of the ith constraint can be notated w_i. The function $C_i(x, y)$ returns the number of times an output candidate y for the input x violates the ith constraint. The *harmony* $H(x, y)$of an output candidate y given the input x is:

$$H(x, y) = \sum_{i}^{N} w_i C_i(x, y)$$

where higher values of $H(x, y)$ are associated with more severe constraint violations. The probability of an output candidate y given input x is

$$P(y|x) = \frac{\exp(-H(x, y))}{\sum_{z \in \Omega} \exp(-H(z, y))}$$

where Ω is the set of all possible output candidates given the input x.

The likelihood of a data set under a MaxEnt model can be calculated by multiplying together the probabilities assigned to each token by the model (or summing them in log space). It is also straightforward to learn constraint weights that optimize fit to a dataset: see Hayes and Wilson (2008) for more details on this learning procedure.

G Estimating $P(\textsc{hc}|x)$

In the lexical and lexical-surface MaxEnt models described in Section 6.3, a logistic regression model was used to estimate $P(\textsc{hc}|x)$, the probability of each root x being in the front/back harmonic class. The model was fitted to predict the proportion of times a root takes a back suffix.

The predictors for the model were:

- Underlying final vowel identity
- Log token count
- Proportion of raised forms
- Whether the root ends in /-nɑmæ/, /-ɑnæ/, or /-ʧæ/
- Interaction terms between underlying vowel identity and log token count, and underlying vowel identity and proportion of raised forms. These were included because the effect of the log token count and proportion of raised forms are different depending on the underyling final vowel.

None of these predictors relate to the specific context in which a root is used: they either capture aggregate properties of its behavior or phonotactic/morphological properties of the root itself. On this basis I assume the predictions of the model are a reasonable approximation of $P(\textsc{hc}|x)$.

The learned coefficients were:

Term	Coefficient
Intercept	3.35
Final vowel (F)	-7.19
Log token count	0.54
Proportion raised	-0.77
Has -*ane*	2.14
Has -*name*	1.38
Has -*che*	2.02
Final vowel (F) * Log token count	-1.00
Final vowel (F) * Proportion raised	3.74

Table 11: Learned coefficients from the logistic regression model for approximating $P(\textsc{hc}|x)$. The coefficients are the estimated change in the log odds of a back suffix given a unit increase in the corresponding variable, where the log odds are defined as $\log \frac{P(\text{back suffix})}{1-P(\text{back suffix})}$. Positive values indicate an increased likelihood of a back suffix, while negative values indicate a decreased likelihood.

References

Abdulla, A., Ebeydulla, Y., & Raxman, A. (2010). *Hazirqi zaman uyghur tili [Modern Uyghur].* Ürümchi: Xinjiang Xelq Neshriyati [Xinjiang People's Publishing House].

Al-Mozainy, H. Q. (1981). *Vowel alternations in a Bedouin Hijazi Arabic dialects* (Unpublished doctoral dissertation). University of Texas, Austin, Austin, TX.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, *5*(2), 183-207.

Baković, E. (2007). A revised typology of opaque generalisations. *Phonology*, *24*(2), 1–43.

Baković, E. (2011). Opacity and ordering. In J. A. Goldsmith, J. Riggle, & A. C. Yu (Eds.), *The Handbook of Phonological Theory* (2nd ed., pp. 40–67). London: Wiley-Blackwell.

Baković, E., & Blumenfeld, L. (2019). Rule interaction conversion operations. *Loquens*, *6*(2), e062.

Becker, M., & Dow, M. (2013). Gender without morphological segmentation in french. In *Phonology 2013.* UMass Amherst.

Becker, M., & Gouskova, M. (2016). Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry*, *47*(3), 391–425.

Beesley, K., & Karttunen, L. (2003). *Two-level rule compiler* (Retrieved from https://web.stanford.edu/ laurik/.book2software/twolc.pdf). Stanford University.

Bermúdez-Otero, R. (2003). The acquisition of phonological opacity. In J. Spenader, A. Eriksson, & Östen Dahl (Eds.), *Variation within optimality theory: Proceedings on the stockholm workshop on 'variation within optimality theory'* (pp. 25–36). Stockholm: Department of Linguistics, Stockholm University.

Bermúdez-Otero, R. (2018). Stratal phonology. In S. Hannahs & A. R. Bosch (Eds.), *The Routledge Handbook of Phonological Theory* (pp. 100–134). Abingdon: Routledge.

Bodrogligeti, A. J. E. (2001). *A grammar of Chagatay.* Muenchen, Germany: LINCOM EUROPA.

Bowers, D. (2019). The Nishnaabemwin restructuring controversy: New empirical evidence. *Phonology*, *36*(2), 187–224.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Burnham, K. P., & Anderson, D. R. (2004). Multimodal inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*, 261–304.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English.* New York: Harper & Row.

Clauson, G. (1972). *An etymological dictionary of pre-thirteenth-century Turkish.* Oxford: Clarendon Press.

Coetzee, A. W., & Kawahara, S. (2012). Frequency biases in phonological variation. *Natural Language and Linguistic Theory*, *31*, 47–89.

Donegan, P. J., & Stampe, D. (1979). The study of natural phonology. In D. A. Dinnsen (Ed.), *Current approaches to phonological theory* (pp. 126 – 173). Bloomington, IN: Indiana University Press.

Engesæth, T., Yakup, M., & Dwyer, A. (2009/2010). *Teklimakandin salam: hazirqi zaman Uyghur tili qollanmisi / Greetings from the Teklimakan: a handbook of modern Uyghur.* Lawrence: University of Kansas Scholarworks.

Goldwater, S., & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In J. Spenader, A. Eriksson, & Östen Dahl (Eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory* (pp. 111–120). Stockholm: Stockholm University, Department of Linguistics.

Hahn, R. F. (1991a). Diachronic aspects of regular disharmony in modern Uyghur. In W. Boltz & M. Shapiro (Eds.), *Studies in the Historical Phonology of Asian Languages.* John Benjamins.

Hahn, R. F. (1991b). *Spoken Uyghur.* Seattle, WA: University of Washington Press.

Hall, D. C., & Ozburn, A. (2018). *When is derived [i] transparent? a subtractive approach to Uyghur vowel harmony.* (Talk presented at the 49th Northeast Linguistics Society Conference (NELS 49), Cornell University, 5-7 October)

Halle, M., Vaux, B., & Wolfe, A. (2000). On feature spreading and the representation of place of articulation. *Linguistic Inquiry*, *31*, 387-444.

Hay, J. (2001). Lexical frequency in morphology: is everything relative? *Linguistics*, *39*(6), 1041–1070.

Hayes, B. (2016). Comparative phonotactics. In *Proceedings of the 50th meeting of the Chicago Linguistics Society* (pp. 265–285).

Hayes, B., & Londe, Z. (2006). Stochastic phonological knowledge: the case of Hungarian vowel harmony. *Phonology*, *23*, 59–104.

Hayes, B., & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, *39*(3), 379–440.

Hayes, B., Zuraw, K., Siptar, P., & Londe, Z. (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language*, *85*, 822–863.

Hooper/Bybee, J. (1976). *An introduction to nature generative phonology.* New York: Academic Press.

Ito, J., & Mester, A. (1999). The phonological lexicon. In N. Tsujimura (Ed.), *The Handbook of Japanese Linguistics* (pp. 62–100). Malden, MA and Oxford, UK: Blackwell.

Jarosz, G. (2014). Serial markedness reduction. In J. Kingston, C. Moore-Cantwell, J. Pater, & R. Staubs (Eds.), *Proceedings of the 2013 Annual Meeting on Phonology.*

Johnson, C. D. (1972). *Formal aspects of phonological decription.* The Hague: Mouton.

Kaplan, R., & Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics*, *20*, 331-378.

Kiparsky, P. (1971). Historical linguistics. In W. Dingwall (Ed.), *A survey of linguistic science* (pp. 576–642). College Park: University of Maryland Linguistics Program.

Kiparsky, P. (1973). Abstractness, opacity, and global rules. In O. Fujimura (Ed.), *Three dimensions of linguistic theory* (pp. 57–86). Tokyo: TEC.

Kiparsky, P. (2000). Opacity and cyclicity. *The Linguistic Review*, *17*, 351–367.

Kirchner, R. (1996). Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry*, *27*(2), 341–350.

Koskenniemi, K. (1983). *Two-level morphology: A general computational model for word-form recognition and production* (Publication 11). University of Helsinki, Department of General Linguistics, Helsinki.

Koskenniemi, K. (1984). A general computational model for word-form recognition and production. In *Coling'84* (pp. 178–181).

Koskenniemi, K. (1986). Compilation of automata from morphological two-level rules. In F. Karsson (Ed.), *Papers from the Fifth Scandinavian Conference on Computational Linguistics* (pp. 143–149).

Kraska-Szlenk, I. (1997). Exceptions in phonological theory. In B. Caron (Ed.), *Proceedings of the 16th International Congress of Linguists.* Oxford: Pergamon.

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan.* Academic Press.

Kupish, T., Geiss, M., Mitrofanova, N., & Westergaard, M. (2022). Structural and phonological cues for gender assignment in monolingual and bilingual children acquiring german. experiments with real and nonce words. *Glossa*, *7*.

Lindblad, V. M. (1990). *Neutralization in Uyghur.* University of Washington.

Linden, K., Silfverberg, M., Axelson, E., Hardwick, S., & Pirinen, T. (2011). HFST – Framework for compiling and applying morphologies. In C. Mahlow & M. Pietrowski (Eds.), *Systems and frameworks for computational morphology* (Vol. 100, pp. 67–85). Communications in Computer and Information Science.

Littell, P., Tian, T., Xu, R., Sheikh, Z., Mortensen, D., Levin, L., ... Hovy, E. (2018). The ARIEL-CMU situation frame detection pipeline for LoReHLT16: a model translation approach. *Machine Translation*, *32*, 105–126.

Łubowicz, A. (2003). *Contrast preservation in phonological mappings* (Unpublished doctoral dissertation). University of Massachusetts Amherst.

Mayer, C. (2021). Capturing gradience in long-distance phonology using probabilistic tier-based strictly local grammars. *Proceedings of the Society for Computation in Linguistics*, *4*(5).

Mayer, C., Major, T., & Yakup, M. (2022). Are neutral roots in Uyghur really neutral? evaluating a covert phonemic contrast. In P. Jurgec et al. (Eds.), *Supplemental Proceedings of the 2021 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America.

Mayer, C., McCollum, A., & Eziz, G. (accepted). Issues in uyghur phonology. *Language and Linguistics Compass*.

Mayer, C., & Zuraw, K. (in prep). *`maxent.ot`: A package for doing maximum entropy optimality theory in r.* Retrieved from `https://github.com/connormayer/maxent.ot`

McCarthy, J. J. (1999). Sympathy and phonological opacity. *Phonology*, *16*, 331–339.

McCarthy, J. J. (2007). *Hidden generalizations: Phonological opacity in optimality theory*. London: Equinox Publishing.

McCollum, A. (2021). Transparency, locality, and contrast in Uyghur backness harmony. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *12*(10), 10.

Meunier, F., Seigneuric, A., & Spinelli, E. (2008). The morpheme gender effect. *Journal of Memory and Language*, *58*(1), 88–99.

Mielke, J., Hume, E., & Armstrong, M. (2003). Looking through opacity. *Theoretical Linguistics*, *29*(1–2).

Moore-Cantwell, C., & Pater, J. (2016). Gradient exceptionality in maximum entropy grammar with lexically specific constraints. *Catalan Journal of Linguistics*, *15*, 53–66.

Nazarov, A. (2019). Formalizing the connection between opaque and exceptionful generalizations. *Toronto Working Papers in Linguistics*, *41*(1).

Nazarov, A. (2020). Bedouin Arabic multiple opacity with indexed constraints in Parallel OT. In H. Baek, C. Takahashi, & A. H.-L. Yeung (Eds.), *Proceedings of the 2019 Annual Meeting on Phonology*.

Nazarov, A. (2021). Learnability of indexed constraint analyses of phonological opacity. In *Proceedings of the society for computation in linguistics* (Vol. 4).

Nazarov, A., & Pater, J. (2017). Learning opacity in Stratal Maximum Entropy Grammar. *Phonology*, *34*, 299–324.

Nazarova, G., & Niyaz, K. (2013). *Uyghur: An elementary textbook*. Washington, DC: Georgetown University Press.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas – Part II. *Language and Linguistics Compass*, *10*(11), 591–613.

Pater, J. (2009). Weighted constraints in generative linguistics. *Cognitive Science*, *33*, 999–1035.

Pater, J. (2010). Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In S. Parker (Ed.), *Phonological argumentation: Essays on evidence and motivation* (pp. 123–154). London: Equinox.

Pater, J. (2014). Canadian raising with language-specific weighted constraints. *Language*, *90*, 230–240.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Cambridge, MA: Blackwell. (Technical Report CU-CS-696-93, Department of Computer Science, University of Colorado at Boulder, and Technical Report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, April 1993.)

R Core Team. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rebrus, P., & Törkenczy, M. (2017). Co-patterns, subpatterns and conflicting generalizations in Hungarian vowel harmony. In H. van der Hulst & A. Lipták (Eds.), *Approaches to Hungarian. Volume 15: Papers from the 2015 Leiden Conference*. Amsterdam: John Benjamins Publishing Company.

Rebrus, P., & Törkenczy, M. (2021). Harmonic Uniformity and Hungarian front/back harmony. *Acta Linguistica Academica*, *68*(1–2), 175–206.

Roettger, T. B. (2019). Researcher degrees of freedom in phonetic research. *Laboratory Phonology*, *10*(1).

Sanders, R. N. (2003). *Opacity and sound change in the Polish lexicon* (Unpublished doctoral dissertation). UCSC.

Speelman, D. (2014). Logistic regression: A confirmatory technique for comparisons in corpus linguistics. In *Corpus methods for semantics: Quantitative studies in polysemy and synonymy* (Vol. 43, pp. 487–533).

Steriade, D. (2000). Paradigm uniformity and the phonetics-phonology interface. In M. Broe & J. Pierrehumbert (Eds.), *Papers in laboratory phonology v* (pp. 313–334). Cambridge, MA: Cambridge University Press.

Sumner, M. (2003). *Testing the abstractness of phonological representations in Modern Hebrew weak verbs* (Unpublished doctoral dissertation). State University of New York at Stony Brook.

van der Hulst, H. (2016). Vowel harmony. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics.* Oxford: Oxford University Press.

Vaux, B. (2000). Disharmony and derived transparency in Uyghur vowel harmony. *Proceedings of NELS 30*, 671-698.

Vaux, B. (2008). Why the phonological component must be serial and rule-based. In B. Vaux & A. Nevins (Eds.), *Rules, constraints, and phonological phenomena* (pp. 20–60). Oxford University Press.

Vaux, B. (2011). Language games. In J. A. Goldsmith, J. Riggle, & A. C. Yu (Eds.), *The Handbook of Phonological Theory* (2nd ed., pp. 722–750). London: Wiley-Blackwell.

Washington, J., Salimzianov, I., Tyers, F. M., Gökırmak, M., Ivanova, S., & Kuyrukçu, O. (2019). Free/open-source technologies for Turkic languages developed in the Apertium project. In *Proceedings of the International Conference on Turkic Language Processing (TURKLANG 2019).*

Zhang, J. (2019). Speakers treat transparent and opaque alternation patterns different – evidence from Chinese tone sandhi. In R. Stockwell, M. O'Leary, Z. Xu, & Z. Zhou (Eds.), *Proceedings of the 36th West Coast Conference on Formal Linguistics.* Somerville, MA: Cascadilla Proceedings Project.

Zuraw, K. (2000). *Patterned exceptions in phonology* (Unpublished doctoral dissertation). University of California, Los Angeles.