

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 of non-local dependencies. However, experimental work has demonstrated on a number of occasions that ‘transparent’ vowels actually undergo harmony (e.g. Gick, Pulleyblank, Campbell, & Mutaka, 2006), re-establishing the role of locality in the analysis of harmony. Transparency has also been shown to be constrained by count effects – a single token of a vowel may be transparent but multiple tokens are not (Hayes & Londe, 2006; Ringen & Kontra, 1989). However, existing work on backness harmony in Uyghur argues that harmony truly skips the high front vowel /i/ and is unaffected by multiple tokens of transparent /i/. 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, as [i] and [u] regularly alternate for harmony. This finding is interpreted as further support for the role of locality in harmony more generally.

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 define 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 (also called transparent) interactions in harmony, 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 may apply to contrastively specified segments, non-contrastive differences cannot play a role in harmony (see also Vaux 2000). 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 early in the phonological derivation. In contrast, allophonic patterns are postlexical, and as a consequence cannot affect the application of harmony. In other analyses, historical contrasts have 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 cannot enforce locality in any phonetically-grounded sense.

This paper addresses two questions as they relate to Uyghur backness harmony, incorporating experimental findings toward the goal of formal analysis (Pierrehumbert, Beckman, & Ladd, 2000). First, how non-local may harmony be? Extant descriptions claim that harmony may skip multiple /i/ vowels, but no phonetic studies have been conducted on the language. Second, what role does contrast play in the operation of Uyghur harmony? Transparency in previous analyses is predicated on the claim that harmony does not output [ɯ] although consonantal context may. This paper thus examines the distribution of high unrounded vowels to assess whether two vowels, [i] and [ɯ] are distinguishable based on consonantal and vocalic contexts within roots and suffixes.

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 on the language. 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, /a 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 Yakup (2005:31), [e] is typically a raised variant of /æ/ or /a/, 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	a	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]~[ɣ].

- (1) Backness harmony in Uyghur
- a. kəl-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-ɣu ‘remain-GER’
 - h. bar-ɣu ‘go-GER’

In addition to the low vowels and the high round vowels, two other sets of vowels can be defined for harmony. First, the mid vowels, /e ø o/ do not typically occur in non-initial syllables in the standard 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 further consider these vowels. Second, previous work reports that the high front unrounded vowel, /i/, is transparent to harmony (Mayer et al. 2019). Root /i/ triggers [+back] vowel suffixes in (2a,b), but front vowel suffixes in (2c,d). Moreover, in polysyllabic roots /i/ is transparent. In disyllabic roots with second-syllable /i/, the backness of the initial-syllable vowel controls the realization of the suffix (2e,f).¹

- (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-ka | ‘sunflower-DAT’ |

In (2), variation occurs across different lexical roots, but the data in (3) demonstrates that a single root may variably trigger front or 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 by particular suffixes, 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, Eckart, & Quasthoff, 2012).

¹ 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, Eckart, & Quasthoff, 2012)
- | | | |
|----|-------------|---------------------|
| a. | ij-lar | ‘work-PL’ |
| b. | tʃi-i-ka | ‘tooth-POSS.3S-DAT’ |
| c. | ij-læ | ‘work-VRB’ |
| d. | tʃi-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 an alternating locative suffix.²

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

In addition to underlying high vowels, high vowels may arise via vowel raising or epenthesis in the language. Epenthetic vowels are inserted to break up illicit coda consonant 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. When the epenthetic vowel is rounded, it undergoes backness harmony, alternating between [y] and [u]. However, when unrounded, the epenthetic vowel surfaces as [i] instead of [u] after back vowels (5f).³ When [i] surfaces, the backness of the preceding vowel still controls the realization of subsequent suffixes, like the locative suffix below.

² 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.

³ 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.	jøl-um-da		‘road-POSS.1S-LOC’
e.	æsæl-im-dæ		‘honey-POSS.1S-LOC’
f.	aslan-im-da		‘kitten-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 smaller phonetic effect 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 +nInj + *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, generally transcribed as [i] or [ɪ].

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 absolute 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 vowels.

Problematically, in all previous work the fact that [u] may not arise via harmony is never juxtaposed to the allophonic backing of /i/ to [ə i ɤ u] in certain consonantal contexts, which I collectively treat as [u] for simplicity. 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, apparent transparency should occur in [balida] 'child-LOC' but not when consonantal effects are taken into account, [baluɫar] 'child-PL.' The same complication holds for front vowel forms, too, predicting [sællidæ] 'turban-LOC' as well as the aberrant [sælluɫær] '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, as noted by Mayer & Major (2018).

Despite the fact that previous work doesn't consider this vowel-consonant interaction, they predict a type of partitioning of the vowel space unattested in other languages. The harmony pattern previously described suggests that the set of vowels that participate in backness harmony, /æ y a u/, and the set of vowels that participate in rounding harmony, /i y u/, partially overlap. Specifically, /y u/ are subject to both harmonies, /i/ is subject to rounding harmony when occurring as an epenthetic vowel,

while /æ α/ are subject to backness harmony only. Aksenova (2018) and Aksenova & Deshmukh (2018) demonstrate that this type of intersecting harmonies produces a computational learning problem. In languages with two active harmony patterns, all attested cases involve a subset-superset relation or the two harmonies operate over non-intersecting sets (e.g. a language with a vowel and a consonant harmony pattern). In comparison, Aksenova (2018) and Aksenova & Deshmukh (2018) argue that the learner must search a much larger hypothesis space to determine the correct partitioning of vowels for both harmonies in a partially intersecting phonological space, creating a problem for the learner.

In fact, if both vowels and consonants condition the backness of a given vowel, harmony in Uyghur is more complex than tier-based strictly local patterns (e.g. consonant harmony), contradicting Heinz's (2018) claim that phonology is maximally tier-based strictly local. 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 the phonetic level. Unlike the small, definitively phonetic effects reported in 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. Whereas the phonological status of the transparent vowels in Gordon and Benus & Gafos’ studies are likely unaffected by harmony, both the phonetic and phonological properties of Kinande /a/ are

affected by harmony. 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 /ə̃/, 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).⁴

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 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 transparency and exceptionality in harmony.

3.2 Distance effects

In addition to pseudo-transparency, phonological evidence supports an additional, distance-based restriction on transparency. For instance, 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)

⁴ Similar results have been reported for consonant and vowel-consonant harmonies (e.g. Walker, 1999; Walker, Byrd, & Mpiranya, 2008).

- (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 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

As described above, previous work suggests that ‘transparent’ vowels often undergo alternations, and their putative invisibility to harmony is constrained by distance. However, the descriptions and analyses of Uyghur seem to flout both generalizations. First, previous work argues that /i/ does not even phonetically vary according to backness harmony. Second, words with multiple transparent /i/ are reported, as in (4), with no apparent effect of count on the behavior of the transparent vowel.

Given that existing work on transparency in Uyghur depends entirely on textual and impressionistic data, this paper investigates the pattern with acoustic data from a production study. Four questions are addressed. First, what is the distribution of [i] and [u] within roots? To answer this question, it is necessary to know whether both [i] and [u] even surface in the language, and if their distribution entirely conditioned by consonantal context or not. Second, is there 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, if the high unrounded vowels do, in fact, alternate for harmony, is this a small, presumably imperceptible alternation, like Hungarian /i/, or is this larger, as in Kinande? Relatedly, if the high unrounded vowels are transparent, does their number affect the realization of following vowels? The first two address the phonological backness of high unrounded vowels in triggering positions. The third and fourth, and fifth 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, /ɑ-æ, ɒ-ø, u-y/. Pictorial prompts were used to avoid an orthographic confound, since Uyghur orthographies do not represent [ʊ]. 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.	baɭ	‘honey’
c.	bæɭ	‘waist’
d.	joɭ	‘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 /ʊ/ in closely related languages that maintain contrastive /ʊ/, Kyrgyz (kr) and Kazakh (kz). Target stimuli with putative /i/ are shown in (9). Two monosyllabic stimuli were selected that are reported to trigger [+back] harmony and have [+back] cognates in Kyrgyz and Kazakh (9a,c). Two monosyllabic stimuli were selected that are reported to trigger [+back] harmony but have [-back] cognates in Kyrgyz and Kazakh, (9b,d). Finally, in (9e,f) two disyllabic target words were selected that reportedly trigger [-back] harmony and correspond to [-back] cognates in Kyrgyz and Kazakh.

(9) Stimuli with putative /i/

	Stimulus	Gloss	Status in Lindblad	Cognates
a.	qiɟ	‘winter’	[+ back]	quɟ (kr), quɪs (kz)
b.	tɟiɟ	‘tooth’	[+ back]	tiɟ (kr), tɪs (kz)
c.	jil	‘year’	[+ back]	ɟɪul (kr), ɟul (kz)
d.	pil	‘elephant’	[+ back]	pil (kr), pɪl (kz)
e.	ilim	‘science’	[-back]	ilim (kr); ɪlɪm, ɪulɪum (kz)
f.	ɟilim	‘paste’	[-back]	ɟelim (kr), ɟielim (kz)

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 /køl/ ‘lake’ and /jol/ ‘road’ are shown below in (10).

(10) Example elicited forms

		/køɫ/	/jøl/
a.	NOM	køɫ	jøl
b.	ACC	køɫ-ni	jøl-ni
c.	LOC	køɫ-dæ	jøl-da
d.	ABL	køɫ-din	jøl-din
e.	PL	køɫ-lær	jøl-lar
f.	PL-LOC	køɫ-lær-dæ	jøl-lar-da
g.	POSS.1S	køɫ-ym	jøl-um
h.	POSS.3S	køɫ-i	jøl-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 (10b,d,h). An epenthetic high vowel occurs preceding the first-person singular possessive suffix after consonant-final roots (10g). Note that Lindblad (1990), Hahn (1991), and Vaux (2000) describe this suffix as more harmonic than other possessive suffixes (4(5)). Thus, if their descriptions are correct, we should expect to see a morpheme-based difference in harmony, with POSS.1S exhibiting rounding alternations while the other suffixes do not. Third, I examined raised vowels, which surface via the reduction of low vowels in medial open syllables, shown above in (6).

The four research questions that drive the acoustic analysis in the paper are articulated in (11). 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.

(11) Research questions:

1. Is F2 of initial [i]-[u] predictable based on consonantal context (allophony)?
2. Is there a relationship between F2 of root-internal [i]-[u] and suffix backness?
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?

These research questions and their answers are linked to related patterns in harmony below in Table 2. If harmony in Uyghur parallels Moro (Ritchart & Rose 2017), then /i/ and /u/ should exhibit a contrastive distribution in initial position, and should alternate for harmony in non-initial positions. If harmony is instead more similar to Kinande (Gick et al. 2006), the distribution of [i] and [u] in initial position should depend on flanking consonants, while the distribution of these sounds in non-initial position should depend on preceding vowel backness. In contrast, if the high unrounded vowel, /i/ and its allophones [i] and [u], are truly transparent to harmony, as in Wolof (Dye 2015), any differences between these two sounds should be describable in terms of adjacent consonants in all positions, with no effect of preceding vowel backness. Finally, if the Uyghur pattern resembles Hungarian, then the realization of [i] and [u] should depend on both flanking consonants and syllable-adjacent vowels in initial positions, with low-level phonetic alternations based on vowel backness (Benus & Gafos, 2007; Hayes & Londe, 2006; Ringen & Kontra, 1989). Further, if Uyghur is like Hungarian, the realization of regularly alternating vowels should depend on how many high unrounded vowels intervene between trigger and target.

Table 2: Schema of possible results

	Contrast neutralization	Allophonic participation	True transparency	Distance-based pseudo- transparency
Is F2 of trigger (root-internal) [i]-[u] predictable based on consonantal context alone?	No	Yes	Yes	Partially
Is there a relationship between F2 of trigger (root-internal) [i]-[u] and suffix backness?	Yes	No	No	Yes, but small
Is F2 of target (non-initial) [i]-[u] predictable based on preceding vowel backness?	Yes	Yes	No	Yes, but differences are small
Is F2 of low vowels following target (non-initial) [i]-[u] predictable based on initial vowel backness?	Yes	Yes	Yes	Yes, modulo distance
Comparison language	Moro	Kinande	Wolof	Hungarian

To examine the first question, root-internal high unrounded vowels were analyzed to determine if consonants exert the only significant effects on vowel F2. Second, the relationship between root-internal F2 of [i]-[u] was compared to suffix backness, both in categorical terms ([+ back] or [-back]) as well as in phonetic terms, i.e. F2). Similarly, non-initial high vowels (underlying, epenthetic, and raised) were analyzed to determine if preceding vowel backness exerts a significant effect on F2 of target high vowels.

To examine the fourth and fifth questions, elicited words fell into one of two conditions, exhibiting either a short- or long-distance dependency (12). 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 low vowel suffixes alternate for harmony in the short-distance condition only, then transparency is distance-based, as in Hungarian.

(12) 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. tɒpi-lir-i-dæ ‘hill-PL-POSS.3S-LOC’
- g. palti-lir-i-da ‘axe-PL-POSS.3S-LOC’
- h. χurmi-lir-i-da ‘persimmon-PL-POSS.3S-LOC’

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 the relational networks in Chunja, Kazakhstan. 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, and dialectical work on the language reports that the variety spoken in Kazakhstan is part of the central dialect region, upon which the standard language is based (Hahn, 1986:36-42, 1991:5-6; Kaydarov, 1970; Yakup, 2005:8-22). All participants reported native fluency in the target language. Most participants also reported fluency in Kazakh and Russian. Speaker participation and informed consent were obtained in accordance with University of California San Diego Linguistic Fieldwork IRB protocol #141520.

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 first three formants and duration were 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. In particular, 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, Machler, Bolker, & Walker, 2015). A mixed effect linear regression was used to predict normalized F2 at vowel midpoint for both high unrounded vowels and subsequent alternating vowels.

5 Results

5.1 Root-internal high unrounded vowels

The first question to address is whether the high unrounded vowels in these roots are describable in terms of consonantal context only (allophony) or if there exist significant differences in F2 that depend on lexical root (contrast). To investigate this, normalized F2 at vowel midpoint ($n = 631$) was analyzed using a model that included fixed effects of lexical root, preceding consonant context, and following consonant context, with by-speaker random slopes for preceding and consonantal context. In order to simplify the fixed effect structure, preceding and following consonant contexts were treated as binary variables based on Hahn's (1991) description, differentiating between backing and fronting environment. Fronting environments included word-initial contexts, immediately following or preceding an alveopalatal or palatal consonant. Backing environments included immediately following a uvular, or preceding or following a lateral or bilabial.

After constructing the baseline model that included both fixed and random effects for preceding and following consonantal context, a second model was constructed without fixed effects for preceding and following context. Model comparisons indicate that the reduced model does not significantly differ from the baseline model [$\chi^2(2) = 1.29$, $p = 0.52$, $\Delta AIC = 2.7$]. Results from the reduced model are reported in Table 3 below. Note that F2 of /ilim/ 'science', /qi/ 'winter', and /tʃi/ 'tooth' differ significantly from the model intercept.

Table 3: Model output for root-internal high unrounded vowels

	Estimate	SE	df	t-value	p (Satterthwaite)
jil (Intercept)	0.02	0.11	13.6	0.15	0.89
pil	-0.13	0.09	261.2	-1.42	0.16
filim	-0.02	0.07	478.7	-0.25	0.80
ilim	0.37	0.07	516.2	5.56	< .001 ***
qif	0.91	0.15	97.1	6.28	< .001 ***
tʃif	1.44	0.16	135.6	9.10	< .001 ***

Although the model without fixed effects for consonants was not significantly different from the baseline model, there was a highly significant difference in model fit when fixed effects for root were removed from the model [$\chi^2(5)=81.11$, $p < .001$, $\Delta AIC = 28.9$]. In tandem with the results reported in Table 3, this suggests that variation for F2 is not entirely explained by immediate consonantal context.

The distribution of F2 (z) by lexeme for the six roots is compared in Figure 1. F2 values range from less than -1z, which is comparable to [u] to more than 2z, a value consistent with a very peripheral [i] vowel quality. In general, though, values tend to fall between a 0.5 and -0.5z. In addition to the global distribution of F2, there is variation between different lexemes. Observe that the highest F2 is associated with the root, /tʃif/ ‘tooth’ while the lowest F2 is associated with /jil/ ‘year’, /pil/ ‘elephant’, and /filim/ ‘paste.’

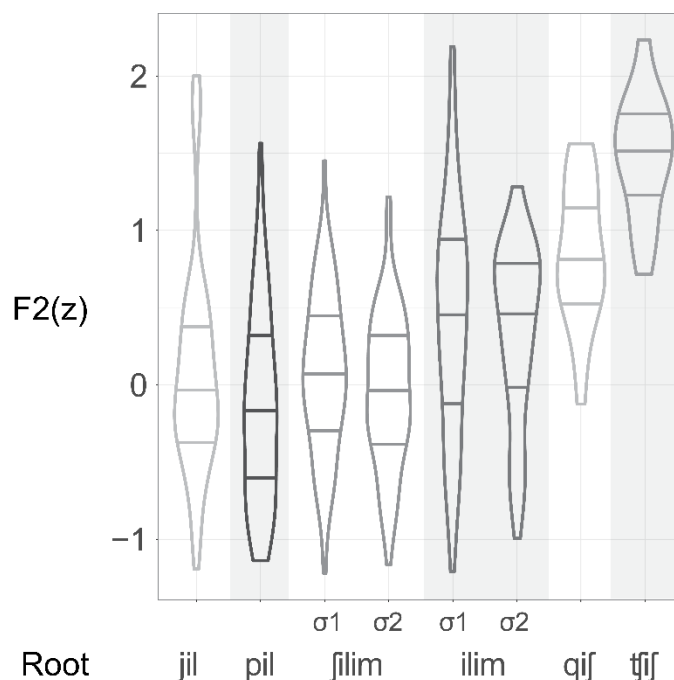


Figure 1: F2 (z) of root internal high unrounded vowels by lexeme and syllable number. Within each distribution, horizontal lines represent median and quartiles.

In particular, note the difference in F2 between the lexemes /ilim/ ‘science/ and /jilim/ ‘paste.’ A post-hoc test using the multcomp package indicates that F2 of /ilim/ is significantly higher than F2 of /jilim/ [$\beta = 0.39$, $z = 6.80$, $p < .001$] (Hothorn, Bretz, & Westfall, 2008; with Bonferroni correction). In this pair, even though only the initial consonant differs, the average difference in F2 is 0.36z, a difference that persists over both the first and second syllables within these two roots. To be clear, there is no obvious consonantal reason for this difference, as the alveopalatals appear to promote higher F2 in /qil/ ‘winter’ and /tɕiɕ/ ‘tooth.’ The lexeme /tɕiɕ/ has, by far, the highest mean F2, and is flanked by two alveopalatals. Significantly, the difference in F2 actually increases in the second syllable of these two disyllabic roots even though the consonantal context is identical. Despite being preceded by the lateral and followed by the bilabial nasal in both roots, mean F2 for of the second-syllable vowel in /ilim/ is 0.39z greater than the second-syllable vowel in /jilim/. One additional difference between /ilim/ and /jilim/ is the range of F2 values they vary over. In /ilim/, F2 values vary greatly, spanning the entire plot. In contrast, values for /jilim/ exhibit more moderate variation.

Boxplots showing by-lexeme variation in F2 (Hz) are shown for each speaker in Figure 2. Since no between-speaker comparisons are made below, F2 is presented in Hertz. In particular, observe variation in the realization of /ilim/ and /jilim/. Speakers 1 and 2 produce

relatively small differences in F2 between these two lexemes. In contrast, Speakers 4, 6, and 8 produce larger differences while Speakers 5, 7, and 9 produce no obvious differences in F2 between these lexemes.⁵ Observe across all nine speakers the relative paucity of tokens from /qij/ ‘winter’ and /tʃij/ ‘tooth.’ Initial-syllable high vowels often elide before voiceless fricatives in the language. For some speakers, the high vowel was regularly elided (e.g. Speakers 3, 4, and 9), while for others the vowel was produced with slightly more consistency (e.g. Speakers 5, 7, and 8). Speaker age and gender are noted in all by-speaker plots below.

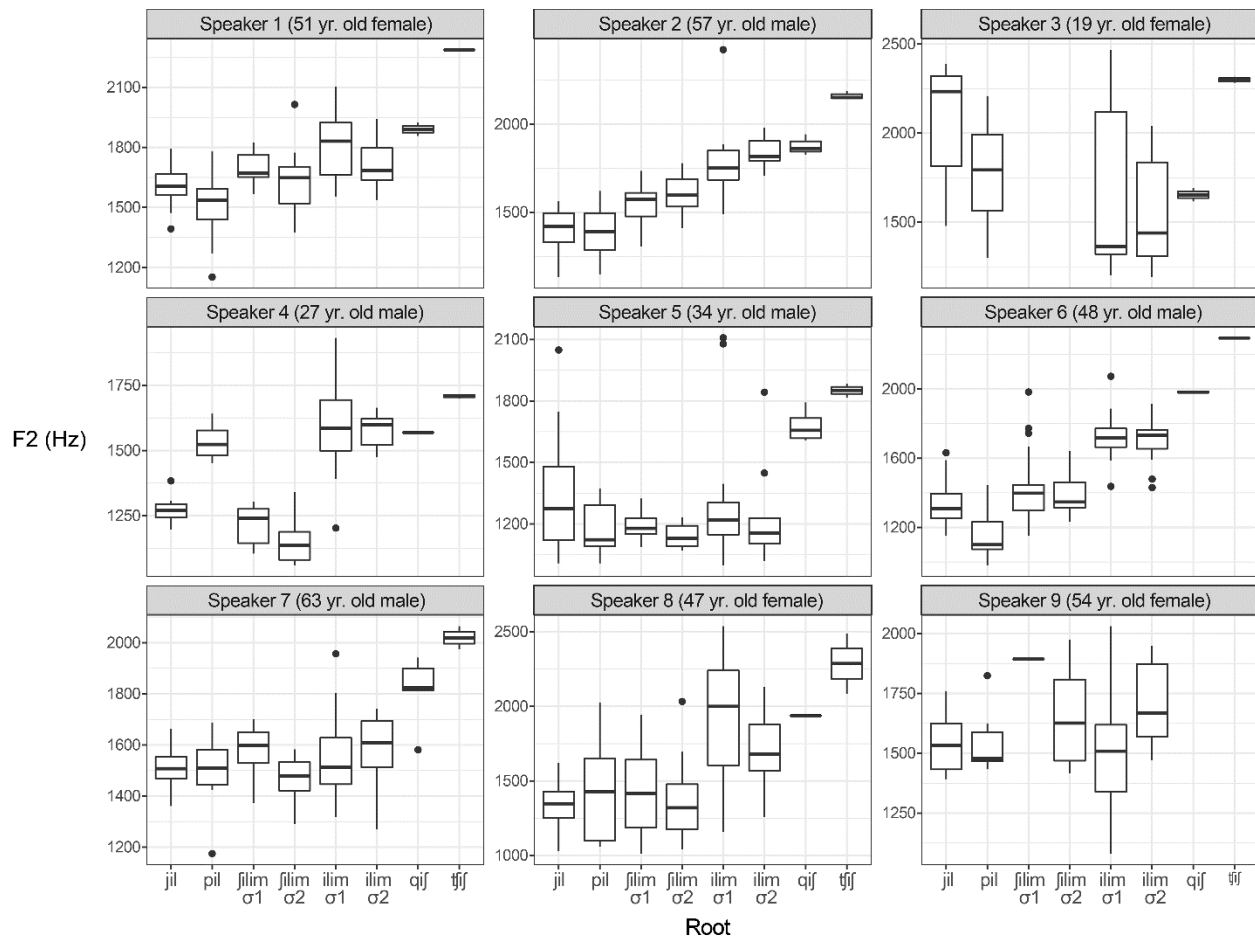


Figure 2: F2 (Hz) of root internal high unrounded vowels by lexeme and syllable number for each speaker

The results above, in particular the differences in F2 between /ilim/ and /jilim/ suggest that there may be a backness contrast among the high unrounded vowels. If a synchronic contrast between the high unrounded vowels, e.g. /i/ and /u/, exists in the language, then it

⁵ Speaker 3 did not produce /jilim/ ‘paste’, and instead produced the Russian /klʲej/.

should manifest itself in both in patterns of root-internal F2 as well as the backness of subsequent suffixes. If both /i/ and /u/ exist in the language, they should trigger [-back] and [+back] suffixes, respectively. In

Figure 3, the percentage of [+back] vowel suffixes produced during elicitation is compared with root-internal F2. It is striking that regardless of root F2, these vowels typically select [+back] suffixes, which is consistent with previous descriptions (e.g. Lindblad 1990; Hahn 1991). There is some variation but no root ever selects [-back] suffixes more often than [+back] suffixes.

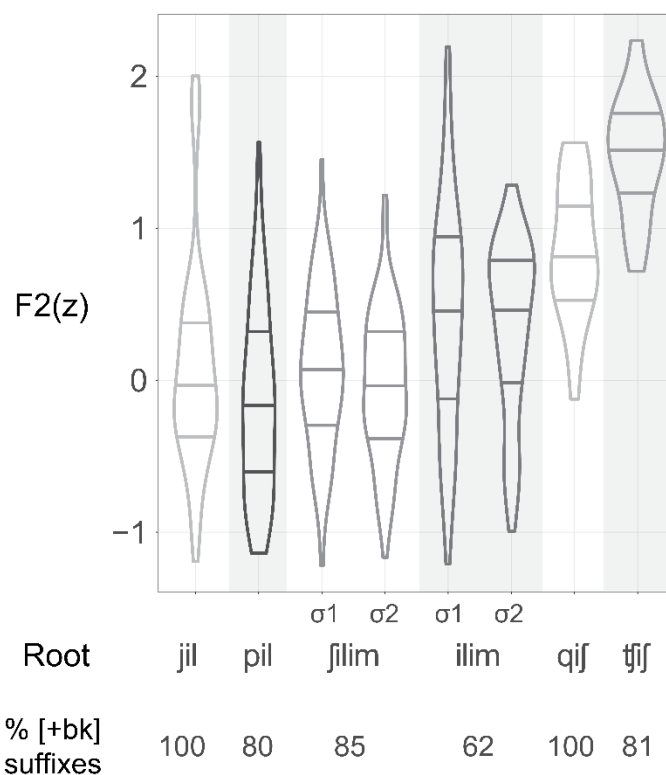


Figure 3: F2 (z) of root internal /i/ in Uyghur by lexeme and syllable number along with percent [+back] suffixes during elicitation. Within each distribution, the middle line represents the median and the two other lines indicate the first and third quartiles

These results are compared with (9) in Table 4. All four monosyllabic roots are classified as [+back] in Lindblad (1990) while the two disyllabic roots are considered [-back]. Additionally, all roots with [+back] cognates in Kyrgyz and Kazakh, i.e. /jil/ and /qil/, trigger [+back] suffixes 100% of the time, whereas roots with [-back] cognates in those languages exhibit variation in suffix selection.

Table 4: Root-internal high unrounded vowels and suffix backness

Root	Gloss	Status in Lindblad (1990)	Cognates		Count [+ back] suffix	Percent [+ back] suffix
			Kyrgyz	Kazakh		
qiʃ	winter	[+ back]	quʃ	quɯs	41/41	100
tʃiʃ	tooth	[+ back]	tiʃ	tɯs	39/48	81
jil	year	[+ back]	ɕʷul	ʒul	39/39	100
pil	elephant	[+ back]	pil	pɯl	28/35	80
ilim	science	[-back]	ilim	ɯlɯm ~ ɯlɯm	23/37	62
ʃilim	paste	[-back]	ɕelim	ʒelim	41/48	85

While it is clear in Table 4 that root-internal [i]-[u] vowels typically triggered [+ back] suffixes for these six lexical items, this does not convey whether or not there is a phonetic relationship between F2 of root-internal [i]-[u] and suffix backness. To assess the significance of F2 of root-final high unrounded vowels on the categorical backness of following low vowel suffixes (n = 153) , a mixed effect logistic regression with random intercepts for speaker and lexeme⁶, increased root-final F2 was predictive of a front vowel suffix [$\beta = 3.07$, $z = 3.49$, $\chi^2(1) = 18.75$, $p < .001$].⁷

In more gradient terms, F2 of root-internal high unrounded vowels was also predictive of suffix F2. When F2 of root-internal high unrounded vowels is higher, so is F2 of the following suffix [$t(144) = 6.07$, $r = .45$, $p < .001$]. The moderate correlation between F2 of root-internal /i/ and F2 of the following suffix is seen in Figure 4. For each lexical item, there is a positive correlation between F2 of the root-final vowel and F2 of the following low vowel suffix. Note that only a few tokens of /tʃiʃ/ and /qiʃ/ are plotted below. As indicated above, this is due to the high rate of vowel elision in these roots. Also, observe that the correlation is weakest for /jil/ and /qiʃ/; there are too few tokens of /tʃiʃ/ to truly support any meaningful correlation. However, the correlation is strongest for /ilim/, /ʃilim/, and /pil/. It may be significant that all of these lexemes have [-back] cognates in related languages.

⁶ A model with a random slope for speaker and root was also tested, but the model failed to converge.

⁷ A separate model using root-initial F2 was also tested, which resulted in the same general relationship between F2 of root-internal [i]-[u] and suffix backness [$\beta = 2.10$ $z = 2.86$, $\chi^2(1) = 11.25$, $p < .001$]. However, the model predicting suffix backness from root-final F2 provided slightly, though not significantly, better model fit [Root-final model AIC = 89.1, Root-initial model AIC = 89.6].

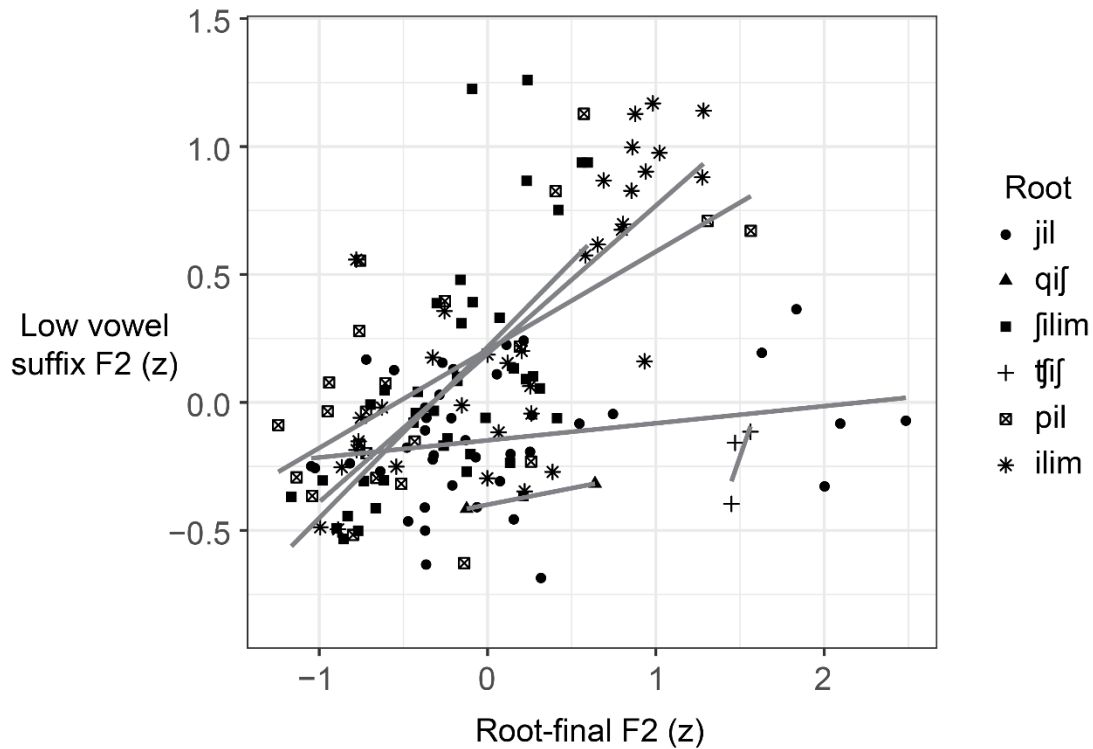


Figure 4: Correlations between F2 (z) of root-final high unrounded vowels and suffixal [ɑ]~[æ] by root

It is also worth pointing out that the correlation seen here does not necessarily indicate the directionality of the relationship. It is possible that selecting a front or back vowel suffix exerts an anticipatory coarticulatory effect on the root-internal vowel.⁸ To determine if suffix backness has a meaningful effect on root-internal F2, unaffixed forms ($n = 116$) were culled from the dataset and plotted in Figure 5. The same generalizations that characterize the distribution of F2 by lexical root in

Figure 3 hold in Figure 5, as well. Specifically, observe the difference between F2 of /ilim/ and /filim/. In unaffixed forms, mean F2 of the vowels in /ilim/ was 0.43z greater than in /filim/, again suggesting a non-allophonic relationship between the vowels in these two roots. More importantly, this suggests that suffix backness does not condition F2 of root-internal high unrounded vowels.

⁸ Thanks to Connor Mayer for pointing this out.

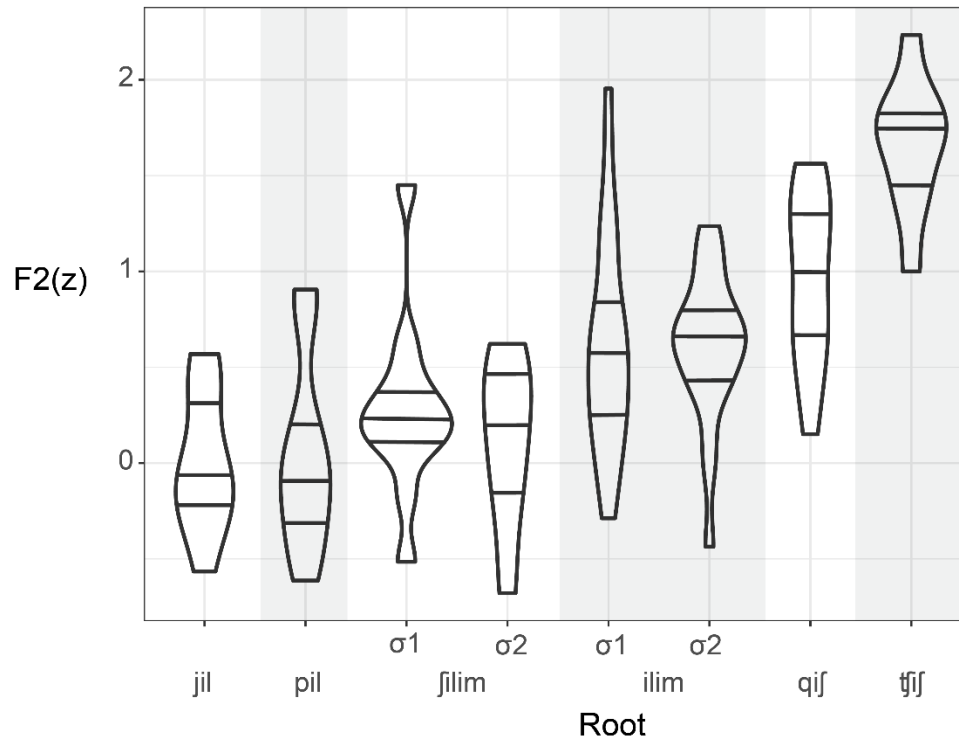


Figure 5: F2 (z) of root internal high unrounded vowels by lexeme and syllable number in unaffixed words. Within each distribution, horizontal lines represent median and quartiles.

To sum up, evidence adduced thus far suggests that F2 of high unrounded vowels in roots is not conditioned by flanking consonants only, but may exhibit lexical contrast, as well. There is some evidence that phonetic backness plays a role in suffix selection, but not to the degree one would expect for a typical contrastive pairing in harmony. In a typical pairing, one would expect root F2 to account for most of the variation in suffix backness. While root F2 is a significant predictor of suffix backness above, and while there is a significant correlation between root F2 and suffix F2, these relationships seem weaker than the relationships typically assumed for a robust phonological contrast in harmony. It is thus unclear whether /i/ has a contrastive [+back] counterpart, /ɯ/, although this seems plausible. If there is a second high unrounded vowel in the language, /ɯ/, it would help explain variation between near-minimal pairs like /ilim/ ‘science’ and /jilim/ ‘paste’, which would be more appropriately transcribed as /ilim/ and /ɣulum/. 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/ and not /ɯ/, since central and back variants are much more common, and occur with a much freer distribution than [i] (see also Hahn 1986 for discussion). When the F2 values in the figures above are consulted, it is clear that vowel qualities ranging from central to back are more common than phonetically front vowels.

5.2 Suffixal and raised /i/

This subsection investigates whether or not high unrounded vowels alternate for harmony in target positions (e.g. suffixes and in 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 unrounded vowel intervenes between a trigger and alternating target low vowel. The second examines results from the long-distance condition, where three high unrounded 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 varies with initial-syllable backness, as well as rounding. Rounding was considered because extant descriptions indicate that the epenthetic vowel of the first-person singular possessive suffix undergoes both backness and rounding harmony while the underlying high vowel of the third-person singular suffix fails to undergo either harmony. Since both the first-person singular possessive suffix and the third-person singular possessive suffix are evaluated, this second predictor, in addition to a predictor for suffix allows the model to differentiate between general effects and effects specific to the first-person singular possessive suffix. The model used predicts normalized F2 at vowel midpoint from root backness, root rounding, and suffix, with interactions between backness and roundness, and roundness and suffix, as well as a random intercept for speaker.

As seen below in Figure 6, F2 of high vowels is significantly lowered after [+back] roots [$\beta = -1.74$, $\chi^2(1) = 82.32$, $p < .001$]. The major conclusion to be drawn from this result is the high unrounded vowels do alternate for backness harmony. Thus, these vowels are not transparent to backness harmony in this condition, but alternate according to the backness of the root vowel. Second, root rounding also significantly lowers F2 of subsequent high vowels [$\beta = -0.47$, $\chi^2(1) = 14.50$, $p < .001$], suggesting that medial high vowels in this condition also alternate for root rounding. The darker grey boxes in Figure 6 show decreased F2, which is consistent with an effect of lip rounding. The interaction of backness and rounding was non-significant.

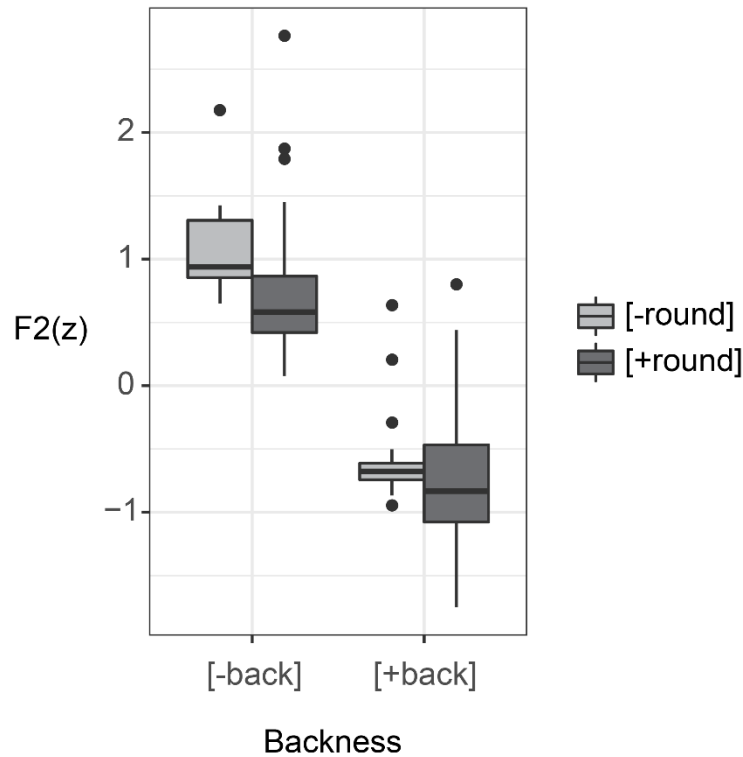


Figure 6: F2 (z) of suffix /i/ by root backness and rounding in short-distance condition

In addition to the effects above, there was also an effect of morpheme, with vowels in the third-person possessive suffix being produced with a significantly higher F2 [$\beta = 1.07$, $\chi^2(1) = 9.38$, $p < .01$], which is observable in Figure 7 below. Note also the difference in the effect of root rounding by backness context below. The effect of rounding is large for the third-person possessive, but smaller in the epenthetic vowel preceding bilabial nasal of the first-person possessive. This interaction between rounding and suffix did not reach significance, but runs counter to previous descriptions that claim that the epenthetic vowel preceding POSS.1s undergoes labial harmony although the underlying high vowel of POSS.3s does not.

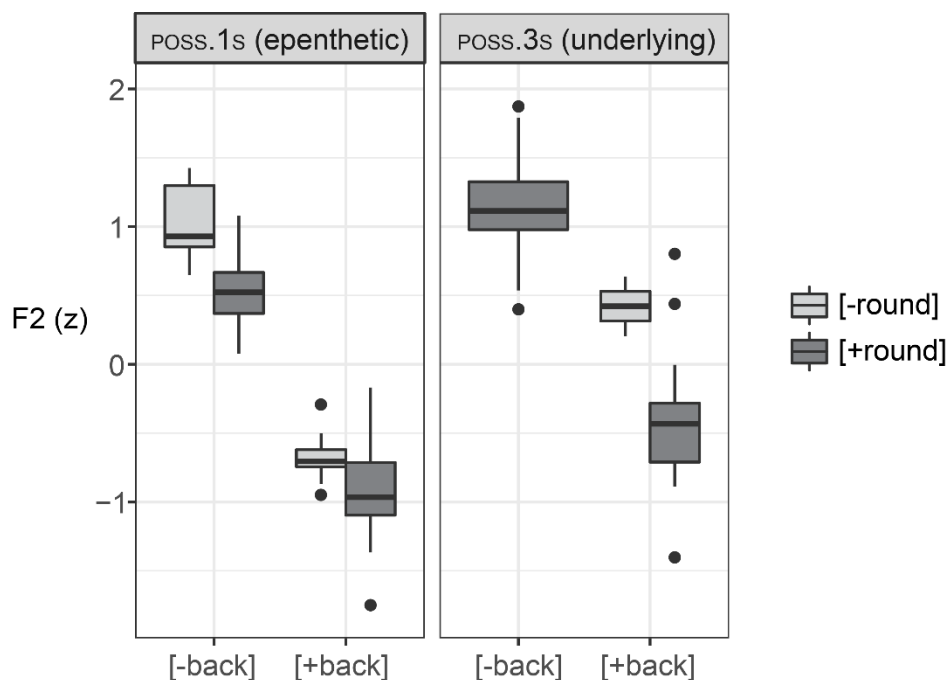


Figure 7: By-suffix plot of F2 (z) based on root backness and rounding in short-distance condition

Results for each speaker are plotted in Figure 8 below. Note that Speaker 1 did not participate in either the short- or long-distance conditions. Although the data is relatively sparse for each speaker, the effect of backness is clear. F2 of high vowels following [+back] roots is much lower than after [-back] roots, further confirming that high vowels alternate for backness harmony in the short-distance condition. In contrast, the effect of rounding appears more variable below. The depression of F2 after round vowels is most noticeable for Speakers 6, 7, and 9, and for many speakers the effect appears negligible.

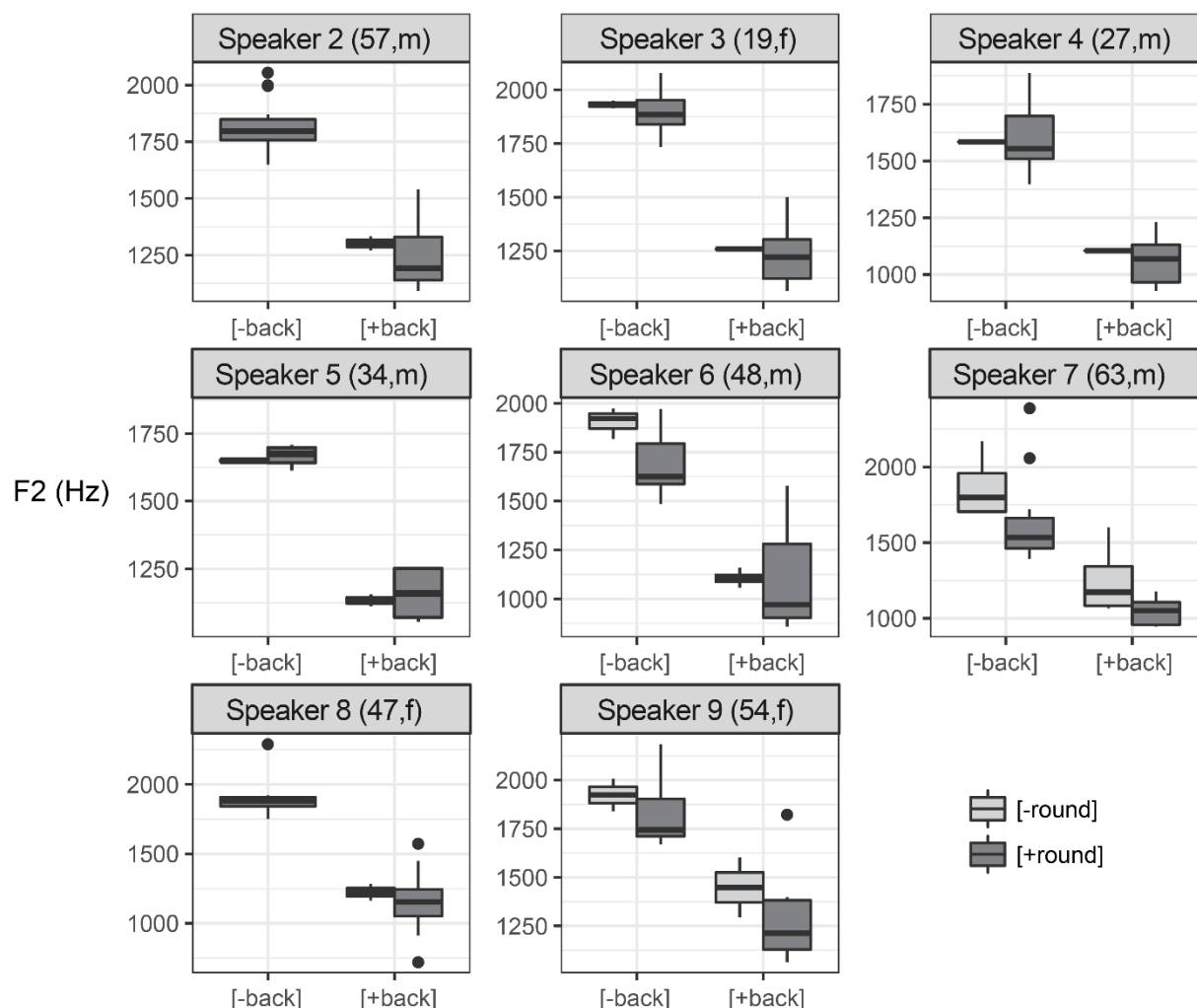


Figure 8: By-speaker plots of F2 (Hz) based on root backness and rounding in short-distance condition

Moving on, it is now crucial to determine whether or not regularly alternating suffixes, here the locative ([*-da*]~[*-dæ*]) and plural ([*-lar*]~[*-lær*]) suffixes alternate for harmony after a high vowel. The statistical model predicted vowel F2 from initial vowel backness and rounding, as well as suffix. No interactions were considered because I had no *a priori* hypotheses concerning the locative or plural suffix's effect on backness or rounding, nor any phonologically significant effect of rounding on low vowel suffixes in the variety of Uyghur under study.⁹

⁹ One would likely predict phonetic effects of consonantal context, but these are not germane for assessing whether these suffixes alternate between /a/ and /æ/ after high vowels. If these suffixes do alternate for harmony, the effect should be large enough to ignore smaller phonetic effects of context.

In Figure 9 it is clear that the low vowels alternate for backness harmony. Statistically, F2 of the locative and plural suffixes is significantly lowered after [+back] roots [$\beta = -0.96$, $\chi^2(1) = 73.22$, $p < .001$]. Unlike the high vowels, F2 of the low vowels was not significantly affected by root rounding, which is consistent with descriptions of the dialect under study, in which rounding harmony targets high vowels only.

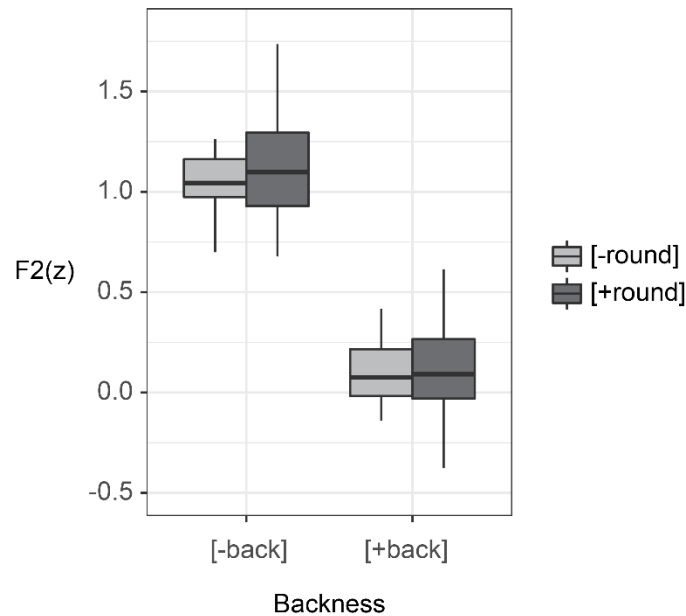


Figure 9: F2 (z) of suffix low vowels by root backness and rounding in short-distance condition

To summarize this subsection, the key finding is that high vowels alternate for both backness and rounding, although the effect for rounding varies greatly between speakers. The acoustic effect is much larger for backness, but that is not surprising since the acoustic effects of lip rounding are typically much smaller than the effects of tongue body backing. 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 2 and 3, and the underlying high vowel of the third-person singular possessive suffix occurred in the fourth syllable. As in Section 5.2.1, a linear mixed effects model was used to predict normalized F2 at vowel midpoint from initial-vowel backness and rounding. An additional fixed effect of syllable number was included in the model to

determine if the second and third high vowels in these sequences pattern like the first high vowel. The model also included interactions between backness and rounding, backness and syllable number, and rounding and syllable number, along with a random intercept for speaker.

As in the short-distance condition, F2 of high vowels ($n = 295$) was significantly lowered after [+back] roots [$\beta = -2.52$, $\chi^2(1) = 235.53$, $p < .001$]. There was a significant effect of syllable number, with F2 of front vowels decreasing as syllable number increases [$\beta = -0.18$, $\chi^2(1) = 11.63$, $p < .001$]. Additionally, there was a significant interaction between backness and syllable number, indicating an asymmetric increase in back vowel F2 in later syllables [$\beta = 0.65$, $\chi^2(1) = 102.24$, $p < .001$]. In other words, F2 of front vowels decreased by 0.18z across syllables while F2 of back vowels increased by 0.47z ($0.65 - 0.18 = 0.47$) across syllables. The magnitude of these positional shifts is quite noticeable in Figure 10. In essence, 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, 2019b). No other effects were significant.

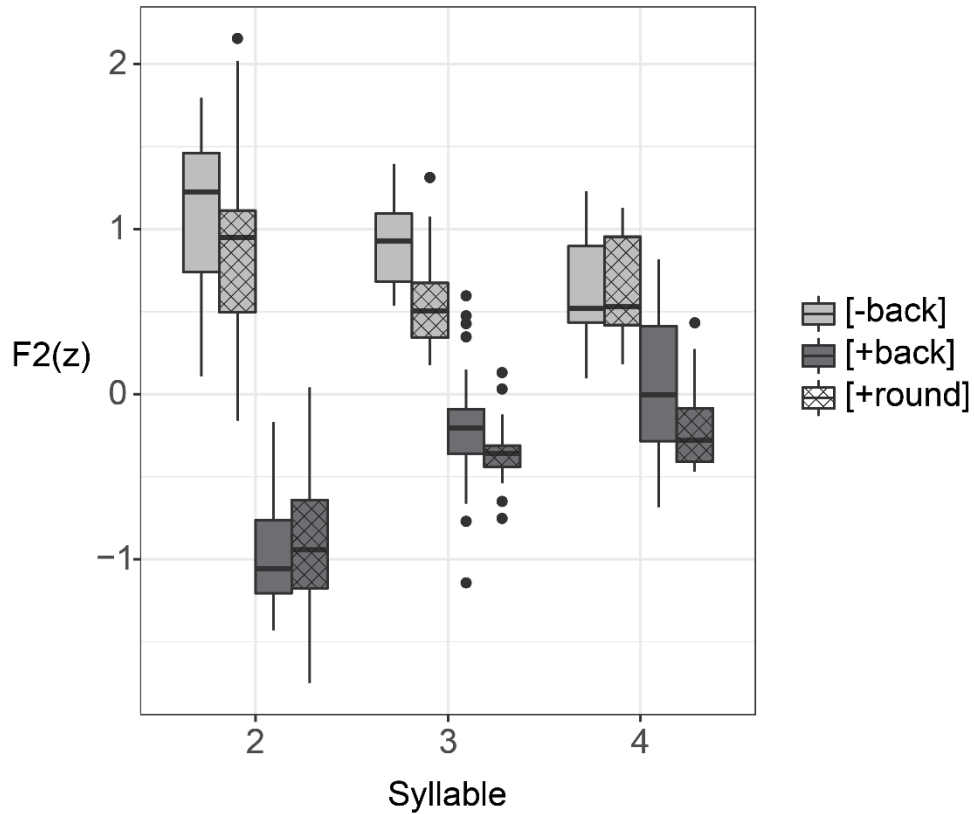


Figure 10: F2 (z) of high vowels by root backness and rounding in long-distance condition

Individual results for each speaker are presented in Figure 11. Observe that for all speakers, the high vowels maintain an acoustic backness distinction across all syllables. Despite this generalization, individual variation is significant. Some speakers produce relatively symmetrical centralization across the three relevant syllables, e.g. Speakers 2 and 8), while some speakers produce the asymmetrical fronting pattern, e.g. Speakers 3 and 7. Interestingly, Speaker 6 fronts all vowels in later syllables.

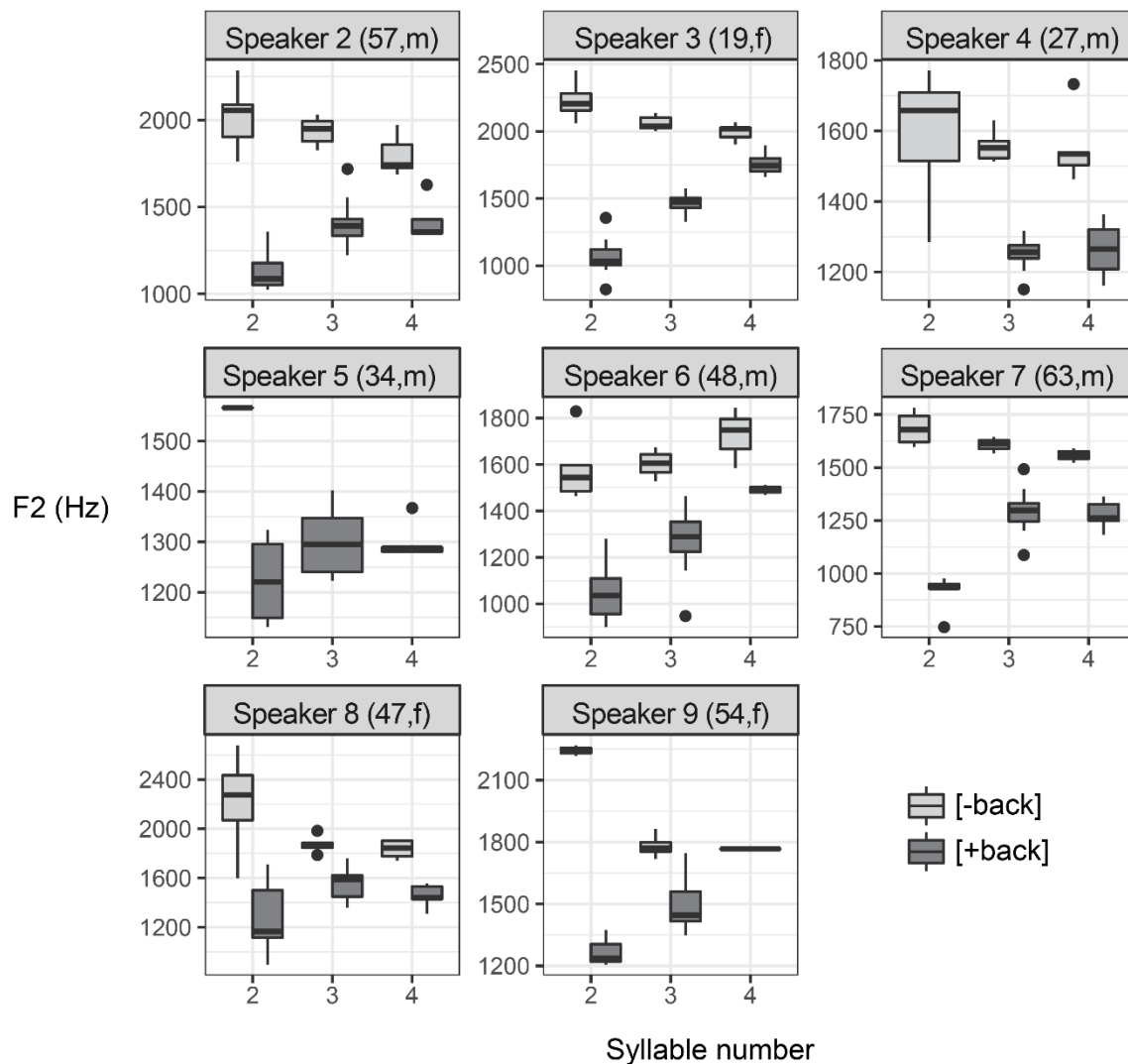


Figure 11: By-speaker plots of F2 (Hz) based on root backness in long-distance condition

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

As in the short-distance condition, F2 of the low vowels was significantly lowered by after [+back] roots [$\beta = -1.18$, $\chi^2(1) = 96.88$, $p < .001$]. In other words, low vowels alternate for backness harmony in the long-distance condition, shown in Figure 12. No other effects were significant.

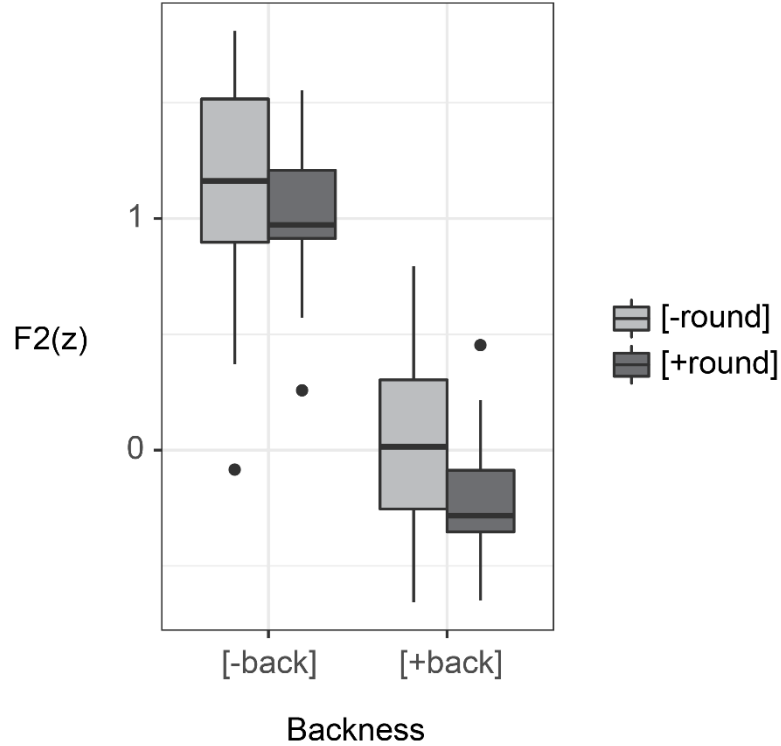


Figure 12: F2 (z) of suffix low vowels by root backness and rounding in long-distance condition

In summary, this subsection has shown that in the long-distance condition, reportedly transparent vowels actually alternate for backness harmony, with no significant effects of rounding harmony. Like the high vowels, low vowels following the high vowels also alternate for harmony. 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, and at least some also exhibit rounding alternations, too.

6 Discussion

The findings from the previous section provide clarity, but also pose intriguing problems for the analysis of harmony in Uyghur. Results are compared to the predictions of the various possible analyses from Table 2 below in Table 5. The behavior of the high unrounded vowels is clear in target positions – they alternate for harmony. As a result, the

transparency-based analyses developed in Lindblad (1990) and Vaux (2000) are unnecessary. Specifically, Lindblad's appeal to a late absolute neutralization rule incorrectly predicts the absence of [u]. As for Vaux (2000), there is no need to appeal to contrast to delimit the operation of the harmony rule, since harmony outputs both [i] and [u]. From a computational perspective, the results above suggest that backness harmony targets all non-initial vowels, while rounding harmony only variably targets high vowels. As a result, the two sets of sounds participating in harmony do not partially intersect, but rather exist in a superset relation. This in turn suggests that learning the harmony pattern is simpler, and moreover, consistent with other known languages with multiple harmony patterns (Aksenova, 2018; Aksenova & Deshmukh, 2018).

The real issue is how to analyze [i] and [u] within roots. Some lexically-specific patterns above, e.g. the difference in F2 for pairs like [ilim] 'science' and [ɟulɪm] 'paste', suggest that these two surface phones are not simply allophones of a single phoneme. However, the general distribution of these sounds in the lexemes tested does not unequivocally support a contrastive relationship between these two sounds, in part because the lexical items with the highest F2 (e.g. /qɪʃ/ 'winter') always triggered [+back] suffixes. If the high unrounded vowels are contrastive, then Uyghur parallels Moro, exhibiting a relatively subtle contrast that manifests itself both in triggers and targets of harmony. However, if the high unrounded vowels are not contrastive within roots, they would only exhibit phonetic alternations in suffixes. If this is the case, then the language resembles Kinande, possessing a surface inventory that is larger than its underlying inventory. Among these two, results suggest that contrastive /i/-/u/ over allophonic [i]-[u] due to the correlations between root-internal vowel backness and suffix backness. If these two are allophonically distributed, then there is no reason to predict that contextual, non-contrastive differences would correspond to differences in suffix selection.¹⁰ It is clear, though, that high vowels are not transparent to harmony, and although there is a gradient reduction in acoustic contrast across syllables, there is no evidence to support a Hungarian-like distance effect in the language.

¹⁰ In Optimality Theory (Prince & Smolensky, 2004), contrast is not a theoretical entity, but rather an epiphenomenon of constraint rankings. As a consequence, initial-syllable allophonic differences are predicted to matter for harmony. If, for instance, a markedness constraint against *il sequences operates over lexical roots, e.g. /jil/ 'year', then this could produce output [ɟul], whose output backness would drive suffix backness. This would likely predict front vowel suffixes after [i] sequences, but this does not conform to the attested data, e.g. [qɪʃ-ta] 'winter-LOC.'

Table 5: Schema of possible results compared against actual results

	Contrast neutralization	Allophonic participation	True transparency	Distance- based pseudo- transparency	Actual results
Is F2 of trigger [i]-[ʊ] predictable based on consonantal context alone?	No	Yes	Yes	Partially	Unclear
Is there a relationship between F2 of trigger [i]-[ʊ] and suffix backness?	Yes	No	No	Yes, but small	Yes, but small
Is F2 of target [i]- [ʊ] predictable based on preceding vowel backness?	Yes	Yes	No	Yes, but differences are small	Yes
Is F2 of low vowels following target [i]-[ʊ] predictable based on initial vowel backness?	Yes	Yes	Yes	Yes, modulo distance	Yes
Comparison language	Moro	Kinande	Wolof	Hungarian	

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 of sounds, shown below in Table 6. 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 [ɪ] as examples of such a salient but non-contrastive relationship. On the other hand, some sounds may exhibit structurally different phonological behavior while being characterized by 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, Baković, Mai, & Meinhardt, 2019). Despite their surface neutralization, these sounds still maintain their abstract structural contrast for [high] for both rounding and ATR harmony.

Table 6: A typology of contrast and distinctiveness (Kiparsky 2015)

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

Despite the appeal of Kiparsky’s proposal, Uyghur [i] and [ɯ] don’t fit well into these categories. The acoustic differences between these two sounds are not terribly salient, as attested by the number of grammars that describe only /i/ in the language. However, the behavior of these two surface sounds is not equivalent to the abstract contrasts in Tutrugbu, /ɪ/-/ɛ/ and /ʊ/-/ɔ/, since clear acoustic differences are maintained. One possibility is that the contrast between these two sounds 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. For the case at hand, [i] and [ɯ] are partially conditioned by flanking consonants in triggering positions. For instance, preceding /l/, [ɯ] is much more common in the initial syllable. However, in suffixes this consonantal conditioning appears ineffectual, since both [il] and [ɯl] 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 *[sælluɯlær]. Stated somewhat differently, [i] and [ɯ] are conditioned by lexical contrast, as well as adjacent consonants and preceding vowel

backness. In initial syllables, preceding vowel backness is absent, leaving adjacent consonants to condition variation between [i] and [u] up to a point. Beyond that, accounting for the different realizations of [ilim] and [ʃulum] suggests the need for lexical contrast. In suffixes, though preceding vowel backness is relevant, overcoming (almost) all potential consonantal effects on vowel quality as well as potential lexical contrasts between /i/ and /u/ in non-initial syllables.

Another possibility is that both the Uyghur system and other outliers, like Russian [i] and [ɨ], derive their behavior from other factors. For instance, Russian [i] and [ɨ] 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, what role does a second language play 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. 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.

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, presumably 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, Garellek, Gillingham, Ackerman, & Malouf, 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 down to minimal pairs and semantic differences? If so, then Uyghur [i] and [u] might not be considered contrastive, since they do not have clear-cut (near-)minimal pairs, although perhaps the

differences in [ilim] ‘science’ and [ʃulum] ‘paste’ suggest the possibility of such pairs. It is my hope that further work can continue 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, in contrast to a number of previous claim. As a result, the formal mechanisms proposed for the analysis of harmony in Uyghur are unnecessary. 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). Typologically, the locality of backness harmony in Uyghur falls nicely in line with previous work investigating putative transparency in harmony, showing that harmony applies more locally than many previous accounts suggest. In addition to issues of locality, the paper raises the question of contrast, how it manifested phonetically, and the potential role for degrees of contrast (Hall 2009, 2013) intermediate between canonical contrast and allophony.

Acknowledgments:

I would like to thank Sharon Rose, Eric Baković, Sarah Creel, Marc Garellek, and Rachel Walker, as well as the audience at the 26th Manchester Phonology Meeting for discussions and a veritable cornucopia of insightful comments. 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.
- Aksenova, A. (2018). *The Hitchhiker's Guide to Harmony Interactions*. 41st Generative Linguistics in the Old World (GLOW). poster. <https://www.aaksenova.com/materials/>
- Aksenova, A., & Deshmukh, S. (2018). *Formal restrictions on multiple tiers*. The Society for Computation in Linguistics. poster. <https://www.aaksenova.com/materials/>
- Archangeli, D., & Pulleyblank, D. (1994). *Grounded phonology*. MIT Press.
- 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>
- Benus, S., & Gafos, A. I. (2007). Articulatory characteristics of Hungarian “transparent” vowels. *Journal of Phonetics*, 35(3), 271–300. <https://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. doi: 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. doi: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://doi.org/10.1515/tlr-2013-0008>
- Hayes, B., & Londe, Z. C. (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology*, 23(1), 59–104. doi:10.1017/S0952675706000765
- Hayes, B., & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39(3), 379–440. <https://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
- Heinz, J., & Lai, R. (2013). Vowel harmony and subsequentiality. *Proceedings of the 13th Meeting on the Mathematics of Language (MoL 13)*, 52–63.
<https://www.aclweb.org/anthology/W13-3006/>
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal*, 50(3), 346–363. <https://doi.org/10.1002/bimj.200810425>
- Kaydarov, A. T. (1970). Uygurovedenie v Kazaxstane [Uyghur studies in Kazakhstan]. In A. T. Kaydarov, G. Sadvakasov, & T. Talipov (Eds.), *Issledovaniya po uygurskomu jazyku [Research on the Uyghur language]* (Vol. 2, pp. 21–30). Nauka.
- Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology Yearbook*, 2(1), 85–138. doi: 10.1017/S0952675700000397
- Kiparsky, P. (2015). Phonologization. Patrick Honeybone and Joseph Salmons (Eds.), *The Oxford handbook of historical phonology*. Oxford University Press.
doi: 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. doi:10.1017/S0954394500000442
- 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://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://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>

- 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*.
<http://dx.doi.org/10.3765/amp.v7i0.4565>
- McCollum, A. G., Baković, E., Mai, A., & Meinhardt, E. (2019). The expressivity of segmental phonology and the definition of weak determinism. UC San Diego.
<http://www.adammccollum.net/research.html>
- 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.
- Nevins, A. (2010). *Locality in vowel harmony*. MIT Press.
- 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.
doi:10.1017/S0022226715000183
- Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint interaction in generative grammar*. Wiley-Blackwell. doi: 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., & 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. doi: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://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://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://doi.org/10.1080/23273798.2017.1359634>
- Szeredi, D. (2016). Exceptionality in vowel harmony (PhD dissertation). New York University.
<https://ling.auf.net/lingbuzz/003148>

- 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).
- 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.
doi:10.1017/S0025100300006423
- Walker, R., Byrd, D., & Mpiranya, F. (2008). An articulatory view of Kinyarwanda coronal harmony. *Phonology*, 25(3), 499–535. doi: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. doi:10.1017/S0952675707001157