

# Transparency, locality, and contrast in Uyghur backness harmony

Adam G. McCollum  
Department of Linguistics  
Rutgers University  
New Brunswick, NJ USA  
adam.mccollum@rutgers.edu

## Abstract:

Theories of vowel harmony have wrestled with the formal challenges of transparency, notably the increased expressivity resulting from non-local dependencies. However, experimental work has demonstrated on a number of occasions that ‘transparent’ vowels actually undergo harmony (e.g. Gick et al., 2006), re-establishing the role of locality in the analysis of harmony. Existing work on backness harmony in Uyghur argues that /i/ is transparent to harmony, with some proposing that this fact falls out from the absence of contrastive /u/ in the language (e.g. Vaux, 2000). This paper examines the distribution of [i] and [u] within roots and suffixes to assess their phonological status, as well as their participation in harmony. Results indicate there are no transparent vowels in Uyghur, suggesting that harmony is truly local in the language. Results for root-internal [i] and [u] are less clear; evidence from most speakers supports an allophonic relationship, while some evidence for other speakers suggests the possibility of contrastive /i/ and /u/.

Keywords: vowel harmony, locality, transparency, contrast, Uyghur

## 1 Introduction

One of the central restrictions imposed by theories of vowel harmony is locality (Goldsmith, 1976; Archangeli & Pulleyblank, 1994; Gafos, 1999; Nevins, 2010). While the representational structures may differ across various theories, each theory demands that locality constrain harmonic interactions. Within computational phonology, locality has been leveraged to defend the relative simplicity of phonological patterns, including harmony (Heinz 2018; Heinz & Lai 2013). When faced with non-local dependencies in harmony (also called transparency), researchers have developed elaborate theoretical mechanisms in order to satisfy a formal definition of locality (e.g. Hayes & Wilson, 2008; Nevins, 2010; Odden, 1994; van der Hulst & van der Weijer, 1995). At a fundamental level though, the empirical question is still largely unaddressed – how local is harmony? Stated differently, what bounds exist on transparency? Phonetic and phonological research on languages with reported transparency has

found that in many cases ‘transparent’ segments actually alternate for harmony (Benus & Gafos, 2007; Ritchart & Rose, 2017). Furthermore, some work argues that transparency may also be constrained by distance. In Hungarian, as the number of transparent vowels increases, these vowels begin to block harmony (Hayes & Londe, 2006; Ringen & Kontra, 1989).

Transparency is also connected to questions concerning contrast and its role in the grammar. Following Calabrese (1995), Nevins (2010) accounts for transparency in harmony by relativizing a search function to operate over either all, marked, or contrastively specified segments. If harmony in a given language applies only to contrastively specified segments, then non-contrastive differences should not play a role in the pattern (Halle et al., 2000; Vaux, 2000; Nevins, 2010). In Lexical Phonology (Kiparsky, 1985; Mohanan, 1982), vowel harmony and contrast are yoked together. Typically, harmony is a lexical pattern, operating over only contrastively specified segments at an early stage in the phonological derivation. In contrast, allophonic patterns are postlexical, and as a consequence cannot affect the application of harmony within the lexical phonology. Historical contrasts have also been used to develop abstract analyses that rely on a larger underlying and intermediate inventory of sounds than found on the surface (Vago, 1973, 1976). Despite enforcing locality during the derivation, these approaches do not demand locality on the surface.

This paper examines backness harmony in Uyghur to determine what role contrast plays in the application of harmony. Transparency in previous analyses of Uyghur hinges on the claim that /i/ has no [+back] counterpart (Halle et al., 2000; Vaux, 2000, Nevins, 2010). Interestingly, these works ignore the role that consonants play in the distribution of high unrounded vowels. As descriptive work notes, [ɯ] is a relatively common surface vowel sound in the language, due to consonantal context. Thus, surface [ɯ] is licit in the language, regardless of its contrastive status. In addition to the significance of contrast and allophony, the paper examines how non-local may harmony be. Extant work describes harmony in Uyghur as able to skip multiple /i/ vowels. If harmony may span any number of transparent vowels, this would offer new insight into the general status of locality in harmony. Alternatively, if harmony is more local than previously described, this would lend credence to claims concerning the ontological locality of phonological patterns. Regardless of the particular implications of the paper, this study presents experimental findings and their significance for more general theoretical analysis (Pierrehumbert et al., 2000).

This paper is structured as follows. In Section 2, I describe Uyghur backness harmony, focusing on the realization of high unrounded vowels, outlining the claims advanced in previous work. Section 3 discusses phonetic and phonological studies on transparency. Section 4 details the methods used during data collection and analysis. Section 5 reports results from

root-internal and suffix data, which I then discuss in Section 6, relating findings to the notions of locality and contrast. Finally, in Section 7, I conclude the paper.

## 2 Uyghur backness harmony

### 2.1 Inventory

The Uyghur inventory includes at least seven contrastive vowels, /ɑ o u æ ø i y/, which are distinguishable in terms of three features, [back], [high], and [round], shown below in Table 1. In addition to these vowels, /e/ is marginal, typically occurring in non-nativized loans, and in the initial-syllable only (Hahn, 1991:37). As noted by Nadzhip (1971:48-49) and Yakup (2005:31), [e] is typically a raised variant of /æ/ or /ɑ/, further suggesting the peripheral status of this sound. In comparison with the common Turkic eight-vowel system (e.g. Menges, 1995), where the underlying inventory is distinguished by these three features, there is one notable gap in the Uyghur inventory – there is no [+back] counterpart of /i/.

Table 1: Uyghur vowel inventory

	[-back]		[+back]	
	[-round]	[+round]	[-round]	[+round]
[+high]	i	y		u
[-high]	æ (e)	ø	ɑ	o

### 2.2. Backness harmony

Backness harmony in Uyghur triggers alternations on the low vowels and the high rounded vowels. In (1a-d), the locative suffix alternates between [-dæ] and [-da] according to the backness of the initial-syllable vowel. Similarly, in (1e-h) the backness of the initial syllable conditions the backness of the vowel of the gerundial suffix, as well as the place of articulation of the preceding dorsal consonant, [g]~[ʁ].

- (1) Backness harmony in Uyghur
- a. køl-dæ 'lake-LOC'
  - b. bæl-dæ 'waist-LOC'
  - c. jol-da 'road-LOC'
  - d. bal-da 'honey-LOC'

e.	kæl-gy	‘come-GER’
f.	bær-gy	‘give-GER’
g.	qal-ƙu	‘remain-GER’
h.	bar-ƙu	‘go-GER’

In addition to the low vowels and the high round vowels, two other sets of vowels can be defined for harmony. The first is comprised of the mid vowels /e ø o/, which do not typically occur in non-initial syllables in most dialects of the language (though see Abdurehim, 2014:70-75 for the non-initial occurrence of [ø o] in the Lopnor dialect). For this reason, the paper does not focus on these vowels. The second set contains /i/, which previous work reports is transparent to harmony (Lindblad, 1990; Hahn, 1991; Vaux, 2000; Mayer & Major, 2018; Mayer et al., 2019). Relatedly, root /i/ triggers [+back] vowel suffixes in some roots (2a,b), but front vowel suffixes in others (2c,d). In polysyllabic roots /i/ is transparent; second-syllable /i/ of disyllabic roots does not affect the realization of following vowels (2e,f).<sup>1</sup>

(2) /i/ in Uyghur roots

a.	it-da	‘dog-LOC’
b.	til-da	‘tongue-LOC’
c.	biz-dæ	‘1P-LOC’
d.	siz-dæ	‘2S.FORM-LOC’
e.	qædir-gæ	‘regard-DAT’
f.	gazir-ƙa	‘sunflower-DAT’

In (2), variation occurs across different lexical roots, but the data in (3) demonstrates that a single root may trigger both front and back suffixes. In (3a,b), the roots /iʃ/ ‘work’ and /tʃiʃ/ ‘tooth’ trigger the [+back] variants of the plural and dative suffixes. Yet, in (3c,d), these roots trigger the [-back] variants of the verbalizer suffix. While this variation may be partially conditioned upon the particular suffixes involved, there is also variation that appears completely independent of suffix type. Observe in (3e,f) that /ilim/ ‘science’ may trigger either front or back variants of the plural suffix. Similar variation is described in Lindblad (1990:26), and is common in the 2016 Wikipedia Uyghur corpus maintained in the Leipzig Corpora Collection (Goldhahn et al., 2012). No other vowels behave like /i/ as a trigger for harmony – the back vowels /ɑ o u/ consistently trigger [+back] suffixes, and the front vowels /æ y ø/ consistently trigger [-back] suffixes.

---

<sup>1</sup> Thanks to Xiayimaierdan Abudushalamu for providing the forms in (2e,f).

- (3) Variable backness harmony after /i/ (retrieved from the 2016 Wikipedia Uyghur corpus in the Leipzig Corpora Collection; Goldhahn et al., 2012)
- |    |             |                     |
|----|-------------|---------------------|
| a. | ij-lar      | ‘work-PL’           |
| b. | tʃij-i-ɤɑ   | ‘tooth-POSS.3S-DAT’ |
| c. | ij-læ       | ‘work-VRB’          |
| d. | tʃij-læ-p   | ‘bite-VRB-CVB’      |
| e. | ilim-lar-ni | ‘science-PL-ACC’    |
| f. | ilim-lær-ni | ‘science-PL-ACC’    |

Like non-initial /i/ in roots, suffix /i/ is also transparent, as seen in (4). In (4a,b), the backness of the roots extends across two /i/ vowels to determine the realization of the locative suffix.<sup>2</sup>

- (4) Transparent suffixal /i/ (Hahn, 1991; Vaux, 2000)
- |    |             |                    |
|----|-------------|--------------------|
| a. | køi-imiz-dæ | ‘lake-POSS.1P-LOC’ |
| b. | jol-imiz-da | ‘road-POSS.1P-LOC’ |

In addition to underlying high vowels, high vowels may surface via vowel raising or epenthesis in the language. Epenthetic vowels are inserted to break up illicit coda clusters. The first-person singular possessive suffix, /-m/, regularly triggers epenthesis when attached to consonant-final roots. In (5a,b) the suffix attaches to a vowel-final root, surfacing without any epenthetic vowel. In (5c-f) however, the suffix is preceded by an epenthetic high vowel. The epenthetic vowel is rounded after [+round] roots (5c,d) and unrounded after [-rounded] roots (5e,f). Lindblad (1990) and Vaux (2000) analyze the epenthetic vowel as copy of preceding backness and rounding, subject to a prohibition on [u], resulting in surface [i] instead of [u] after back vowels (5f).<sup>3</sup> When [i] surfaces, the backness of the preceding vowel still controls the realization of subsequent suffixes, like the locative suffix.

---

<sup>2</sup> Somewhat contrary to the data in (4), Hahn (1998:385) indicates that the initial vowel of the first-person plural possessive suffix does undergo rounding harmony (e.g. [kølymizdæ] ‘lake-POSS.1P-LOC’), analyzing it as epenthetic rather than underlying. This does not, however, affect the status of the second vowel in this possessive suffix, since it is always [i] and is always skipped by harmony.

<sup>3</sup> Low vowels are variably fronted and raised before high vowel suffixes /bæl-i/ [bæli]~[beli] ‘waist-POSS.3S’.

- (5) Transparent epenthetic [i]
- |    |            |                      |
|----|------------|----------------------|
| a. | sællæ-m-dæ | ‘turban-POSS.1S-LOC’ |
| b. | bala-m-da  | ‘child-POSS.1S-LOC’  |
| c. | køl-ym-dæ  | ‘lake-POSS.1S-LOC’   |
| d. | jol-um-da  | ‘road-POSS.1S-LOC’   |
| e. | bæl-im-dæ  | ‘waist-POSS.1S-LOC’  |
| f. | bal-im-da  | ‘honey-POSS.1S-LOC’  |

As for vowel raising, low vowels in medial open syllables raise to [+high]. To help keep track of underlying vowel qualities, both underlying and surface forms are presented in (6). In (6a), the unaffixed root for ‘child’ shows that the word-final vowel is low. The underlying height of the root-final vowel is also evident in (6b,c), where this vowel occurs in a closed syllable. Yet, when this vowel occurs in a medial open syllable, as in (6d), it raises to [i]. The same generalizations hold for all underlying low vowels, as in (6e-j).

- (6) Transparent raised vowels
- |    |              |            |                      |
|----|--------------|------------|----------------------|
| a. | /bala/       | [bala]     | ‘child’              |
| b. | /bala-m/     | [balam]    | ‘child-POSS.1S’      |
| c. | /bala-m-dæ/  | [balamda]  | ‘child-POSS.1S-LOC’  |
| d. | /bala-dæ/    | [balida]   | ‘child-LOC’          |
| e. | /bala-lær/   | [balilar]  | ‘child-PL’           |
| f. | /sællæ/      | [sællæ]    | ‘turban’             |
| g. | /sællæ-m/    | [sællæm]   | ‘turban-POSS.1S’     |
| h. | /sællæ-m-dæ/ | [sællæmdæ] | ‘turban-POSS.1S-LOC’ |
| i. | /sællæ-dæ/   | [sællidæ]  | ‘turban-LOC’         |
| j. | /sællæ-lær/  | [sællilær] | ‘turban-PL’          |

Whether underlying, epenthetic, or derived by raising, Lindblad (1990:13) and Vaux (2000:2) suggest that /i/ does not even exhibit small, phonetic effects for harmony. Lindblad (1990:13) writes:

The choice of allophones is based on the immediate phonetic environment, and especially the adjacent consonants. Thus, for example, the genitive suffix +nɪŋ + *is always pronounced with schwa as its vowel* (Hahn 1986:46), regardless of its backness value as revealed by harmonic processes [emphasis mine].

Lindblad's claim is noteworthy, since it suggests that backed variants of /i/ may surface due to consonantal effects, but not due to harmony. The distribution of the high unrounded vowels is noted most extensively in Hahn (1986:43-48, 1991:34-37). Generalizing over Hahn's (1991:34-37) complex set of allophonic statements, backed allophones of /i/, including [ə i ɤ u], appear following dorsal obstruents, laryngeals, and before /l/ and /ŋ/. In all other contexts, /i/ surfaces as a true front vowel, usually transcribed as [i] or [ɪ]. To simplify discussion throughout, all front allophones of /i/ are transcribed as [i], while all back allophones of /i/ are transcribed as [u].

### 2.3 Implications of previous analyses

In most work on the language, underlying /u/ is preserved, but surface [u] is neutralized with [i], in line with Vago's (1973) abstract analysis of Hungarian (Hahn, 1991; Lindblad, 1990; Yakup, 2005). The late neutralization rule crucially allows harmony to operate locally during the lexical stratum of the phonology, only to be masked by the late neutralization of /u/ and /i/. In Vaux (2000), a different analysis proposed, which requires harmony to operate over contrastively specified vowels only (see also Calabrese, 1995; Nevins, 2010). Under Vaux's analysis, since /i/ does not have a [+back] harmonic counterpart, the harmony rule effectively skips the high unrounded vowel.

Problematically, in all previous work the fact that [u] may not arise via harmony is never juxtaposed to the allophonic backing of /i/ to [u] in certain consonantal contexts. For instance, Hahn (1986, 1991) reports that /i/ surfaces as a back vowel before /l/, which suggests that the transcription of the forms in (6e,j) should be [baluɫar] rather than [balilar] 'child-PL' and [sælluɫær] rather than [sællilær] 'turban-PL' respectively. When the role of adjacent consonants is considered, the transparency of [bali-da] 'child-LOC' is masked by the allophonic backing of /i/ before the lateral in [balu-lar]. The same complication holds for front vowel forms, too. Given the effects of flanking consonants, transparent, but harmonic /i/ in [sælli-dæ] 'turban-LOC' should be transparent and disharmonic [u] in [sællu-lær] 'turban-PL'. In both sets of forms, the predicted allophonic backing of /i/ to [u] before /l/ obscures the harmony pattern. If this is the true state of affairs in the language, with both vowel and consonantal features driving high unrounded vowel backness, transparency appears more complex than other cases of transparency in harmony.

In fact, if both vowels and consonants condition the backness of a given vowel, harmony in Uyghur is more complex than other vowel harmony patterns (e.g. consonant harmony; Mayer & Major, 2018), contradicting expressivity claims in Heinz (2018). If Uyghur exhibits the pattern described in previous work, it offers significant

insight into the complexity of harmony patterns. Conversely, if the reported patterns diverge from the pattern recorded in a controlled study, this would offer further evidence in favor of a simpler, more local interpretation of transparency in harmony.

### 3 Previous work on transparency in harmony

Reported transparency has prompted a number of experimental studies. Two important points have emerged from this literature: ‘transparent’ vowels often actually phonetically alternate for harmony, and transparency may be subject to distance-based constraints.

#### 3.1 Pseudo-transparency

First, in a number of languages with reported transparency, phonetic studies have shown that the ‘transparent’ vowels actually alternate for harmony (Gordon, 1999; Gick et al., 2006; Benus & Gafos, 2007; Ritchart & Rose, 2017). Gordon (1999) examines putatively transparent /e/ and /i/ in Finnish, reporting small (up to 100 Hz) differences in the second formant based on backness context. Similarly, Benus & Gafos (2007) find small articulatory and acoustic differences in Hungarian /i/ based on backness context, arguing that these vowels are not truly transparent at the phonetic level. Unlike the small, decidedly phonetic effects reported for Finnish and Hungarian, Gick et al. (2006) and Ritchart & Rose (2017) report much more salient alternations in Kinande and Moro. In Kinande, the low vowel /a/ is produced with lower F1, and with significantly advanced tongue root in [+ATR] contexts; Gick and colleagues propose that underlying /a/ phonologically alternates with surface [ə]. Even more strikingly, Ritchart & Rose (2017) demonstrate that the vowel once thought to be transparent in Moro, /ə/, is actually two distinct vowels, /ə/ and /ɘ/, which form a contrastive pairing for height harmony, differing both as triggers and targets of harmony. Like in Kinande, the results in Ritchart & Rose’s study support a phonological distinction, and not just a low-level phonetic difference based on harmonic context. These findings support theoretical claims of ‘strict locality’, which demand that all segments within a harmony domain alternate phonetically for harmony (Gafos, 1999; Ní Chiosáin & Padgett, 2001).<sup>4</sup>

Yet, some more recent results cast doubt on the larger claim that all transparency is false. Dye (2015) examines ultrasound and acoustic evidence from Pulaar and Wolof ATR harmony. She finds that both /i/ and /u/ fail to undergo even low-level phonetic alternations based on ATR context in Wolof. Additionally, Szeredi (2016) questions the perceptual significance of the types of low level effects reported in Hungarian and Finnish. His findings

---

<sup>4</sup> Similar results have been reported for some consonant and vowel-consonant harmonies (e.g. Walker, 1999; Walker et al., 2008).

suggest that Hungarian speakers do not perceive the small phonetic differences based on vowel backness context (cf. Benus & Gafos, 2007), arguing rather that other sublexical cues support learning transparent and exceptional patterns in harmony.

### 3.2 Distance effects

In addition to pseudo-transparency, phonological evidence supports an additional, distance-based restriction on transparency. In Hungarian, a single /i/, /i:/, /e/, or /e:/ is transparent (with, as noted above, small phonetic effects) to backness harmony. This is demonstrated in (7) below. In (7a,b), a [+back] root is followed by each of the two diminutive suffixes, which both surface as [i] irrespective of root backness. In (7c,d) both suffixes are followed by the regularly alternating dative suffix, which agrees with the initial-syllable vowel for [back]. Thus, when a single /i/ occurs, it is transparent to harmony. However, in (7e,) we see a different pattern when harmony must span two consecutive transparent vowels. When a regularly alternating suffix is preceded by two transparent vowels (7e), that suffix is far more likely to surface as [-back], regardless of initial vowel backness (Gafos & Dye, 2011; Hayes & Londe, 2006; Ringen & Kontra, 1989; Siptár & Törkenczy, 2000; Vago, 1980)

- (7) Transparency and a count effect (Ringen & Kontra, 1989; Gafos & Dye, 2011)
- |    |                               |                   |
|----|-------------------------------|-------------------|
| a. | mam-i                         | ‘mom-DIM’         |
| b. | mam-tʃi                       | ‘mom-DIM’         |
| c. | mam-i-nak                     | ‘mom-DIM-DAT’     |
| d. | mam-tʃi-nak                   | ‘mom-DIM-DAT’     |
| e. | mam-i-tʃi-nak ~ mam-i-tʃi-nek | ‘mom-DIM-DIM-DAT’ |

This effect is distinct from the low-level phonetic alternations reported in Benus & Gafos (2007). In (7f), the final syllable is output as [a] or [e] (phonetically [ɔ] or [ɛ]), not slight variants of a single category, which suggests a decidedly phonological effect. If these vowels were entirely transparent, it should not matter how many of them intervene between a trigger and target. However, the fact that their number does matter, suggests that they are not entirely transparent, a point made by Ringen & Kontra (1989), Hayes & Londe (2006), as well as Gafos & Dye (2011).

### 3.3 Research questions

While some previous work argues that all ‘transparent’ vowels alternate for harmony (Gafos, 1999; Ní Chiosáin & Padgett, 2001), Dye’s (2015) findings undermine any

simple reanalysis of putative transparency. Further, in some cases these alternations appear to be purely phonetic (Finnish), while in others they are allophonic (Kinande), or even evidence for an additional, unreported vowel contrast (Moro). Thus, the range of actual phonetic and phonological variation discernible among reportedly transparent is not at all clear.

Additionally, data from Hungarian suggests that it may be overly simplistic to categorize elements as categorically transparent or not, since the invisibility of the front unrounded vowels in Hungarian is constrained by distance. For these reasons, data from more languages is necessary to evaluate extant theoretical claims, and to further understand the amount and types of variation attested in these types of harmony patterns.

In addition to the general need to collect more data, Uyghur provides a good test case for experimental examination. Previous research on the language, while cognizant of claims concerning strict locality (Vaux, 2000: footnote 2), argues that the realization of /i/ is unaffected by flanking vowel backness. Thus, Uyghur may present another case like Wolof, where certain vowels are invisible to harmony. Second, reports indicate that transparency persists across multiple vowels, although strings of three or more medial /i/ vowels are never discussed. Given Uyghur's agglutinative morphology and morphotactics, it is possible to produce words with longer sequences of medial /i/ vowels, which offers a fuller comparison against a language like Hungarian, where phonological transparency is sensitive to distance.

Existing work on transparency in Uyghur depends entirely on textual and impressionistic data. This paper, however, investigates the pattern with acoustic data from a production study. In addition, previous work has not examined the reportedly variable behavior of lexical items as a way to examine the Uyghur harmony system. This study uses acoustic data paired with counts of [ $\pm$ back] suffix selection to provide a fuller picture of harmony in the language. Specifically, four questions are addressed. First, do different members of a near-minimal pair select different classes of suffixes? Second, does F2 of root-internal [i]-[u] differ between members of a near-minimal pair? These two questions probe whether there is a relationship between the backness of root-internal [i]-[u] and the backness of following vowels? Third, do non-initial [i] and [u] alternate for harmony? Fourth, do low vowels alternate for harmony when preceded by medial high unrounded vowels? The first and second questions address the behavior of high unrounded vowels in trigger positions. The third and fourth questions address the behavior of high unrounded vowels in target positions. These possibilities are discussed and operationalized in the following section.

## 4 Methods

### 4.1 Stimuli

Participants were presented a set of pictures corresponding to the Uyghur nouns containing the three uncontroversial harmonic pairings in the language, /a-æ, o-ø, u-y/. Pictorial prompts were used to avoid an orthographic confound, since Uyghur orthographies do not represent [u]. Target words were derived from monosyllabic and disyllabic roots, as shown in (8). Monosyllabic roots ended either in a sibilant or a liquid (8a-g), and disyllabic roots contained two vowels that agreed for the feature [high] (8h-n).

(8) Example stimuli

Monosyllabic roots

a.	baʃ	‘head’
b.	bal	‘honey’
c.	bæl	‘waist’
d.	jol	‘road’
e.	køɭ	‘lake’
f.	qul	‘slave’
g.	gyl	‘flower’

Disyllabic roots

h.	bala	‘child’
i.	palta	‘axe’
j.	sællæ	‘turban’
k.	χorma	‘persimmon’
l.	tøpæ	‘hill’
m.	qurum	‘soot’
n.	jyzym	‘grape’

Additionally, stimuli with putative /i/ were drawn from lexical items that either, one, exhibit variation in Hahn (1991) or Lindblad (1990), or two, are cognates with /u/ in closely related languages that maintain contrastive /u/, Kyrgyz and Kazakh. The full set of root-internal /i/ stimuli is shown in Table 2. Two monosyllabic stimuli were selected that are reported to trigger [+back] harmony and have [+back] cognates in Kyrgyz and Kazakh, /qiʃ/ and /jil/. Two monosyllabic stimuli were selected that are reported to trigger [+back] harmony but have [-back] cognates in Kyrgyz and Kazakh, /tʃiʃ/ and /pil/. Finally, two disyllabic target words were selected that reportedly trigger [-back] harmony and correspond to [-back] cognates in Kyrgyz and Kazakh /ilim/ and /jilim/.<sup>5</sup> Table 2 also includes predicted surface quality for each word based on Hahn’s (1991) detailed discussion of /i/ allophony. For the present study, it is relevant to note that Hahn predicts that /i/ is fronted around (alveo) palatals and is backed

---

<sup>5</sup> As a reviewer points out, another logically possible combination is possible: /i/ vowels that trigger [-back] suffixes, although their cognates in related language are [+back]. I do not know of any such examples, nor are any reported in (Lindblad, 1990: ch. 5).

around dorsals and laterals. The elicited words thus exemplify a range of surface qualities, varying in backness from [i] to [ʏ].

Table 2: Stimuli with putative /i/, their predicted phonological and phonetic properties, with cognates in Kyrgyz and Kazakh (indicated by kr and kz, respectively)

Root	Gloss	Phonological status in Lindblad (1990)	Predicted surface quality in Hahn (1991)	Cognates
/qiʃ/	winter	[ + back]	[qʏʃ]	quʃ (kr), quʃ (kz)
/tʃiʃ/	tooth	[ + back]	[tʃɪʃ]	tiʃ (kr), tɪʃ (kz)
/jil/	year	[ + back]	[jɪl]	ɟɪwɪl (kr), ɟɪwɪl (kz)
/pil/	elephant	[ + back]	[pɪl]	pɪl (kr), pɪl (kz)
/ilim/	science	[-back]	[ɪlɪm]	ilim (kr), ɪlɪm ~ wɪlɪm (kz)
/ʃilim/	paste	[-back]	[ʃɪlɪm]	ɟɛlɪm (kr), ɟɪlɪm (kz)

Observe that the stimuli in Table 2 form three near-minimal pairs. Previous work on Uyghur has focused over larger sets of lexical items, devoting almost no discussion to the potential value of pairs such as those in Table 2. In Structuralist and post-Structuralist linguistic work, the use of minimal and near-minimal pairs is a primary heuristic for discovering phonological contrasts. As such, evaluating a small number of pairings, their acoustic characteristics, and their behavior with respect to harmony offers potential insight to complement existing work on the language.

#### 4.2 Task

Each session was divided into training and recording phases. During the training phase, participants were taught a small set of pictorial-lexical correspondences. As an example, a photo of a flower prompted the word /gyl/ ‘flower’ while a photo of lake prompted the word /køl/ ‘lake.’ In addition, participants were also taught a set of pictorial-grammatical correspondences involving number, case, and possession. For instance, a downward red arrow indicated locative case while two outward pointing red arrows indicated ablative case. The training phase typically lasted less than 5 minutes. After participants completed training, the recording phase began. Throughout each session, participants were presented images on a laptop computer screen that showed both a picture representing a lexical item with a grammatical prompt from the training phase. Thus, a picture of a single flower with a

downward pointing red arrow would prompt the word /gyl-dæ/ ‘flower-LOC’. When speakers were unable to guess the target word from the prompt, they were given either the equivalent Russian word or a paraphrase in the target language.

Roots were elicited in four cases (nominative, accusative, locative, and ablative), singular and plural numbers, and in first- and third-person possessive forms. Example inflected forms from the roots /bæl/ ‘waist’ and /bal/ ‘honey’ are shown below in (9). Participants were prompted to produce each word only once. If a speaker produced a particular word several times though, all repetitions were analyzed.

(9) Example elicited forms

		/bæl/	/bal/
a.	NOM	bæl	bal
b.	ACC	bæl-ni	bal-ni
c.	LOC	bæl-dæ	bal-da
d.	ABL	bæl-din	bal-din
e.	PL	bæl-lær	bal-lar
f.	PL-LOC	bæl-lær-dæ	bal-lar-da
g.	POSS.1S	bæl-im	bal-im
h.	POSS.3S	bæl-i	bal-i

I examined three types of high vowel targets, underlying, epenthetic, and raised. Underlying high vowels are present in three of the suffixes elicited, the ablative, accusative, and third-person possessive (9b,d,h). An epenthetic high vowel occurs preceding the first-person singular possessive suffix after consonant-final roots (9g). Third, I examined raised vowels, which surface via the reduction of low vowels in medial open syllables, shown in (6). These raised vowels were either the second syllable of a disyllabic root, or the plural suffix.

There are two basic questions that motivate the paper – what is the relationship between the high unrounded vowel root-internally, and how do the high unrounded vowel participate in harmony. These two questions are, in turn, framed in more specific, testable terms in the four questions in (10). Vowel backness is operationalized as variation in the second formant (F2), since this is the primary acoustic manifestation of varying tongue body backness. Since back vowels exhibit lower F2 than front vowels, for any alternating pair, F2 of a given vowel should be lower in a [+back] context and higher in a [-back] context.

(10) Research questions

Root-internally: what is the relationship between the high unrounded vowels?

1. Do members of the pair trigger different suffix allomorphs?
2. Does F2 of root-internal [i]-[u] differ across members of a near-minimal pair?

Non-initially: how do the high unrounded vowels participate in harmony?

3. Is F2 of non-initial [i]-[u] (underlying, epenthetic, and raised) predictable based on backness of the initial-syllable vowel?
4. Is F2 of low vowels following medial [i]-[u] predictable based on initial-syllable vowel backness?

The lexical items used in the study form three near-minimal pairs, qif-tʃiʃ, jil-pil, and ilim-ʃilim. Using these pairs, it is possible to probe the nature of contrast and the role of adjacent consonants on root-internal [i]-[u]. I evaluate the acoustics of root-internal high unrounded vowels to determine if there exist spectral differences, specifically F2 differences, that are not attributable to consonantal context. In addition, using pairs allows one to connect potential differences in suffix allomorph selection and the acoustics of root-internal [i]-[u].

The first two questions, which concern the nature of root-internal [i] and [u], yield four logically possible interpretations, which are shown in Table 3. If each member of the near-minimal pair triggers the same set of suffix allomorphs, then there is no evidence for a contrast between pair members, and the relationship between [i] and [u] is at most allophonic. If, however, different roots trigger different suffix allomorphs, this would suggest either a transparent (with lexical idiosyncrasy; as in Hungarian) or contrastive relationship between the high unrounded vowels (Moro). Section 5.1 address the behavior of [i] and [u] as triggers of harmony.

Table 3: Four possible relationships between root-internal high unrounded vowels

Interpretation		Do members of the pair trigger different suffix allomorphs?	Does F2 of root-internal [i]-[u] differ across members of the pair?
Non-contrastive	Non-allophonic	No	No
	Allophonic	No	Yes
Transparent/Lexical idiosyncrasy		Yes	No
Contrastive		Yes	Yes

In addition to their contrastive status and behaviors as triggers, Section 5.2 discusses the behavior of the high unrounded vowels as targets of harmony. Do these vowels alternate for backness harmony, and do subsequent vowels alternate for harmony? Four possibilities are noted in Table 4. If these vowels fail to alternate, but allow root backness to propagate to more peripheral suffixes, as reported in previous research, then the high unrounded vowels are transparent. If these vowels fail to alternate, but stop [back] spreading, then they are blockers. As a third possibility, both the high unrounded vowel and subsequent suffixes may alternate for harmony. This state of affairs would suggest that these vowels are not transparent, but alternate for harmony. Finally, it is possible that these vowels alternate for harmony but prevent backness from spreading to subsequent suffixes. Jurgec (2011) calls this type of vowel an “icy target.”

Table 4: Four possible behaviors of the high unrounded vowels as targets of harmony

	Is F2 of non-initial [i]-[u] predictable based on the backness of the initial-syllable vowel?	Is F2 of low vowels following medial [i]-[u] predictable based on initial-syllable vowel backness?
Transparent	No	Yes
Blocking	No	No
Alternate	Yes	Yes
Icy target	Yes	No

To examine the behavior of non-initial high unrounded vowels, elicited words fell into one of two conditions, exhibiting either a short- or long-distance dependency (11). The short-distance condition involved only a single high unrounded vowel between trigger and

alternating target vowel. In contrast, the long-distance condition involved three high unrounded vowels intervening between trigger and alternating target vowel. If the alternation of low vowel suffixes varies across these two conditions, then transparency is distance-based, as in Hungarian. Speaker 1 did not participate in the short- and long-distance conditions, producing data only for the root-internal analysis in Section 5.1.

(11) Short- and long-distance conditions (transcriptions based on previous work)

Short-distance

- a. bæl-i-dæ ‘waist-POSS.3S-LOC’
- b. tʃyʃ-i-dæ ‘dream-POSS.3S-LOC’
- c. bæl-im-dæ ‘waist-POSS.1S-LOC’
- d. tʃyʃ-ym-dæ ‘dream-POSS.1S-LOC’

Long-distance

- e. sælli-lir-i-dæ ‘turban-PL-POSS.3S-LOC’
- f. tʊpi-lir-i-dæ ‘hill-PL-POSS.3S-LOC’
- g. palti-lir-i-da ‘axe-PL-POSS.3S-LOC’
- h. χormi-lir-i-da ‘persimmon-PL-POSS.3S-LOC’<sup>6</sup>

Sessions were conducted in a quiet room. Participants wore a Shure-SM10A unidirectional head-mounted microphone, and all data were recorded to a Marantz PMD 661 MKII digital recorder at a sampling rate of 44.1 kHz. Each session lasted between 45 and 90 minutes.

#### 4.3 Participants

Participants were recruited through existing relational networks in Chunja, Kazakhstan. Speaker participation and informed consent were obtained in accordance with University of California San Diego Linguistic Fieldwork IRB protocol #141520. Nine Uyghur speakers (five females; mean age: 44.4 years; range: 19-63 years) participated in the study. All speakers were from Chunja or immediately surrounding villages. All participants reported native fluency in the target language, as well as fluency in Kazakh and Russian. When asked which of the three languages they speak best, Speakers 1, 2, 4, and 9 reported Uyghur; Speaker 6 reported Russian;

---

<sup>6</sup> Participants produced both /χurma/ and /χorma/ for ‘persimmon’. The spectral differences between [u] and [o] were not large in the collected data, and there are reports of initial-syllable /o/ raising to [u] when followed by high vowels in some Uyghur dialects (Yakup, 2005:63-64; Abdurehim, 2014:82).

and Speakers 3, 5, 7, and 8 reported that they speak all three equally well. This is unsurprising in Chunja, the seat of the Uyghur district in the Almaty region. In Chunja, Uyghur is spoken and written in public, and used in more domains than elsewhere in Kazakhstan, where proficiency in Uyghur is often secondary to proficiency in another language. According to the 2009 national census, 83.8% of Uyghurs indicated the Uyghur was their mother tongue. For comparison, 93.7% of Uyghurs reported fluency in Kazakh, and 95.8% reported fluency in Russian (Smajylov, 2011). The proficiency of Kazakhstani Uyghurs in Kazakh and Russian differentiates them from those living in China, who are often bilingual in Uyghur and Mandarin, with Mandarin encroaching in many linguistic domains.

Since almost all theoretical work on Uyghur phonology has focused on the standard variety, it is important to note that the variety spoken in Kazakhstan is part of the central dialect region, upon which the standard language is based (Kaydarov, 1970; Hahn, 1986:36-42, 1991:5-6; Yakup, 2005:8-22). Further, previous research indicates that one of the traits that Kazakhstani Uyghur shares with the standard variety, which has been the focus of most theoretical work, is the behavior of /i/ (Baskakov, 1970; Kaydarov, 1970; Yakup, 2005:8). Therefore, results reported below should be generally relevant for existing theoretical proposals concerning harmony and transparency in Uyghur.

#### 4.4 Segmentation

All sound files were segmented in Praat (Boersma & Weenink, 2015). The beginning and end of each vowel was set to the onset and offset of the second formant. In cases where the second formant persisted across flanking consonants, abrupt changes in energy or formant frequencies were used to indicate vowel onset and offset.

#### 4.5 Statistical analysis

After segmentation, the second formant was measured at vowel midpoint. To facilitate across-speaker comparisons, the data were z-score normalized (Lobanov, 1971). Four tokens of /a o u æ ø y/ were used for normalization. In addition, eight tokens of the high unrounded vowels, four tokens with relatively high F2 and four with relatively low F2 were included for normalization.

Outliers were inspected for measurement errors. A number of errors were found with [u], where the formant tracker in Praat failed to distinguish the first two formants. In these cases formant frequencies were hand measured at the approximate vowel midpoint. The data were analyzed in R (R Core Team, 2017), using the lme4 package (Bates et al., 2015). A mixed effect linear regression was used to predict normalized F2 at vowel midpoint for both high

unrounded vowels and subsequent alternating vowels. Significance was assessed using likelihood ratio tests.

## 5 Results

### 5.1 Root-internal high unrounded vowels

The first question to address is whether members of each near-minimal pair exhibit different patterns of suffix allomorph selection. Aggregate counts of suffix selection (n = 296) are shown in Table 5. Interestingly, roots with [+back] cognates in Kyrgyz and Kazakh, /jil/ and /qiʃ/, triggered [+back] suffixes 100% of the time, whereas roots with [-back] cognates in those related languages exhibited significant variation. Note however that every root tended to trigger [+back] suffixes, with /ilim/ being the most likely to precede a [-back] suffix. Observe also that there is no obvious correlation between the surface forms predicted in Hahn (1991) and the backness of the following suffix.

Table 5: Root-internal high unrounded vowels and suffix backness

Root	Gloss	Phonological status in Lindblad (1990)	Predicted surface quality in Hahn (1991)	Cognates		Count [+back] suffix	Percent [+back] suffix
				Kyrgyz	Kazakh		
/qiʃ/	winter	[+back]	[qʁʃ]	qʁʃ	qʁs	47/47	100
/tʃiʃ/	tooth	[+back]	[tʃɪʃ]	tɪʃ	tɪs	43/56	77
/jil/	year	[+back]	[jɪl]	ɟʒul	ʒul	45/45	100
/pil/	elephant	[+back]	[pɪl]	pɪl	pɪl	32/43	74
/ilim/	science	[-back]	[ɪlɪm]	ilim	ɪlɪm ~ ɪwulum	28/48	58
/ʃilim/	paste	[-back]	[ʃɪlɪm]	ɟʒelim	ʒielɪm	45/57	79

Comparisons within each pair in Table 5 do, however, suggest some potential phonological differences – /jil/ always behaves like a [+back] vowel, while /pil/ does so only 74% of the time; /ʃilim/ behaves like a back vowel 79% of the time, while /ilim/ does so only 58% of the time. Given the variation in Table 5, it is necessary to determine if these patterns

generalize across all nine speakers, or if suffix selection is affected by significant inter-speaker variation. Table 6 presents suffix selection counts for each speaker and root, indicating that suffix selection does vary dramatically across speaker and lexical root. For example, Speakers 1, 2, 8, and 9 produce only [+back] suffixes after /ilim/, while speakers 4, 6, and 7 produce only [-back] suffixes after this root. Observe also variation for /tʃiʃ/, /pil/, and /ʃilim/, which always precede [+back] suffixes for some speakers, while these roots may precede [+back] or [-back] roots for some speakers.<sup>7</sup>

Table 6: Counts of [+back]/total suffixes for each root and speaker. Light grey boxes indicate  $50\% \leq [+back] \text{ suffixes} < 100\%$ ; darker grey boxes indicate  $[+back] \text{ suffixes} < 50\%$ .

	1	2	3	4	5	6	7	8	9	Total
qiʃ	5/5	6/6	6/6	6/6	5/5	6/6	3/3	5/5	5/5	47/47
tʃiʃ	3/7	11/12	6/8	6/6	2/5	4/6	4/4	4/5	3/3	43/56
jil	4/4	6/6	7/7	4/4	3/3	6/6	5/5	5/5	5/5	45/45
pil	3/3	5/7	2/7	2/5	4/4	6/6	3/4	5/5	2/2	32/43
ilim	6/6	2/2	6/8	0/4	5/6	0/9	0/4	6/6	3/3	28/48
ʃilim	4/6	9/12	0/0	6/6	3/5	5/10	6/6	6/6	6/6	45/57
Total	25/31	39/45	27/36	24/31	22/28	27/43	21/26	31/32	24/24	240/296

Results in Table 6 suggest that the phonological behavior of these lexical items is speaker-dependent. For Speakers 8 and 9, there is no evidence for any phonological difference between these six roots. Comparisons between each near-minimal pair for Speaker 2 also fail to suggest any phonological distinctions – all pairs show the same basic tendency for roots to trigger [+back] suffixes. However, results for Speaker 1 and Speakers 3-7 suggest a different state of affairs for these speakers.

For Speakers 1 and 5, /tʃiʃ/ triggers [-back] suffixes more often than [+back] suffixes, in contrast to /qiʃ/, which always triggers [+back] suffixes. Despite a difference in suffix patterning here, previous research has argued that the dorsal consonants act as triggers of harmony, causing any root with a uvular to behave as phonologically [+back] (Hahn, 1991; Mayer & Major, 2018). As such, it is not possible to determine whether the vowel or the uvular consonant acts as the trigger of harmony here.

<sup>7</sup> No tokens occurred for Speaker 3 because she produced the Russian word /klʲej/ rather than /ʃilim/ ‘paste’.

Speakers 3 and 4 produce [-back] suffixes after /pil/ in a majority of cases, but always produce [+back] suffixes after /jil/. Most intriguing, though, is the comparison between /ilim/ and /jilim/. Speakers 4 and 7 produce only [-back] suffixes after /ilim/ and only [+back] suffixes after /jilim/. Speaker 6 also produces only [-back] suffixes after /ilim/, but suffix backness is variable after /jilim/. These results suggest that these roots are treated differently in across speakers. Further, these data support a possible contrast between the vowels in /pil/ and /jil/ for Speakers 3 and 4, and a possible contrast between the vowels in /ilim/ and /jilim/ for Speakers 4, 6, and 7.

If these patterns of suffix backness selection are evidence for contrast, acoustic differences should exist between these particular roots for these particular speakers. Before examining speaker-specific acoustic data, observe the aggregate patterns for F2 for each lexical root in Figure 1. In this violin plot, the vertical length and shape of each violin represents the distribution of F2 for each lexical root. Horizontal lines within each violin indicate median and interquartile range for each root. F2 values range from less than -1z, which is comparable to a very posterior [u] to more than 2z, a value consistent with a very peripheral [i]. While the range of effects is similar to Hahn's description of high unrounded vowel allophony, a more detailed picture is evident below. The realization of these vowels is more affected by the following vowel than the preceding vowel, with a following lateral resulting in a more backed vowel, even in a root like /jil/, with a preceding palatal.<sup>8</sup> In general, though, values tend to fall between a 0.5 and -0.5z. In addition to the global distribution of F2, there is noteworthy variation between lexical roots. Observe that the highest F2 is associated with the root, /tʃiʃ/ 'tooth' while the lowest F2 is associated with /jil/ 'year', /pil/ 'elephant', and /jilim/ 'paste.'

Figure 1 presents aggregate patterns of suffix backness selection. If a general contrast existed for all six lexical items for all speakers, then we would expect the leftmost roots to exhibit the highest percentage of [+back] suffixes, and the rightmost roots to exhibit the highest percentage of [-back] suffixes. No such general contrast exists, and no obvious phonetic correlation between F2 and suffix backness is evident in the aggregated data.

---

<sup>8</sup> In a regression predicting F2 from preceding and following consonantal effects alone, following consonant was the only significant predictor, with vowels differing by 0.75z based on following context, (alveo)palatal vs. bilabial or lateral [ $\chi^2(1) = 7.37, p < .01$ ].

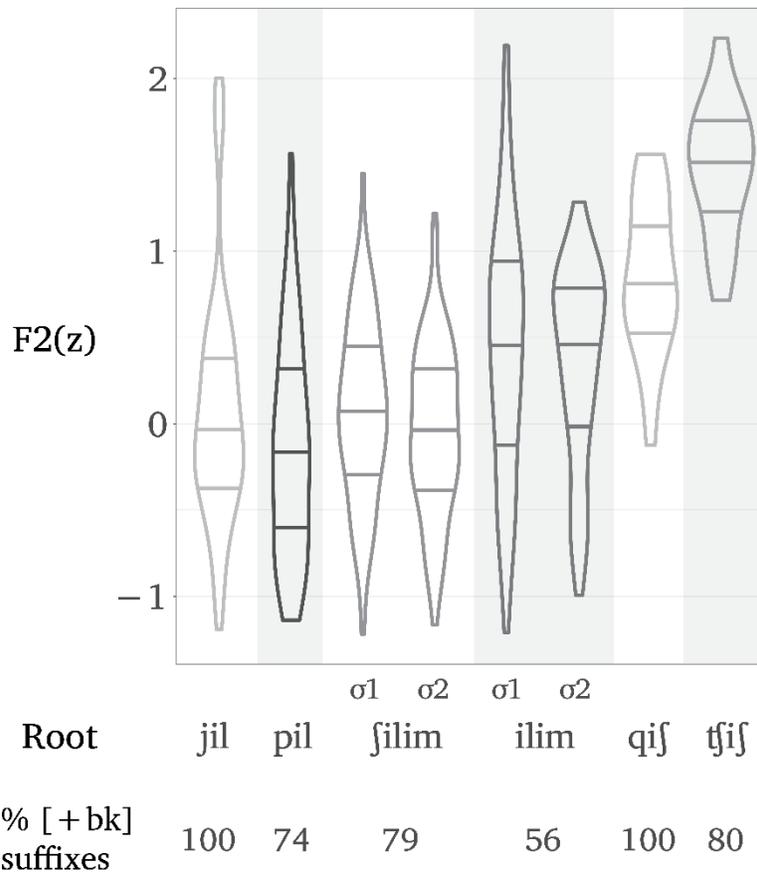


Figure 1: F2 (z) of root internal /i/ (n = 631) by lexeme along with percent [+back] suffixes during elicitation. Within each distribution, the horizontal lines represent the median and interquartile range.

By-speaker boxplots showing variation across the six lexical items are presented in Figure 2. Since no between-speaker comparisons are made below, F2 is presented in Hertz. Observe across all nine speakers the relative paucity of tokens from /qif/ ‘winter’ and /tif/ ‘tooth.’ Initial-syllable high vowels often elide before voiceless fricatives in the language. For some speakers, the high vowel was frequently elided (e.g. Speakers 3, 4, and 9; others have also reported devoicing in this context), while for others the vowel was produced with slightly more consistency (e.g. Speakers 7 and 8). Speaker age and gender are noted in all by-speaker plots.

In order to determine which vowels of the three near-minimal pairs exhibit significant F2 distinctions for each speaker, a model was constructed to predict F2 (Hz) from speaker, lexical root, and their interaction, along with random intercepts for speaker and root. From this model, post-hoc comparisons were implemented in the emmeans package with Bonferroni correction for multiple comparisons (Lenth et al., 2018). The comparisons noted above – /pil/

and /jil/ for Speakers 3 and 4, and /ilim/ and /ʃilim/ for Speakers 4, 6, and 7 – were tested. The roots /tʃiʃ/ and /qiʃ/ were not considered due to the fact that /q/ actively triggers harmony in the language, as well as the frequent elision of these vowels before /ʃ/. If these pairings are truly contrastive for these speakers, then I predict an acoustic difference between these items, specifically higher F2 for those items more frequently triggering [-back] suffixes.

Table 7: Post-hoc comparisons for potential speaker- and pair-specific contrasts (Bonferroni-adjusted  $\alpha = .01$ )

Pair	Speaker	Estimate	SE	$z$	$p$
/pil/-/jil/	3	-299	139	-2.15	.03
	4	259	169	1.53	.13
/ilim/-/ʃilim/	4	398	130	3.05	.002
	6	301	120	2.51	.01
	7	55	125	0.44	.66

As for /pil/ and /jil/, F2 for Speaker 3 trends toward a significant difference, but in the opposite direction of predictions. If the vowels in /pil/ and /jil/ behave like a front and back vowels respectively, one would expect F2 of the vowel in /pil/ to show higher F2. Yet, for Speaker 3 F2 for /pil/ is actually lower than F2 for /jil/ [ $\beta = -299, z = -2.15, p = .03$ ]. The acoustic realization of the vowels in these two roots does however conform to predictions for Speaker 4. For Speaker 4, /pil/ triggers [-back] suffixes 60% of the time, while /jil/ always triggers [+back] suffixes. Based on Table 7, the vowel in /pil/ is produced with around 250 Hz greater F2 than the vowel in /jil/, although this difference does not reach significance [ $\beta = 259, z = 1.53, p = .13$ ]. Note that this phonetic difference runs counter to very general tendencies we would expect from coarticulation, since /p/ would likely depress F2 while /j/ would increase it, all else being equal. No other speakers produced /pil/ with a more anterior vowel quality than /jil/.

Moving on to /ilim/ and /ʃilim/, recall that Speakers 4, 6, and 7 always produced [-back] suffixes after /ilim/. Speakers 4 and 7 always produced [+back] suffixes after /ʃilim/, and Speaker 6 did so 50% of the time. There is a significant difference between the vowels in /ilim/ and /ʃilim/ for Speaker 4, and the difference for Speaker 6 trends toward significance [Speaker 4:  $\beta = 398, z = 3.05, p = .002$ ; Speaker 6:  $\beta = 301, z = 2.51, p = .01$ ]. Finally, for Speaker 7 there is no meaningful acoustic difference between the vowels of /ilim/ and /ʃilim/ [ $\beta = 54, z = 0.44, p = .66$ ].

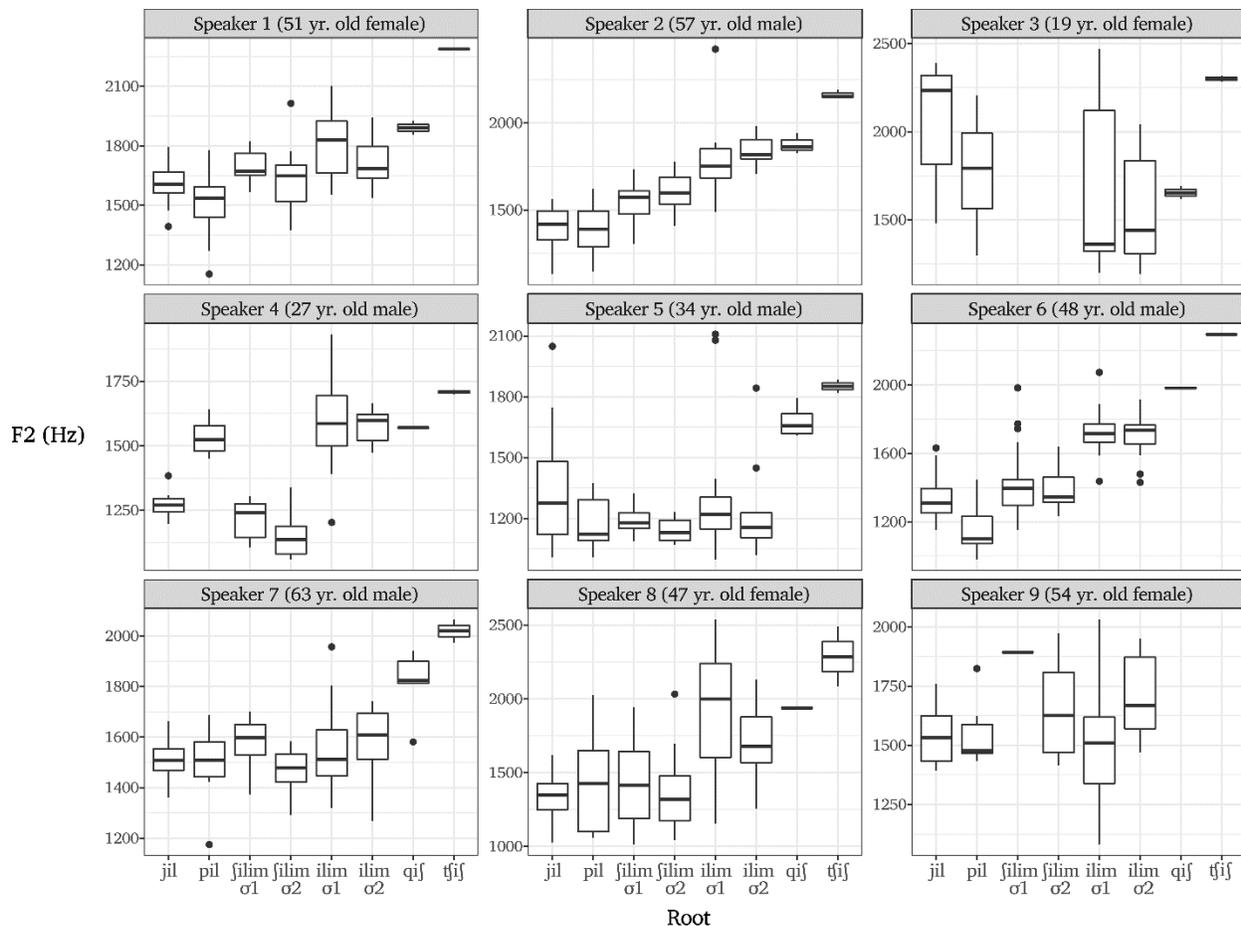


Figure 2: F2 (Hz) of root internal high unrounded vowels by lexeme and syllable number for each speaker (n = 631)

The results above suggest that the status of the high unrounded vowels varies significantly between speakers. Speakers 8 and 9 appear to treat all high unrounded vowels equivalently, triggering [+back] suffixes across all roots tested. Speaker 4 however, shows a pattern that supports a phonological contrast for /ilim/ and /filim/, which would be better transcribed as /ilim/ and /ɣulum/, since they trigger front and back suffixes, respectively, and are marked by significantly different F2. Data from Speaker 6 is also suggestive in this regard, although /ɣulum/ optionally triggers [-back] suffixes for this speaker, and the acoustic differences between the vowels of these two words only trend toward significance. Importantly, differences in F2 correspond to phonological patterning for these speakers – roots that trigger [-back] vowels are produced with higher F2. Other speakers, however, produced data that are consistent with previous descriptions, specifically claims of lexical idiosyncrasy. For instance, Speaker 7 produces [-back] suffixes after /ilim/ but [+back] suffixes after /filim/ with no

reliable acoustic differences for the vowels in these two roots. It is reasonable to tentatively conclude that for those speakers each lexical item is specified for a particular harmonic value independent of surface [i]-[u]. If there is only one contrastive high unrounded vowel, then there is more freedom along the F2 dimension for variation. The extent of variation seen in Figure 1 is potentially consistent with a general lack of and /i/-/u/ contrast, although this variation appears slightly less rampant in the by-speaker plots in Figure 2. As an example, even for Speaker 4, whose production data best supports a contrast between /i/ and /u/, /qi/ and /tʃi/ have high F2 but behave like [+back] roots.<sup>9</sup>

Evidence thus far points to three types of systems: (1) those where /i/ and /u/ contrast, (2) those where [i] and [u] do not contrast, and roots exist with idiosyncratic harmonic specifications independent of surface [i]-[u], and (3) those where [i] and [u] do not contrast and all suffixes after [i] and [u] are always [+back]. As a final note before moving on, if one posits that there is only one high unrounded vowel phoneme in the language, it is not at all clear that it should be /i/ (with a default surface quality of [ɪ]) and not /u/. Central and back variants are much more common, occur with a much freer distribution than [i] (see also Hahn 1986 for discussion), and [i]-[u] most often trigger [+back] suffixes. As support for relative backness of most high unrounded vowels, observe the vowel plot in Figure 3, which shows mean F2 of the vowels in these roots in the context of the larger vowel inventory. For clarity, ellipses are omitted around roots with [i]-[u], as their variation for F2 is evident in the figures above. Observe that F2 for most of these six roots falls in between the front and back rounded vowels. Although Hahn (1991) suggests the default realization of /i/ is [ɪ], for these roots the default is more likely a more central vowel, like [ə] or [ɘ].<sup>10</sup> Additionally, if one posits /u/ rather than /i/, then the tendency to trigger [+back] suffixes receives a more straightforward explanation, and is no less consistent with the relative backness of these roots in Figure 3.

---

<sup>9</sup> It is possible that consonantal influence on [i]-[u] is not allophonic, but contrast neutralizing, masking the true harmonic value of each vowel in the lexemes tested. It would take a much larger corpus of data, with an extensive range of root-internal high unrounded vowels to investigate this possibility.

<sup>10</sup> A few other comments on Figure 3 are deserved. The increased F1 of /qi/ is also reported in the literature (Hahn 1986, 1991). Additionally, the vowel transcribed here as /æ/ is known to vary between [æ] and [ɛ]. For the two back rounded vowels, observe the significant overlap in their distributions. This is reminiscent of the back rounded vowels in Kazakh (McCollum 2018). As in Kazakh, the durations of /u/ and /o/ differ substantially. Mean duration of root-internal /u/ was 65 ms. (SD = 29), and mean duration of root-internal /o/ was 109 ms. (SD = 38).

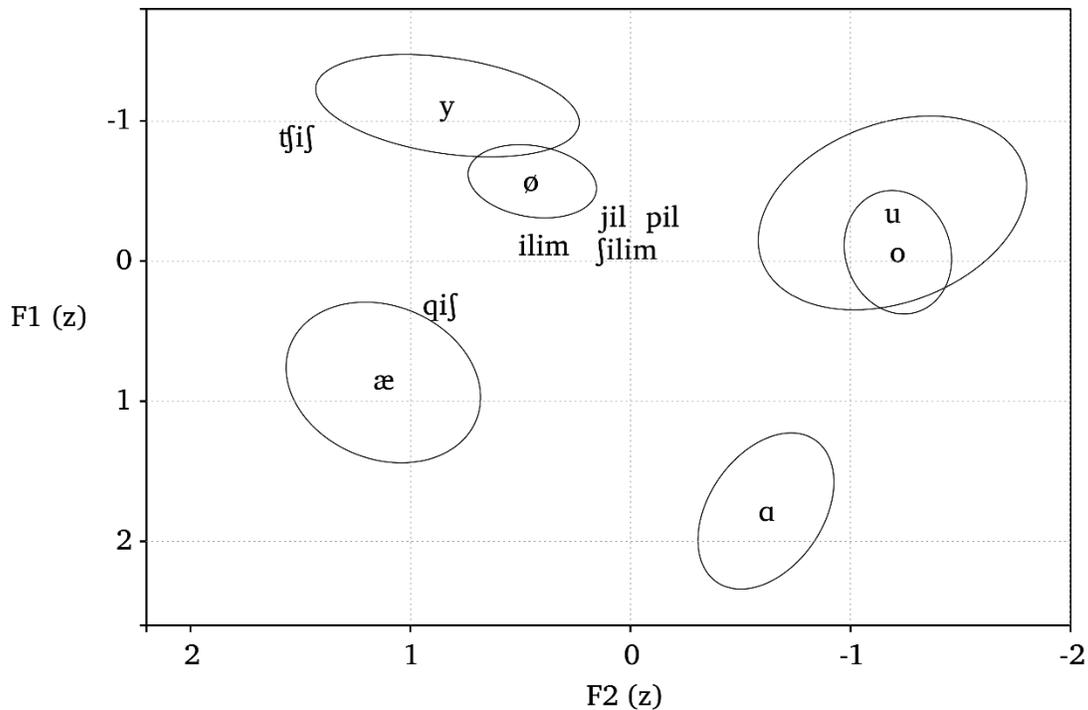


Figure 3: Mean F1-F2 (z) centroid values for root-internal high unrounded vowels compared to other root-internal vowels (n = 2,975; with 1 SD ellipses)

## 5.2 Suffixal and raised /i/

This subsection investigates whether or not high unrounded vowels alternate for harmony in target positions (e.g. in suffixes and raised vowel contexts). To that end, the subsection is broken up into two parts. The first part examines results from the short-distance condition, where a single high vowel intervenes between a trigger and alternating target low vowel. The second examines results from the long-distance condition, where three high vowels intervene between a harmony trigger and alternating low vowel target.

### 5.2.1 Short-distance condition

In this condition, second-syllable high vowels from three-syllable words were compared to determine whether or not F2 covaries with initial-syllable backness. As a reminder, example forms from (11a-d) are repeated in (12).

(12) Example forms for the short-distance conditions (transcriptions based on previous work)

- a.    bæl-i-dæ       ‘waist-POSS.3S-LOC’
- b.    tʃyʃ-i-dæ      ‘dream-POSS.3S-LOC’
- c.    bæl-im-dæ     ‘waist-POSS.1S-LOC’
- d.    tʃyʃ-ym-dæ    ‘dream-POSS.1S-LOC’

Since previous work reports that the epenthetic vowel of POSS.1s /-m/ alternates for rounding, and partially for backness, [-im]~[-ym]~[um] (12c,d), while the underlying high vowel of POSS.3s /i/ and raised vowels do not (12a,b), the model incorporated a three-way distinction in target type (epenthetic, underlying, raised). Target types were treatment coded, with the epenthetic vowel of POSS.1s serving as the default. The model thus predicts normalized F2 at vowel midpoint from root backness, target type, and their interaction, along with a random intercept for speaker.

In Table 8, observe that F2 of the epenthetic vowel in POSS.1s is significantly lowered after back vowels [ $\beta = -1.50$ ,  $\chi^2(1) = 192.29$ ,  $p < .001$ ]. This is consistent with previous descriptions of the language. In addition, the two interaction terms also indicate that the underlying high vowel of POSS.3s and raised vowels are also realized with significantly lower F2 in back contexts. Since the interaction of backness and POSS.3s is non-significant, the model suggests that underlying /i/ should be treated the same as the epenthetic vowel for harmony [ $\beta = 0.02$ ,  $\chi^2(1) = 0.01$ ,  $p = .92$ ]. For raised vowels, the effect of initial vowel backness is even larger, as evident in the interaction term, indicating an even larger acoustic difference based on harmonic context [ $\beta = -0.28$ ,  $\chi^2(1) = 5.66$ ,  $p = .02$ ].

These generalizations are observable in Figure 4, where it is clear that high vowel F2 depends on initial vowel backness. In addition, the magnitude of this difference is consistent with those of other vowels in the language. Here, F2 of /i/ in [-back] context ranges from around 0.6 to around 1.2z; this is the same range of F2 value for /æ ø y/ in Figure 3 above. Further, in [+back] contexts, F2 of /i/ ranges from around -0.9 to -0.4z. These values fall in between F2 values for root-internal /a o u/ shown in Figure 3. The size of this effect thus supports the conclusion that the high unrounded vowels do, in fact, alternate for harmony in the short-distance condition. This conclusion is further supported by the uniformity of the across speakers in Figure 5.

Table 8: Regression model output for high vowels in the short-distance condition (n = 285)

	Estimate	SE	$\chi^2$	<i>p</i>
Intercept	0.61	0.07	27.90	< .001
Backness	-1.50	0.09	192.29	< .001
POSS.3s	0.53	0.10	27.88	< .001
Raised	0.50	0.09	30.65	< .001
Backness: POSS.3s	0.02	0.15	0.01	.92
Backness: Raised	-0.28	0.12	5.66	.02

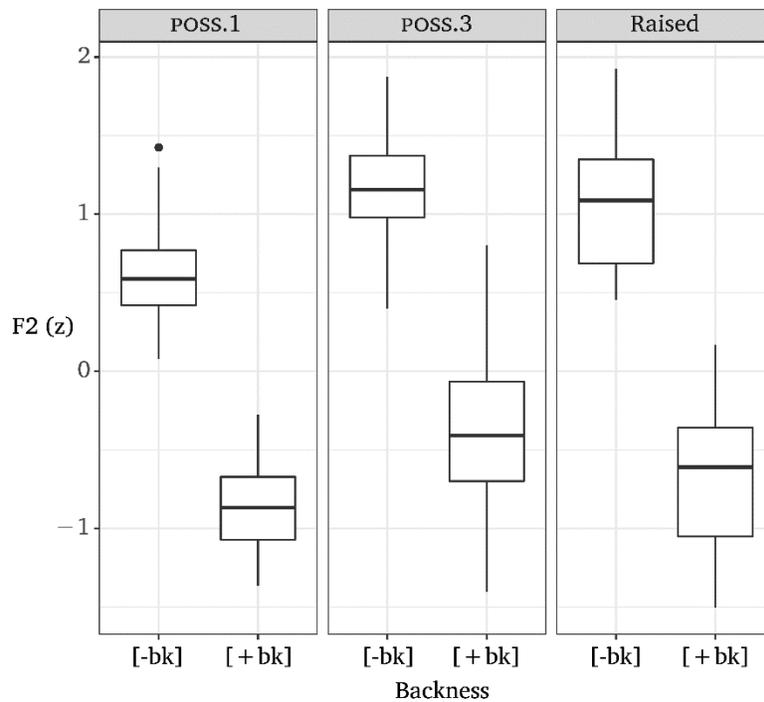


Figure 4: By-target plot of F2 (z) of /i/ based on root backness in the short-distance condition

Moving on, it is now necessary to determine whether or not regularly alternating suffixes, here the locative ([-dɑ]~[-dæ]) and plural ([-lɑr]~[-lær]) suffixes, alternate for harmony after a high vowel. The model predicted low vowel F2 from initial vowel backness and suffix. Suffixes were treatment coded, with locative serving as the default value.

Model output in Table 9 shows that initial vowel backness exerts a significant lowering effect on F2 of the low vowel of the locative suffix [ $\beta = -0.97$ ,  $\chi^2(1) = 128.22$ ,  $p < .001$ ]. Since the effect for the plural morpheme and the interaction between backness and plural are both non-significant, the behavior of both low-vowel morphemes is equivalent – they both undergo

harmonic alternations in the short-distance condition [Plural:  $\beta = -0.20$ ,  $\chi^2(1) = 2.90$ ,  $p = .09$ ; [Backness:Plural:  $\beta = -0.16$ ,  $\chi^2(1) = 1.29$ ,  $p = .26$ ]. This is also evident in Figure 6, as F2 of each suffix's low vowel is substantially lowered after [+back] vowels. The size of the acoustic shift in Figure 6 is again similar to the size of the difference between the low vowels within roots, supporting the phonological nature of this alternation. When Figures 4 and 6 are compared, the behavior of /i/ parallels that of /æ/ and /a/; there is a clear shift between vowels qualities that is consistent with a phonological backness alternation.

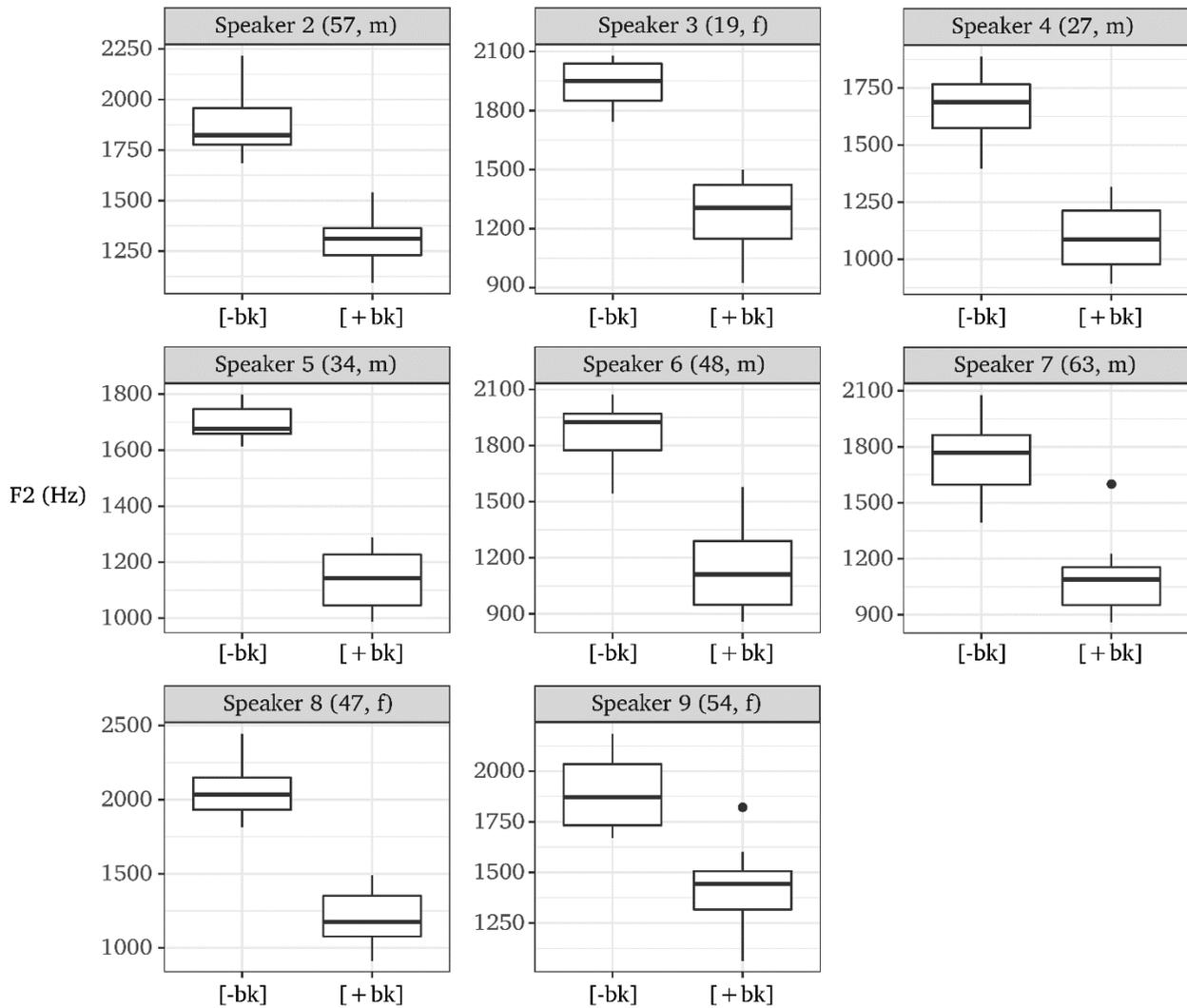


Figure 5: By-speaker plots of F2 (Hz) based on root backness in the short-distance condition. Speaker 1 did not participate in this portion of the experiment (n = 285).

Table 9: Regression model output for low vowels in the short-distance condition

	Estimate	SE	$\chi^2$	<i>p</i>
Intercept	1.01	0.06	36.38	< .001
Backness	-0.97	0.07	128.40	< .001
Plural	-0.20	0.12	2.90	.09
Backness:Plural	-0.16	0.15	1.29	.26

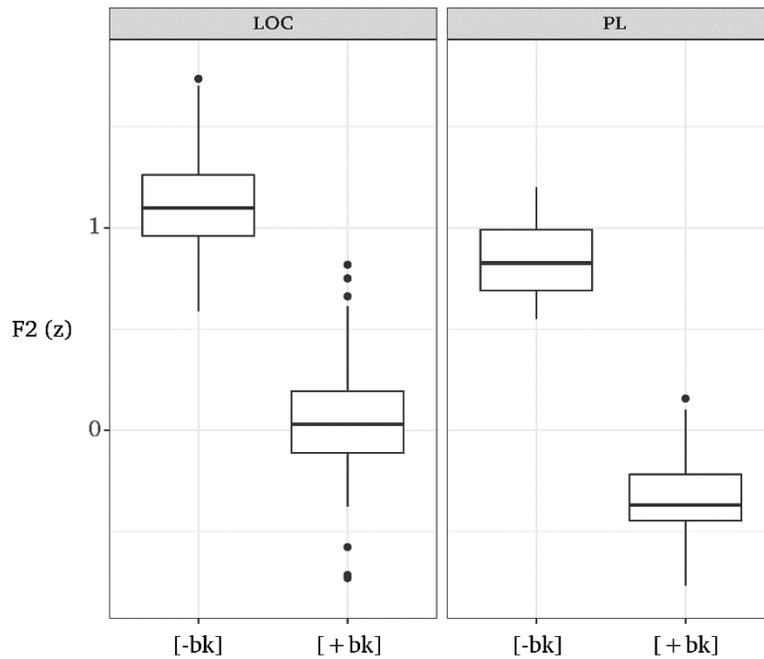


Figure 6: By-suffix plot of F2 (z) based on root backness in the short-distance condition

To summarize, the key finding from this subsection is that high vowels alternate for backness. Given that the high unrounded vowels alternate for backness harmony, it is unsurprising that following low vowels also alternate for harmony. In short, results from the short-distance condition support an analysis whereby both high and low vowels regularly undergo backness harmony.

### 5.2.2 Long-distance condition

In this condition, the medial three syllables of five-syllable words consisted of high vowels derived from vowel raising as well as an underlying high vowel. The raised vowels occurred in syllables two and three, and the underlying high vowel of the third-person singular

possessive suffix, POSS.3s, occurred in the fourth syllable. Example forms from (11) are repeated in (13).

(13) Example forms for the long-distance conditions (transcriptions based on previous work)

- a. sælli-lir-i-dæ ‘turban-PL-POSS.3S-LOC’
- b. tøpi-lir-i-dæ ‘hill-PL-POSS.3S-LOC’
- c. palti-lir-i-da ‘axe-PL-POSS.3S-LOC’
- d. χormi-lir-i-da ‘persimmon-PL-POSS.3S-LOC’

As in Section 5.2.1, a linear mixed effects model was used to predict normalized F2 at vowel midpoint from initial-vowel backness. An additional fixed effect of syllable number was included in the model to determine if the third- and fourth-syllable high vowels in these sequences pattern like the second-syllable high vowel. Since the patterning of high vowels in Hungarian categorically depends on their number, syllable number was treated as a categorical rather than linear predictor. Syllable number was treatment-coded, with the second syllable serving as the default. The model also included interactions between backness and syllable number, along with a random intercept for speaker.

As in the short-distance condition, F2 of high vowels in the second syllable was significantly lowered after [+back] roots [ $\beta = -1.96$ ,  $\chi^2(1) = 353.44$ ,  $p < .001$ ]. Syllables 3 and 4 each also exerted an influence on vowel F2, depressing F2 in both front and back vowel contexts [Syllable 3:  $\beta = -0.26$ ,  $\chi^2(1) = 10.98$ ,  $p < .001$ ; Syllable 4:  $\beta = -0.34$ ,  $\chi^2(1) = 15.78$ ,  $p < .001$ ]. Interestingly, initial vowel backness interacted significantly with position, with F2 in [+back] contexts significantly increasing in third- and fourth-syllable vowels [Backness:Syllable 3:  $\beta = 0.98$ ,  $\chi^2(1) = 82.61$ ,  $p < .001$ ; Backness:Syllable 4:  $\beta = 1.26$ ,  $\chi^2(1) = 104.75$ ,  $p < .001$ ]. In essence, the high unrounded vowels are realized with higher F2, and presumably with less posterior lingual positions, in syllables three and four.

Table 10: Regression model output for high vowels in the long-distance condition (n = 294)

	Estimate	SE	$\chi^2$	<i>p</i>
Intercept	1.00	0.08	30.89	< .001
Backness	-1.96	0.07	353.44	< .001
Syllable 3	-0.26	0.08	10.98	< .001
Syllable 4	-0.34	0.08	15.78	< .001
Backness:Syllable 3	0.98	0.10	82.61	< .001
Backness:Syllable 4	1.26	0.11	104.75	< .001

The asymmetric fronting effect is manifest in Figure 7, where F2 varies by both backness and syllable number. Generally, F2 of both back and front vowels shifts toward a more central value. Significantly, the magnitude of the back vowel shifts far exceeds that of the front vowels. This conforms to the general pattern of asymmetric fronting of the Uyghur vowel space reported in McCollum (2019a,b).

In McCollum (2019a,b), all back vowels in Uyghur are reported to undergo stepwise fronting in non-initial syllables while front vowels exhibit no general shift in the vowel space. McCollum’s results for /ɑ-æ/ and /u-y/ are presented in Figure 8. The key takeaway from these plots is that the F2 distinctions between each pair diminish in later syllables. Further, this reduction of acoustic backness distinctions occurs via the fronting of the back vowels and not a symmetrical centralization/reduction of the vowel space.

Observe the F2 values across each pairing in Figure 8. The [-back] variants of /i/ are as front as phonologically front vowels, and the [+back] variants of /i/ are as back as /ɑ/. There are two additional points worth making here. Figure 8 does not contain any root-internal [i]-[ɯ], thus the absence of first-syllable vowels in the middle panel. All second-syllable vowels are from raised low vowels or suffixal high vowels. Second, the middle panel of Figure 8 presents fifth-syllable /i/, as well. These data points come from the ablative suffix /-din/ in five-syllable words. These were not included in the analysis shown above, but serve to further demonstrate the fronting pattern, and the alternation between [i] and [ɯ].

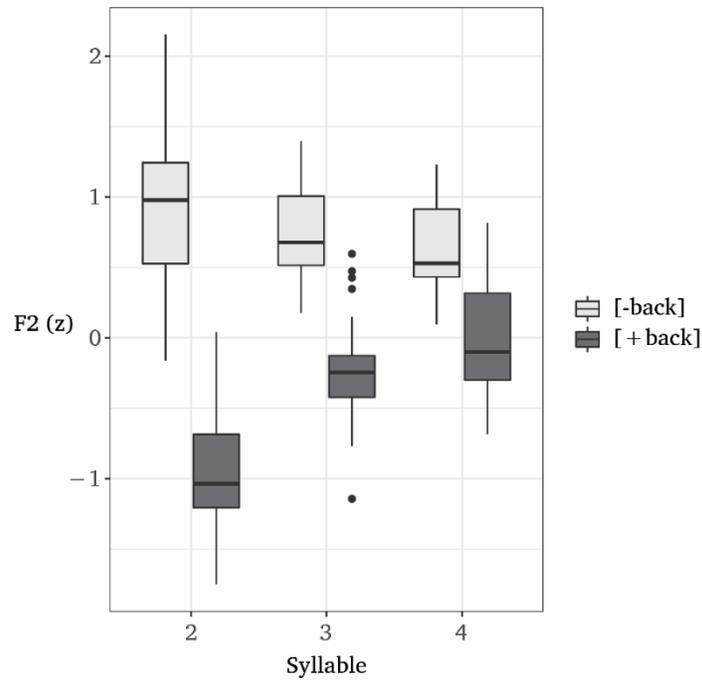


Figure 7: F2(z) of high vowels by root backness and syllable in the long-distance condition

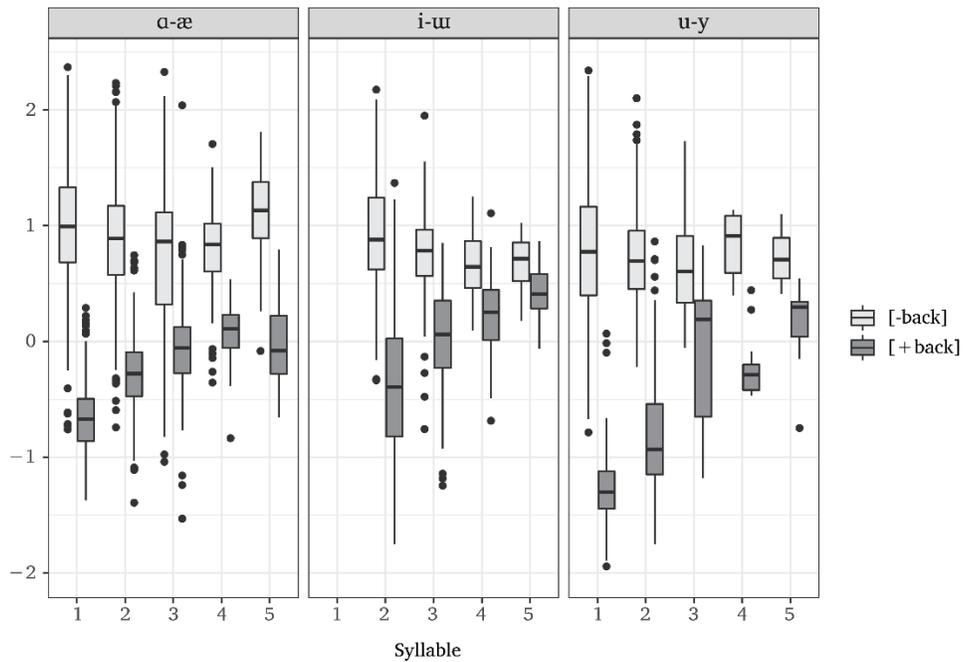


Figure 8: F2(z) for /a-æ/ and /u-y/ from McCollum (2019a)

Individual results for each speaker are presented in Figure 9. Observe that for all speakers, the high vowels maintain an acoustic backness distinction across all syllables. Despite this generalization, individual variation is noteworthy. Some speakers produce relatively

symmetrical centralization across the three relevant syllables, e.g. Speakers 2 and 8, while some speakers produce an asymmetrical fronting pattern, e.g. Speakers 3 and 7. Speaker 4 shows less centralization of vowel F2 than other speakers, and Speaker 6 fronts all vowels in later syllables. To some degree, all speakers elided high vowels, but some, e.g. Speaker 9, did so very regularly, resulting in the sparse plots below.<sup>11</sup>

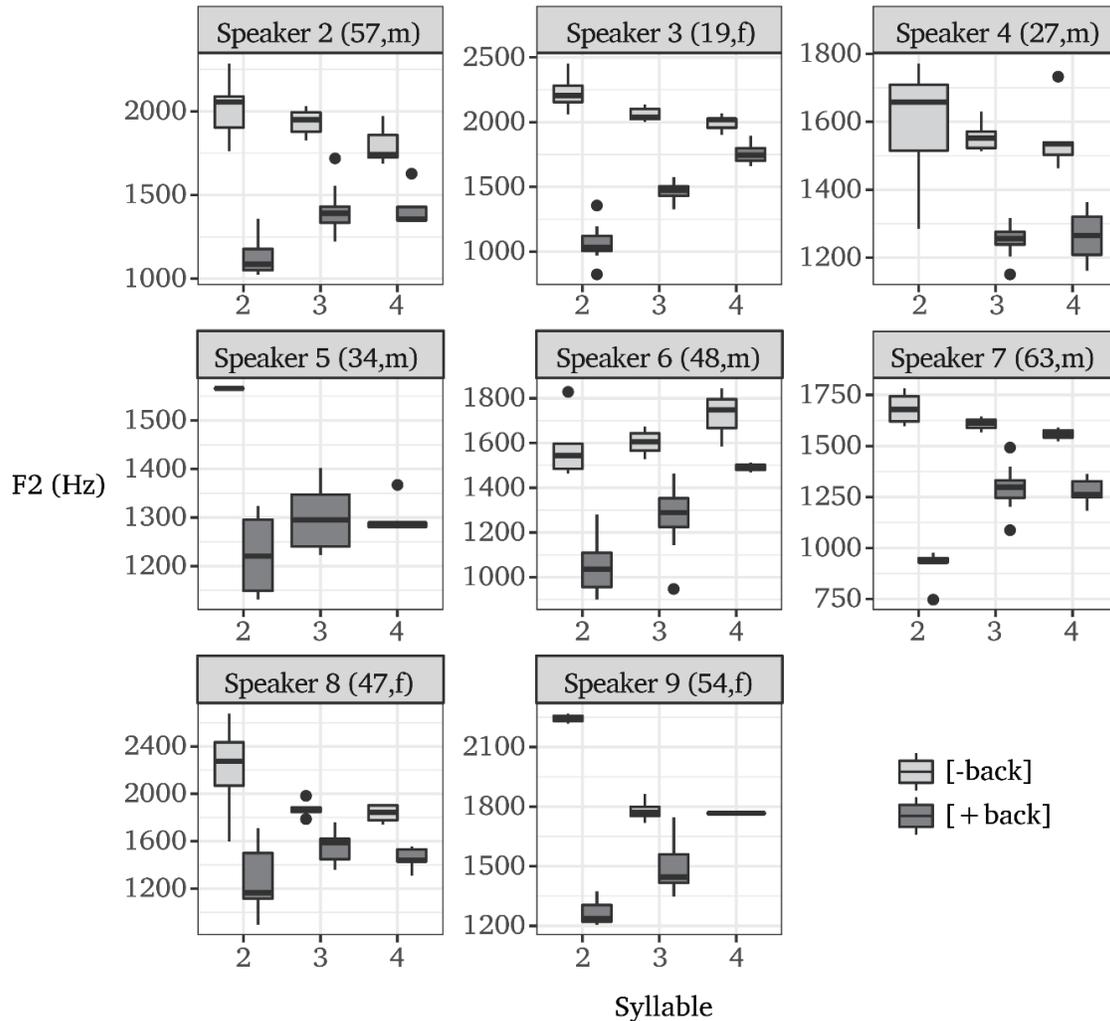


Figure 9: By-speaker plots of F2 (Hz) based on root backness in the long-distance condition. Speaker 1 did not participate in this portion of the experiment.

<sup>11</sup> Speaker 5 did not produce many [-back] vowels in this condition because he preferred other, [+back] lexemes to the [-back] lexemes targeted for elicitation. For instance, when trying to elicit /tɔpæ/ ‘hill’ or ‘summit’, he preferred /tʃoqqɑ/ ‘summit’ instead. For /sællæ/ ‘turban’, he preferred /tʃalma/, a Turkic synonym that has been adopted into Russian.

To determine whether or not low vowels alternate for harmony in the long-distance condition, fifth-syllable low vowels in the locative suffix ([-dɑ]~[-dæ]) were examined. Root vowel backness was used to predict low vowel F2. As above, the model also included a random intercept for speaker.

Once again, F2 of the low vowels was significantly lowered after [+back] roots [ $\beta = -1.14, \chi^2(1) = 117.70, p < .001$ ]. In other words, low vowels alternate for backness harmony in the long-distance condition, shown in Figure 10.

Table 11: Regression model output for low vowels in the long-distance condition (n = 95)

	Estimate	SE	$\chi^2$	<i>p</i>
Intercept	1.14	0.10	26.20	< .001
Backness	-1.13	0.07	117.70	< .001

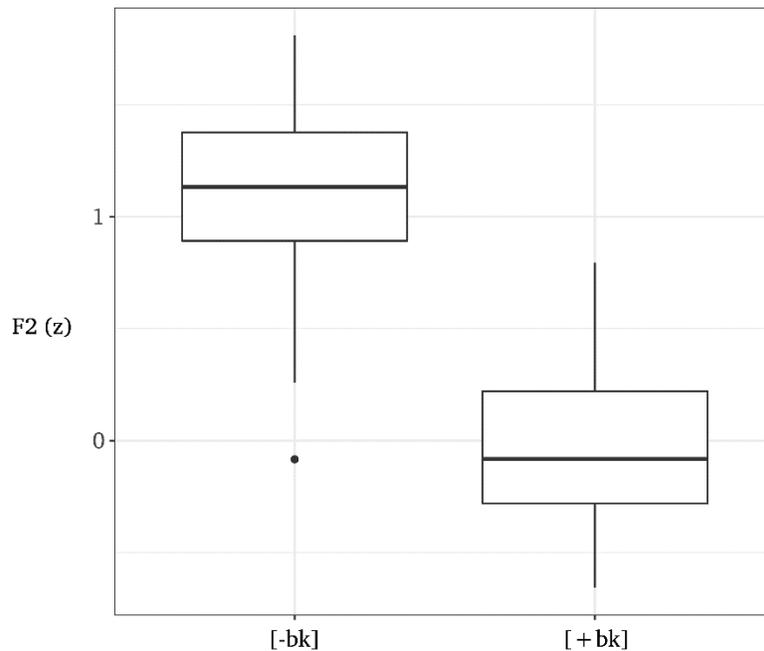


Figure 10: F2 (z) of suffix low vowels by root backness in the long-distance condition

In summary, this subsection has shown that in the long-distance condition, F2 of reportedly transparent vowels systematically varies by initial-vowel backness, suggesting that these vowels alternate for harmony. This subsection has also demonstrated that F2 of these vowels depends on position, with third- and fourth-syllable [u] exhibiting higher F2 than in

earlier syllables. Additionally, F2 of low vowels following high vowels also varies in accordance with initial-vowel backness. These results mirror results from the short-distance condition, confirming that high vowels do in fact alternate for harmony in non-initial syllables. Moreover, high vowels in epenthetic, underlying, and raising contexts all alternate for harmony. These results indicate quite strongly that non-initial high vowels exhibits backness alternations.

While it is clear that high vowels systematically vary by backness context, there are multiple possible explanations for this variation. These alternations may be phonological and contrast-neutralizing, or they may be allophonic in nature. Alternatively, these effects may be phonetic in nature, due to coarticulatory influences from flanking [ $\pm$  back] vowels. The next section discusses these possibilities and their implications.

## 6 Discussion

The findings from the previous section provide clarity, but also pose intriguing problems for the analysis of harmony in Uyghur. The behavior of the high unrounded vowels is relatively clear in target positions – they exhibit surface alternations. However, a reviewer asks if these surface alternations should be considered phonological, or if they are better analyzed as phonetic, due to coarticulation with flanking [ $\pm$  back] vowels. There are several reasons to believe they are phonological in nature. First, the magnitude of these alternations is comparable to the uncontested alternations for /a-æ/ and /u-y/, shown in Figure 3 and Figure 8. The vowels output by harmony are genuinely front and back vowels. Second, their behavior across syllables mirrors that of phonologically-alternating vowels, as seen in Figure 8. The asymmetric fronting pattern described for [i]-[u] parallels the patterns for other, alternating vowels in Uyghur. Third, these vowels trigger structure-preserving phonological alternations. The consonant of the gerundial and dative suffixes alternates between [g] and [ɣ] (1e-h; also [bal-ɣɑ] ‘honey-DAT’ and [bæl-gæ] ‘waist-DAT’). When preceded by high vowels, the dative suffix shows the same behavior, [bal-lur-u-ɣɑ] ‘honey-PL-POSS.3S-DAT’ and [bæl-lir-i-gæ] ‘waist-PL-POSS.3S-DAT.’ If these are phonetic effects, then harmony must operate over long distances, yielding forms like [bal-lir-i-ɣɑ]. Phonetics is then responsible for the alternation between [i] and [u], filling in unassimilated /i/ with its actual pronounced backness. Rather than requiring phonology and phonetics to do the same basic work, it is more parsimonious in my view to treat all non-initial vowel alternations as phonological.

Also, the alternations attested in Section 5.2 have significant implications for existing analyses. For one, the transparency-based analyses developed in Lindblad (1990) requires revision. Specifically, Lindblad’s appeal to a late neutralization rule incorrectly predicts the absence of [u], although this vowel regularly occurs after [+ back] vowels. One must either

supplement the derivational analysis with later, allophonic rules dictating the realization of [i] and [u], or discard the neutralization rule mapping |u| to [i]. As for Vaux (2000), as well as related analyses in Halle et al. (2000) and Nevins (2010), these analyses rely on the claim that [i] and [u] are non-contrastive, with harmony operating between contrastively specified vowels only. If the application of harmony is truly defined by the contrastive status of these vowels, this in turn predicts that the different speakers examined here should demonstrate qualitatively different patterns in Section 5.2. More concretely, if a speaker does not display a contrast between the two high unrounded vowel in Section 5.1, then no phonetic differences should emerge in the non-initial positions tested in Section 5.2. Conversely, if a speaker does exhibit a contrast among these vowels in root-internal positions, then these vowels should alternate for harmony. This is clearly not consistent with the data; [i] and [u] regularly alternate for harmony for all speakers regardless of their speaker-specific contrastive status.

Despite this challenge, it is possible in a contrast-based analysis to generate the attested pattern. In Vaux's analysis, this would require an additional harmony rule operating post-lexically on /i/ to derive surface [i] and [u], just like the additional late rule necessary for Lindblad's analysis. Problematically, this results in three different backness harmony rules for Vaux; a cyclic rule, a post-cyclic rule, and a third, post-lexical rule. Similar issues arise with Nevins' (2010) analysis. Under Nevins' search and copy analysis of Uyghur, alternating vowels search for the harmonic feature in nearest contrastively specified vowel. Given the present results, though, this again requires an additional search to derive the backness of [i] and [u]. This sort of post-lexical search, where all vowels (or perhaps just /i/) search for the harmonic feature largely duplicates the behavior of the lexical search and copy procedure. The other option for such an analysis is to abandon contrast as the delimiting factor in non-initial alternations, and allow all alternating vowels to search for the nearest value of [back], regardless of the trigger's contrastive status.

The real issue is how to analyze [i] and [u] within roots. I cannot offer any broad-reaching generalities because the behavior and status of these surface vowels appears to be speaker-specific. For some speakers, there is no evidence for a contrastive relationship between these two vowels, i.e. Speakers 8 and 9. For Speaker 4, however, there is evidence for a contrast between /ilim/ and /ʃilim/. For the other speakers, the data are less clear, with either weak evidence for contrast (/ilim/ and /ʃilim/ for Speaker 6), or a lexically-specific, idiosyncratic relationship between root-internal [i]-[u] and suffix backness, i.e. Speakers 1-3 and 5. It is worth noting, although briefly, that participants who produced the most evidence for contrastive /i/-/u/ were all males (Speakers 4, 6, and 7). There is no immediately obvious evidence for any age- or socioeconomic-based generalizations, though. Moreover, since all speakers grew up and reside around Chunja, I doubt these differences derive from dialectical

differences. At this point, I cannot conjecture whether or along what lines these differences persist among Uyghur speakers. Future work is definitely necessary.

Comparing the distribution of [i] and [u] within roots to the distribution of these vowels in suffixes suggests an additional challenge for analysis. Within roots, the distribution of these vowels appears largely constrained by adjacent consonants. However, adjacent consonants do not appear to play the same role in suffixes. For instance, [i] is almost entirely absent before root-internal laterals; instead [u] is more common. Before the initial lateral of the plural suffix, though, [i] occurs after [-back] vowels, as in [sælli-lær] ‘turban-PL’. In Optimality Theory (Prince & Smolensky 2004), if some contextual markedness constraint limits the occurrence of [il], its effect must be limited to roots. Consonantal effects seem more significant within initial syllables than across morpheme boundaries. This runs counter to a range of proposals, which claim that a given distribution of sounds should be least constrained in prominent positions, e.g. roots, initial syllables (Trubetskoy, 1969; McCarthy & Prince, 1995:116-117; Steriade, 1995; Beckman, 1997; cf. Noske, 2000; Urbanczyk, 2007).

More generally, the distribution of root-internal high unrounded vowels forces questions like, what is contrast? Are there degrees of contrast? Kiparsky (2015) argues that structural contrast should be divorced from perceptual distinctiveness, producing a four-way typology, shown in Table 12. If contrast and distinctiveness are not yoked, then two sounds may be distinctive in very perceivable ways without exhibiting actual linguistic contrast. Kiparsky (2015) points to Russian [i] and [ɨ] as examples of such a salient but non-contrastive relationship. On the other hand, some sounds may exhibit structurally different phonological behavior while exhibiting the same acoustic and/or articulatory properties. Kiparsky calls this near-merger, but to distinguish this category from the typical use of that term (e.g. Labov et al., 1991; Yu, 2007), I refer to this as abstract contrast. The vowels [ɛ] and [ɔ] in Tutrugbu illustrate this sort of relationship. The front vowel [ɛ] is the surface output of both /ɪ/ and /ɛ/, while [ɔ] is the surface output of /ʊ/ and /ɔ/ (McCollum & Essegbey, 2020; McCollum et al., 2020). Despite their surface neutralization, these sounds in Tutrugbu still maintain their abstract structural contrast for [high] for both rounding and ATR harmony.

Table 12: A typology of contrast and distinctiveness (Kiparsky, 2015)

	Contrastive	Non-contrastive
Distinctive	phoneme (English /t/ and /d/)	quasi-phoneme (Russian [i] and [i])
Non-distinctive	abstract contrast (Tutrugbu /ɛ/-/ɪ/ and /ɔ/-/ʊ/)	allophone (English [t] and [t <sup>h</sup> ])

Framed within Kiparsky’s typology, the relationship between [i] and [ɯ] is likely allophonic for most speakers consulted, since the high unrounded vowels don’t exhibit terribly salient acoustic differences (supported by the inability of both foreign and native-speaker linguists to note the non-initial alternations discussed in Section 5.2). That being said, for other speakers, the relationship between the high unrounded vowels aligns more closely with other of Kiparsky’s categories. For Speaker 4, these two vowels are likely contrastive. For Speaker 7, who selects [-back] suffixes after /ilim/ but [+back] suffixes after /jilim/, the relationship may be more akin to the abstract contrast in Tutrugbu, since there are no meaningful acoustic differences between these two roots, despite their bifurcated phonological behavior.

Another possibility is that the relationship between [i] and [ɯ] simply isn’t categorical, but rather contrast and allophony exist on a continuum. This is the very point argued in Hall (2009, 2013; see also Goldsmith, 1995). Hall’s proposal offers a way to compare degrees of contrast between these two vowels, but also for the comparison of contrast across languages. Hall’s model uses entropy and probability to define a continuum of phonological relationships intermediate between complete contrast and complementarity. Returning to the issue of [i]-[ɯ] and the lateral, [ɰ] is much more common before /l/ in the initial syllable. However, in non-initial syllables this consonantal conditioning appears ineffectual, since both [il] and [ɰl] occur widely in non-initial syllables, e.g. [paltɯ-lar] ‘axe-PL’ and [sælli-lær] ‘turban-PL.’ If consonantal context drives wholesale allophony between these two sounds, we might predict the unattested \*[sællɰ-lær]. It is possible that [i] and [ɰ] are conditioned by lexical contrast, adjacent consonants, and preceding vowel backness. In initial syllables, preceding vowel backness is absent, leaving lexical contrast and adjacent consonants to condition variation between [i] and [ɰ], at least for some speakers, due to acoustic and phonological differences noted above. In suffixes, though preceding vowel backness is relevant, overcoming potential consonantal effects on vowel quality as well as potential lexical contrasts between /i/ and /ɰ/ in non-initial syllables.

One final possibility is that both Uyghur [i]-[u] and other outliers, like Russian [i] and [i̯], derive their behavior from other factors. For instance, Russian [i] and [i̯] are represented orthographically, and figure meaningfully in Russian pedagogy. In contrast, Uyghur [i] and [u] are not represented orthographically. Uyghur, in both contemporary Xinjiang and Kazakhstan, is represented by scripts that do not convey a distinction between the two sounds. If non-linguistic factors may play a role in the psychological reality of a sound pair, then perhaps orthography offers some explanatory power. Additionally, it is also worth asking what role a second language plays in the maintenance or loss of contrast. Kazakhstani Uyghurs speak both Uyghur and Kazakh, and the genetic and structural similarities of the two languages may influence one another (Kaydarov, 1970:25). As noted above, the backness distinction between the high unrounded vowels in Kazakh is quite robust, and given the lexical similarities between the two, it is plausible that a contrast in Kazakh may help to maintain a contrast in Kazakhstani Uyghur. Such a proposal does, though, presuppose cultural and linguistic affinity that may not actually exist. It is well known that some speech communities enhance their linguistic distinctives in order to separate themselves from a related group, so it is not at all clear that structural similarities would play a contrast-preserving role in this case.<sup>12</sup>

Moreover, root-internal [i] and [u] raise the question, how does one investigate structural contrast from a phonetic point of view? In the clearest of cases, the answer seems simple. One examines the distribution of sounds along some continuum to determine whether their distributions are similar or sufficiently distinct. Pairing such a production-based approach with perception testing should, in many cases, offer a relatively clear picture of the relationship between two sounds. However, other factors, particular lexical factors, are known to play a significant role in production and perception. Issues like neighborhood density (Scarborough & Zellou, 2013), contextual predictability (Seyfarth, 2014), morphological constituency (Plag et al., 2017), and intra-paradigmatic relationships (Seyfarth et al., 2018) all affect the realization of segments, and if the contrast is relatively subtle, then it could become quite difficult to tease apart the different possible effects inherent within a set of data points. Does contrast really boil

---

<sup>12</sup> One other possibility is that the researcher's personal history among Kazakhs and greater proficiency in the Kazakh language resulted in an artefactual pattern of more Kazakh-like speech among the participants in this study. Although I cannot rule this out, there are several reasons to doubt this. First, I did not speak Kazakh during data collection, but rather Russian and Uyghur (not well, mind you) to avoid this very scenario. Second, if the speakers were accommodating me by speaking more Kazakh-like, I would expect them to produce Kazakh lexical items rather than imitating lower-level Kazakh phonetic patterns (e.g. medial [i]-[u] alternations, by-syllable back vowel fronting). To my knowledge, no speaker ever produced a Kazakh lexeme in place of a Uyghur lexeme. Rather, if they could not retrieve a Uyghur word or were simply code-switching during conversation, they used Russian.

down to minimal pairs and semantic differences? If so, then for most speakers consulted Uyghur [i] and [ɯ] are not contrastive. However, for at least one speaker the relationship between these two vowels appears contrastive, and for others, the relationship is murky. I must rely on future work to refine our understanding of contrast and how to evaluate it from an empirical point of view.

## 7 Conclusion

This paper has demonstrated that the high unrounded vowels do, in fact, alternate for backness harmony in Uyghur, suggesting that backness harmony operates locally, affecting all non-initial vowels. This suggests that previous analyses, especially those dependent on contrast to delimit the operation of harmony in the language, can be simplified. Furthermore, the finding that harmony is local further exemplifies the usefulness of experimental methods to more accurately document empirical patterns, providing a strong basis for formal analysis (Pierrehumbert et al., 2000). While the surface alternation between [i] and [ɯ] in non-initial positions is clearly supported, the relationship between these vowels in roots is less clear. For most speakers consulted, the relationship is likely allophonic, while for others these two vowels may exhibit a structural contrast. Generally, the paper questions the nature of contrast, how it is manifested phonetically, and the potential for degrees of contrast intermediate between categorical contrast and allophony (Hall 2009, 2013).

### Acknowledgments:

I would like to thank Alan Yu and two anonymous reviewers for their feedback, which has significantly improved the paper. Also, discussions with Sharon Rose, Eric Baković, Sarah Creel, Marc Garellek, Rachel Walker, and the audience at the 26<sup>th</sup> Manchester Phonology Meeting have been incredibly helpful. Any errors are my own.

### Funding:

This work was supported by a University of California President's Dissertation Year Fellowship.

## References

- Abdurehim, E. (2014). *The Lopnor dialect of Uyghur: A descriptive analysis*. Publications for the Institute of Asian and African Studies, 17. University of Helsinki.
- Archangeli, D., & Pulleyblank, D. (1994). *Grounded phonology*. MIT Press.
- Baskakov, N. A. (1970). Nekotorye zadachi izucheniya sovremennogo ujugurskogo yazyka [Some tasks of studying the modern Uyghur language]. In A. T. Kaydarov, G. Sadvakasov, & T. Talipov (Eds.), *Issledovaniya po ujugurskomu jazyku [Research on the Uyghur language]* (Vol. 2, pp. 7–16). Alma-Ata: Nauka.
- Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.  
<http://dx.doi.org/10.18637/jss.v067.i01>
- Beckman, J. N. (1997). Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology*, 14(1), 1–46. <https://dx.doi.org/10.1017/S0952675797003308>
- Benus, S., & Gafos, A. I. (2007). Articulatory characteristics of Hungarian “transparent” vowels. *Journal of Phonetics*, 35(3), 271–300. <https://dx.doi.org/10.1016/j.wocn.2006.11.002>
- Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer (Version 5.4.18).  
<http://www.fon.hum.uva.nl/praat/>
- Calabrese, A. (1995). A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry*, 373–463.
- Dye, A. (2015). Vowel Harmony and Coarticulation in Wolof and Pulaar: An Ultrasound Study (PhD dissertation). New York University.
- Gafos, A. I. (1999). *The articulatory basis of locality in phonology*. Garland Publishing.
- Gafos, A. I., & Dye, A. (2011). Vowel harmony: Opaque and transparent vowels. *The Blackwell Companion to Phonology*, 1–26. <https://dx.doi.org/10.1002/9781444335262>
- Gick, B., Pulleyblank, D., Campbell, F., & Mutaka, N. (2006). Low vowels and transparency in Kinande vowel harmony. *Phonology*, 23(1), 1–20.  
<https://dx.doi.org/10.1017/S0952675706000741>
- Goldhahn, D., Eckart, T., & Quasthoff, U. (2012). Building Large Monolingual Dictionaries at the Leipzig Corpora Collection: From 100 to 200 Languages. *LREC*, 29, 31–43.  
<http://wortschatz.uni-leipzig.de/en>
- Goldsmith, J. A. (1976). Autosegmental phonology (PhD dissertation). MIT.
- Goldsmith, J. A. (1995). Phonological theory. *The handbook of phonological theory* (pp. 1–23). Wiley-Blackwell.
- Hahn, R. (1986). Modern Uighur language research in China: Four recent contributions examined. *Central Asiatic Journal*, 30(1/2), 35–54.

- Hahn, R. (1991). *Spoken Uyghur*. University of Washington Press.
- Hall, K. C. (2009). A probabilistic model of phonological relationships from contrast to allophony (PhD dissertation). The Ohio State University.  
<https://blogs.ubc.ca/kathleencurriehall/dissertation/>
- Hall, K. C. (2013). A typology of intermediate phonological relationships. *The Linguistic Review*, 30(2), 215–275. <https://dx.doi.org/10.1515/tlr-2013-0008>
- Halle, M., Vaux, B., & Wolfe, A. (2000). On feature spreading and the representation of place of articulation. *Linguistic Inquiry*, 31(3), 387–444.
- Hayes, B., & Londe, Z. C. (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology*, 23(1), 59–104.  
<https://dx.doi.org/10.1017/S0952675706000765>
- Hayes, B., & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3), 379–440.  
<https://dx.doi.org/10.1162/ling.2008.39.3.379>
- Heinz, J. (2018). The computational nature of phonological generalizations. Larry M. Hyman and Frans Plank (Eds.), *Phonological Typology*, 126–195. De Gruyter.  
[http://jeffreyheinz.net/papers/heinz\\_papers.html](http://jeffreyheinz.net/papers/heinz_papers.html)
- Heinz, J., & Lai, R. (2013). Vowel harmony and subsequentity. *Proceedings of the 13th Meeting on the Mathematics of Language (MoL 13)*, 52–63.  
<https://www.aclweb.org/anthology/W13-3006/>
- Kajdarov, A. T. (1970). Ujgurovedenie v Kazaxstane [Uyghur studies in Kazakhstan]. In A. T. Kaydarov, G. Sadvakasov, & T. Talipov (Eds.), *Issledovaniya po ujugurskomu jazyku [Research on the Uyghur language]* (Vol. 2, pp. 21–30). Alma-Ata: Nauka.
- Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology Yearbook*, 2(1), 85–138. <https://dx.doi.org/10.1017/S0952675700000397>
- Kiparsky, P. (2015). Phonologization. Patrick Honeybone and Joseph Salmons (Eds.), *The Oxford handbook of historical phonology*. Oxford University Press.  
<https://dx.doi.org/10.1093/oxfordhb/9780199232819.001.0001>
- Labov, W., Karen, M., & Miller, C. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change*, 3(1), 33–74.  
<https://dx.doi.org/10.1017/S0954394500000442>
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R Package Version*, 1(1).
- Lindblad, V. M. (1990). Neutralization in Uyghur (MA thesis). University of Washington.

- Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, 49(4B), 606–608.  
<https://dx.doi.org/10.1121/1.1912396>
- Mayer, C., & Major, T. (2018). A challenge for tier-based strict locality from Uyghur backness harmony. *Formal Grammar 2018*, 10950, 62–83. Springer.  
[https://dx.doi.org/10.1007/978-3-662-57784-4\\_4](https://dx.doi.org/10.1007/978-3-662-57784-4_4)
- Mayer, C., Major, T., & Yakup, M. (2019). Wug-testing Uyghur vowel harmony. Presented at the 27th Manchester Phonology Meeting. slides available at:  
<https://linguistics.ucla.edu/people/grads/connormayer/research.html>
- McCarthy, J. J., & Prince, A. S. (1995). Faithfulness and reduplicative identity.  
<https://dx.doi.org/doi:10.7282/T31R6NJ9>
- McCollum, A. G. (2018). Vowel dispersion and Kazakh labial harmony. *Phonology* 35(2), 287–326. <https://dx.doi.org/10.1017/S0952675718000052>
- McCollum, A. G. (2019a). Gradience and locality in phonology: Case studies from Turkic vowel harmony (PhD thesis). University of California, San Diego.  
<https://escholarship.org/uc/item/7sx31303>
- McCollum, A. G. (2019b). Gradient morphophonology: Evidence from Uyghur vowel harmony. *Proceedings of the Annual Meeting on Phonology*.  
<https://dx.doi.org/10.3765/amp.v7i0.4565>
- McCollum, A. G., Baković, E., Mai, A., & Meinhardt, E. (2020). Unbounded circumambient patterns in segmental phonology. *Phonology* 37(2), 215–255.  
<https://dx.doi.org/10.1017/S095267572000010X>
- McCollum, A. G., & Essegbey, J. (2020). Initial prominence and progressive vowel harmony in Tutrugbu. *Phonological Data & Analysis* 2(3), 1–37.  
<https://dx.doi.org/10.3765/pda.v2art3.14>
- Menges, K. H. (1995). *The Turkic languages and peoples: An introduction to Turkic studies* (Vol. 42). Otto Harrassowitz Verlag.
- Mohanan, K. P. (1982). *Lexical phonology* (PhD Thesis). Massachusetts Institute of Technology.
- Nadzhip, E.N. (1971). *Modern Uigur*. (D. Segal, trans.). Moscow: Nauka.
- Nevins, A. (2010). *Locality in vowel harmony*. MIT Press.
- Noske, M. (2000). [ATR] harmony in Turkana: A case of FAITH SUFFIX > > FAITH ROOT. *Natural Language & Linguistic Theory*, 18(4), 771–812.  
<https://dx.doi.org/10.1023/A:1006474124675>
- Odden, D. (1994). Adjacency parameters in phonology. *Language*, 289–330.

- Plag, I., Homann, J., & Kunter, G. (2017). Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics*, 53(1), 181–216.  
<https://dx.doi.org/10.1017/S0022226715000183>
- Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint interaction in generative grammar*. Wiley-Blackwell. <https://dx.doi.org/10.1002/9780470756171>
- R Core Team. (2017). R: A language and environment for statistical computing (Version 3.4.2). R foundation for statistical computing. <https://www.r-project.org/>
- Ringen, C. O., & Heinämäki, O. (1999). Variation in Finnish vowel harmony: An OT account. *Natural Language & Linguistic Theory*, 17(2), 303–337.
- Ringen, C. O., & Kontra, M. (1989). Hungarian neutral vowels. *Lingua*, 78(2–3), 181–191.
- Ritchart, A., & Rose, S. (2017). Moro vowel harmony: Implications for transparency and representations. *Phonology*, 34(1), 163–200.  
<https://dx.doi.org/10.1017/S0952675717000069>
- Scarborough, R., & Zellou, G. (2013). Clarity in communication: “Clear” speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America*, 134(5), 3793–3807.  
<https://dx.doi.org/10.1121/1.4824120>
- Seyfarth, S. (2014). Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition*, 133(1), 140–155.  
<https://dx.doi.org/10.1016/j.cognition.2014.06.013>
- Seyfarth, S., Garellek, M., Gillingham, G., Ackerman, F., & Malouf, R. (2018). Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*, 33(1), 32–49. <https://dx.doi.org/10.1080/23273798.2017.1359634>
- Smajylov, A. A. (ed.). (2011). Qazaqstan Respublikasyndaghy ulttyq quram, dini nanyim zhane tilderdi menggeru [Ethnic composition, religious beliefs, and language proficiency in the Republic of Kazakhstan]. Qazaqstan Respublikasy Statistika Agenttigi [Statistical Agency of the Republic of Kazakhstan]. <https://stat.gov.kz/census/national/2009>
- Steriade, D. (1995). Underspecification and markedness. *The handbook of phonological theory*, J. A. Goldsmith (Ed.), pp. 114–174.. Cambridge, MA: Blackwell.
- Szeredi, D. (2016). Exceptionality in vowel harmony (PhD dissertation). New York University. <https://ling.auf.net/lingbuzz/003148>
- Trubetskoy, N. S. (1969). *Principles of Phonology* (C. Baltaxe, trans.). University of California Press.(Original work published 1939).
- Urbanczyk, S. (2006). Reduplicative form and the root-affix asymmetry. *Natural Language & Linguistic Theory*, 24(1), 179–240. <https://dx.doi.org/10.1007/s11049-005-4373-x>

- Vago, R. M. (1973). Abstract vowel harmony systems in Uralic and Altaic languages. *Language* 49(3), 579–605.
- Vago, R. M. (1976). Theoretical implications of Hungarian vowel harmony. *Linguistic Inquiry*, 7(2), 243–263.
- van der Hulst, H., & van der Weijer, J. (1995). Vowel harmony. J. A. Goldsmith (Ed.), *The handbook of phonological theory*, pp. 495–534. Cambridge, MA: Blackwell.
- Vaux, B. (2000). Disharmony and derived transparency in Uyghur vowel harmony. *Proceedings of the North East Linguistic Society*, 30, 672–698.
- Walker, R. (1999). Guaraní voiceless stops in oral versus nasal contexts: An acoustical study. *Journal of the International Phonetic Association*, 29(1), 63–94.  
<https://dx.doi.org/10.1017/S0025100300006423>
- Walker, R. (2012). Vowel harmony in optimality theory. *Language and Linguistics Compass*, 6(9), 575–592. <https://dx.doi.org/10.1002/lnc3.340>
- Walker, R., Byrd, D., & Mpiranya, F. (2008). An articulatory view of Kinyarwanda coronal harmony. *Phonology*, 25(3), 499–535. <https://dx.doi.org/10.1017/S0952675708001619>
- Yakup, A. (2005). *The Turfan dialect of Uyghur*. Harrassowitz Verlag.
- Yu, A. C. L. (2007). Understanding near mergers: The case of morphological tone in Cantonese. *Phonology*, 24(1), 187–214. <https://dx.doi.org/10.1017/S0952675707001157>