Vowel Dispersion and Kazakh Labial Harmony

Adam G. McCollum,
UC San Diego

Abstract:

This paper uses novel data showing gradient labial harmony in Kazakh to compare Kaun's (1995) feature-based analysis with a dispersion-based analysis in a Maximum Entropy Harmonic Grammar. The paper demonstrates that the dispersionary analysis better fits the Kazakh data than Kaun's analysis and then extends the dispersionary analysis to account for four languages with harmony patterns different from that in Kazakh. The paper also argues that the dispersion-based account provides a better analysis of the typology of labial harmony than Kaun's feature-based analysis.

1 Introduction

Cross-linguistically, labial (or rounding) harmony is subject to a number of restrictions (Steriade 1981; Kaun 1995, 2004; Walker 2001). Most importantly, the class of trigger and target vowels is restricted by height and/or backness. Kaun (1995, 2004) contends that the best triggers are perceptually weak, with harmony extending the temporal span of these weak contrasts. Thus, vowels whose [rd] feature is only weakly realized are more likely to trigger harmony, or in her terms, "bad vowels spread" (1995:1).

In this paper I argue, like Kaun (1995, 2004), that perceptual weakness motivates labial harmony. However, I contend that perceptual weakness is more appropriately defined in dispersion-based terms (Liljencrants & Lindblom 1972), and not the feature-based representations used in Kaun's analysis. I present four kinds of evidence to justify this claim. First, using gradient labial harmony data from colloquial Kazakh, I show that the dispersion-based model outperforms Kaun's model in a Maximum Entropy Harmonic Grammar (Goldwater & Johnson 2003). Second, I show that the proposed model also accounts for categorical harmony in four non-Turkic languages. Third, I discuss labial harmony in Mayak (Andersen 1999), which cannot be accounted for under Kaun's analysis but falls out naturally from the dispersion-based analysis. Lastly, I outline the typological predictions of the dispersion-based analysis, which provide new insight into the typology of labial harmony missed by Kaun's analysis.

The paper is organized as follows. In §2, I describe vowel harmony in colloquial Kazakh. In §3, I discuss previous work on labial harmony, in particular, Kaun (1995, 2004). In §4, I analyze Kazakh using Kaun's feature-based analysis, and in §5, I demonstrate that the dispersion-based analysis is superior. §6 extends the analysis to account for harmony in four non-Turkic languages, and §7 discusses the typological predictions of the proposed analysis. In §8, I relate this analysis to other dispersion-theoretic work, and in §9, I conclude the paper.

2 Kazakh Vowel Harmony

2.1 Background and Vowel Inventory

Kazakh, like most Turkic languages, exhibits palatal, or backness harmony, and labial harmony, which is the focus of this paper. Complex words are formed via suffixation. As a result, these two harmonies extend rightward from roots to suffixes.

Writers have disagreed on the number of vowel phonemes in the language, ranging from eight to eleven (Menges 1947; Dzhunisbekov 1972; Kirchner 1998; Kara 2002; McCollum 2015; Muhamedowa 2015). Writers typically agree, though, on a common set of eight vowels, often transcribed as /ɑ ɯ ɔ ʊ e ɪ ø ʏ/. I transcribe these vowels as /ɑː ə ɔː ɔ ɪː ɪ ʏː ʏ/, with a length distinction replacing what has traditionally been a height distinction. I justify this through an acoustic analysis of the vowel inventory presented below. Three additional vowels might be phonemic in the language, /i æ u/. As these three vowels are lexically rare and peripheral to the harmony system (Dzhunisbekov 1972:39-41; Kirchner 1998:322), they are excluded from further discussion.

Data was collected from eleven native speakers, (mean age 33.5 yrs, range 19-46 yrs) in southeastern Kazakhstan, using a semi-formal, conversationally-based elicitation in the target language. Acoustic data is presented below. Root vowel data (N=2,490) was elicited to assess the general shape of the vowel space. Mean F1-F2 for the eight vowels under study is shown in Figure 1.

As noted above, the distinction between high and mid vowels in Turkic has been neutralized in colloquial Kazakh. The historical high vowels have lowered, resulting in a length distinction where historical mid vowels are long and historical high vowels are short (Johanson 1998:94; Kara 2002:9).

Figure 1: Mean root vowel F1-F2 (z) with 1 standard deviation ellipses

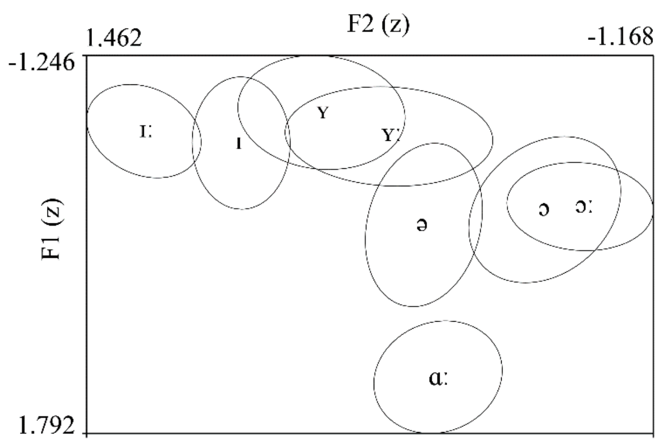


Figure 1 shows that vowel height, defined in terms of F1, is insufficient to differentiate these vowels. Instead, vowel length distinguishes the historical high from non-high vowels. The historical high vowels are significantly shorter (Dzhunisbekov 1972:75; Kirchner 1998:319;

Washington 2016; McCollum & Chen 2016). Table 1 presents duration data (N=32) for eight vowels produced in analogous environments by Speaker 11.

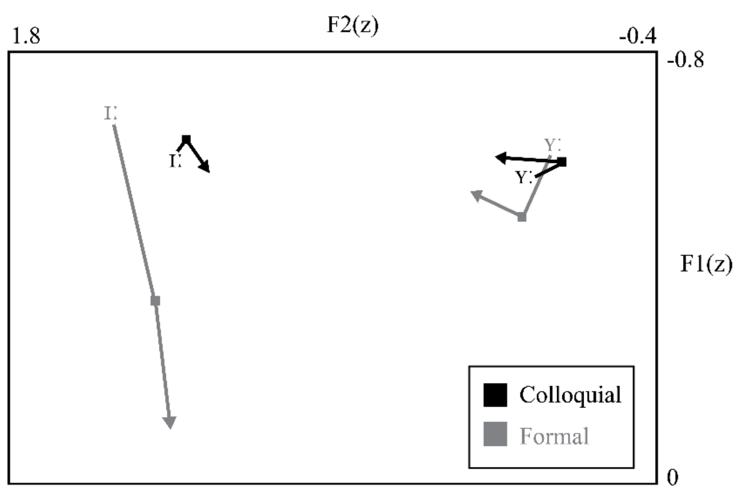
Table 1: Mean duration (in ms. with standard deviations) of initial-syllable vowels

Vowel	Duration (SD)	Vowel	Duration (SD)
ʏ	22.5 (8.0)	ɔ	43.3 (14.5)
ɪ	47.7 (10.9)	ə	30.1 (12.5)
ʏː	90.1 (18.5)	ɔː	78.4 (27.1)
ɪː	69.7 (23.0)	ɑː	62.4 (8.0)

When averaged within the two putative lengths, mean duration was 35.9 ms for the short vowels, and 75.2 ms for the long vowels. This is further confirmed in McCollum & Chen (2016), which found the average duration of initial-syllable /ə/ and /ɑː/ were 32.5 and 78.5 ms, respectively. The absolute duration of the Kazakh vowels is short, but the duration ratio of long-to-short vowels in the fieldwork data is 2.1:1, which parallels length contrasts in a number of other languages (Hadding-Koch & Abramson 1964; Abramson & Ren 1990).

In the formal register, the long vowels /ɪː ʏː ɔː/ are reportedly distinguished from the short vowels by diphthongization, particularly in word-initial position. In formal speech, /ɪː/ is realized as [jɪː], while /ʏː/ and /ɔː/ are realized as [wʏː], and [wɔː], respectively (Dzhunisbekov 1980:21; Vajda 1994:619-624). In colloquial speech, though, these vowels do not typically surface as diphthongs. Figure 2 compares the realization of initial-syllable /ɪː ʏː/ in colloquial speech with formal speech collected in the UC San Diego phonetics lab. Vowels were measured in matching syllabic contexts, and measurements were taken at the 25, 50, and 75% points of each vowel. In colloquial speech, F1-F2 of both vowels show only slight movement, while in formal speech, the spectral properties of /ɪː/ are far more dynamic.

Figure 2: Initial-syllable /ɪː ʏː/ in colloquial and formal Kazakh



With these facts in mind, I treat these vowels as monophthongs, and use the features, [back], [round] and [long] to represent the eight vowels that participate in harmony, shown in Table 2.

Table 2: Feature chart

	-bk		+bk	
	-rd	+rd	-rd	+rd
[-long]	ɪ	ʏ	ə	ɔ
[+long]	ɪː	ʏː	ɑː	ɔː

2.2 Vowel Harmony

2.2.1 Palatal Harmony

Kazakh words usually contain vowels from only one palatal set.[1] As palatality is determined by the root, and labiality is at least partially dependent on the root, only one feature, length, is fully contrastive in non-initial syllables. In (1-2), both tauto- and heteromorphemic palatal harmony are presented, assuming that non-initial root vowels are, like suffixes, targets for harmony (Harrison & Kaun 2000; Kabak & Weber 2013; cf. Clements & Sezer 1982).

In (1), all possible combinations of vowel length are exemplified for unrounded roots. Harmony among front vowels is shown in (1a-d), and harmony among back vowels is demonstrated in (1e-h).

(1) Root-internal harmony

a.	tɪzɪm	'list'	e.	qəzəɫ	'red'
b.	tɪlɪːk	'wish'	f.	qərɑːn	'hawk'
c.	ɪːsɪk	'door'	g.	qɑːzə	'horse sausage'
d.	ʒɪːbɪː	'arrow'	h.	qɑːrɑː	'black'

In (2), harmony between unrounded roots and suffixes is shown. In the locative suffix (2a-d), long back [ɑː] alternates with long front [ɪː]. In the accusative suffix (2h-j), short back [ə] alternates with short front [ɪ]. Furthermore, iterative palatal harmony is exemplified in (2e-g).

(2) Suffix harmony

a.	bɪz-dɪː	'1P-LOC'	h.	bɪz-dɪ	'1P-ACC'
b.	kɪːz-dɪː	'time-LOC'	i.	kɪːz-dɪ	'time-ACC'
c.	qəz-dɑː	'girl-LOC'	j.	qəz-də	'girl-ACC'
d.	qɑːz-dɑː	'goose-LOC'	k.	qɑːz-də	'goose-ACC'

e. tɪzɪm-dɪː-gɪ-lɪːr-ɪ-m-ɪz-dɪː 'list-LOC-REL-PL-POSS-1-PL1-LOC'
f. tɪrkɪː-l-dɪ-ŋ-ɪz-dɪːr 'register-PASS-PST-2-FORM-PL2'
g. qərɑːn-dɑːr-ə-m-əz-dɑː 'hawk-PL-POSS-1-PL1-LOC'

In summary, palatal harmony in Kazakh is iterative and not sensitive to length restrictions. As will be evident in the next subsection, labial harmony is subject to restrictions not enforced on palatal harmony.

[1] Exceptions include compounds, unassimilated loans, and a few invariant morphemes.

2.2.2 Labial Harmony

Menges (1947:59-64) and Korn (1969:101-102) report that labial harmony is robust, but more recent writers have noted increasing restrictions on harmony (Dzhunisbekov 1972; Kirchner 1998; Kara 2002; McCollum 2015).[2] Specifically, labial harmony in contemporary Kazakh is confined to a smaller domain of application (Balakaev 1962:100-105; Dzhunisbekov 1972; Vajda 1994:633-634; McCollum 2015; cf. Abuov 1994), and harmony occurs more consistently within roots than across morpheme boundaries (Vajda 1994; McCollum 2015).

Within roots, several generalizations emerge from (3). First, short vowels (3a-d) are more likely to undergo harmony than long vowels (3e-h). Second, harmony is generally triggered by the front vowels, /ʏ ʏː/, as well as the short back vowel, /ɔ/, but not by the long back vowel, /ɔː/, as in (3d). The frequency data reported below are counts of impressionistic transcriptions after initial data collection. These judgements were then confirmed by comparing acoustic properties of non-initial vowels to those of underlying (i.e. initial) vowels (Zsiga 1997:234-235).

(3) Root-internal harmony

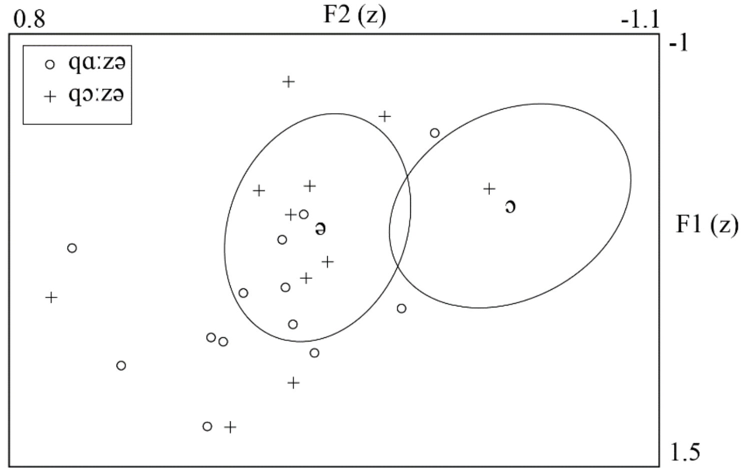
	Attested variants	Frequency of harmony	Gloss
a.	ʒʏzʏk ~ ʒʏzɪk	(10/11)	'ring'
b.	kʏːsʏk ~ kʏːsɪk	(10/11)	'desert carrot'
c.	qɔɫɔn ~ qɔɫən	(16/17)	'colt'
d.	qɔːzɔ ~ qɔːzə	(1/10)	'lamb'
e.	tʏlʏːk ~ tʏlɪːk	(1/11)	'graduate'
f.	tʏːbɪː	(0/11)	'hill'
g.	qɔɫɑːq	(0/10)	'ear'
h.	bɔːɫɑːt	(0/10)	'steel'

The inertness of /ɔː/ as a trigger for harmony is further shown below. Figure 3 plots F1-F2 of the second-syllable vowels in /qɑːzə/ 'horse sausage', represented by ○, and /qɔːzə/ 'lamb', represented by +. Assuming the F1-F2 values of vowels derived from harmony approximate those of underlying vowels, then with the input, /qɔːzə/ labial harmony should produce vowels more similar to /ɔ/ than /ə/.[3] However, in almost every case, F1-F2 of the short back vowels after /ɔː/ more closely resemble initial /ə/. The second-syllable vowel after /ɔː/ approximates F1-F2 of initial /ɔ/ only once.

[2] Abuov (1994) describes literary Kazakh. Menges and Korn do not identify the register described.

[3] While lip rounding depresses both F2 and F3 (Ladefoged 2001:41, 46), F3 is not a reliable predictor of rounding in Kazakh (McCollum 2015).

Figure 3: F1-F2 scatterplot of second-syllable /ə/ in [qɔːzə], represented by +, and [qɑːzə], represented by ○, (N=24) compared to mean /ə/ and /ɔ/ in initial syllables, shown with 1 standard deviation ellipses (N=365).



Labial harmony on suffixes, shown in (4), is even more restricted. The short vowel of the past tense suffix (4a-h) and the long vowel of the conditional suffix (4i-l) typically surface as unrounded regardless of initial vowel roundness. In longer words (4m-o) the same generalizations hold- outside of roots, harmony is infrequent, even after /ʏ/. If every surface [ʏ] vowel triggered harmony, we would expect the accusative to undergo assimilation in (4m), [ʒʏzʏk-tɪ] 'ring-ACC.' Since it does not, harmony is triggered by the initial syllable only.

(4) Suffix harmony

	Attested variants	Freq. of harmony	Gloss
a.	kʏl-dʏ ~ kʏl-dɪ	(7/22)	'laugh-PST.3'
b.	ʏːl-dʏ ~ ʏːl-dɪ	(1/20)	'die-PST.3'
c.	qɔr-dɔ ~ qɔr-də	(1/18)	'construct-PST.3'
d.	qɔːs-tɔ ~ qɔːs-tə	(1/20)	'add-PST.3'
e.	kʏl-dʏ-m ~ kʏl-dɪ-m	(4/12)	'laugh-PST-1S'
f.	ʏːl-dʏ-m ~ ʏːl-dɪ-m	(3/11)	'die-PST-1S'
g.	qɔr-dɔ-m ~ qɔr-də-m	(3/10)	'construct-PST-1S'
h.	qɔːs-tɔ-m ~ qɔːs-tə-m	(2/12)	'add-PST-1S'
i.	kʏl-sɪː	(0/15)	'laugh-COND'
j.	ʏːl-sɪː	(0/14)	'die-COND'
k.	qɔr-sɑː	(0/13)	'construct-COND'
l.	qɔːs-sɑː	(0/13)	'add-COND'
m.	ʒʏzʏk-tʏ ~ ʒʏzʏk-tɪ	(1/9)	'ring-ACC'
n.	ʒʏzʏk-tɪːr-dɪ	(0/23)	'ring-PL-ACC'
o.	qɔr-dɔ-ŋ-əz-dɑːr ~ qɔr-də-ŋ-əz-dɑːr	(1/13)	'construct-PST-2-FORM-PL2'

The data in (4) above suggest a coarticulatory effect of labial consonants, since frequency of harmony increases in (4e-h), where the suffix is a labial consonant. However, labials do not, on their own, trigger harmony, as in words like [tɪzɪm], (1a), and [ʒɪːbɪː], (1d), (see Clements & Sezer 1982 for Turkish). More generally, the data in (3), along with Figure 3, indicate that /ɔː/ does not usually trigger harmony. In contrast, /ʏ/ /ɔ/ and /ʏː/ do trigger harmony to varying degrees.

Across morpheme boundaries, harmony is more frequently attested in a class of suffixes like the converbial and passive-reflexive suffixes, /-p/ and /-n/~/-l/, respectively. The converbial suffix is realized as /-p/ in (5a-b). After consonant-final roots, this suffix results in an illicit coda cluster, which is repaired by epenthesis (5c-h). The epenthetic vowel is always short, and agrees with the root in backness, and optionally in rounding (5e-g). After /ʏ/, /ɔ/, /ʏː/, and /ɔː/, the epenthetic vowel is rounded 84%, 72%, 54%, and 7% of the time, respectively.

(5) Converbial suffix

	Attested variants	Frequency of harmony	Gloss
a.	qɑːrɑː-p		'look-CVB'
b.	sʏːjlɪː-p		'speak-CVB'
c.	qɑːt-əp		'harden-CVB'
d.	kɪːs-ɪp		'cut-CVB'
e.	kʏl-ʏp ~ kʏl-ɪp	(26/31)	'laugh-CVB'
f.	qɔr-ɔp ~ qɔr-əp	(13/18)	'construct-CVB'
g.	ʏːl-ʏp ~ ʏːl-ɪp	(10/18)	'die-CVB'
h.	qɔːs-ɔp ~ qɔːs-əp	(1/15)	'add-CVB'

In summary, labial harmony typically applies to short vowels within roots, except after /ɔː/. Harmony applies less often to suffixes. Initial vowels exert varying degrees of assimilatory force, with /ʏ/ triggering harmony more often than /ɔ/, /ʏː/, or /ɔː/. From the data presented above, the trigger strength hierarchy emerges: ʏ > ɔ > ʏː > ɔː, or perhaps ʏ > {ɔ, ʏː} > ɔː, since the difference between /ɔ/ and /ʏː/ is small. Generally, short vowels are better triggers than long vowels, and front vowels are better triggers than back vowels. In the following section I relate these findings to the cross-linguistic asymmetries discussed in Kaun (1995, 2004).

3 Background

3.1 Labial Harmony

Most analyses of labial harmony have converged on two separate generalizations. First, labial harmony is parasitic, in that its application depends on the agreement of some other feature(s), in many cases, vowel height (Ultan 1973:55; Steriade 1981; Cole & Trigo 1988; van der Hulst & Smith 1988; Cole & Kisseberth 1994:§5; van der Hulst & van der Weijer 1995:523; Finley 2008b:§9; Jurgec 2011:ch. 8). For instance, in Kachin Khakass (Korn 1969), labial harmony applies only when both trigger and target are high.

Second, phonetic grounding has factored significantly into numerous accounts of labial harmony (Kaun 1995, 2004; Walker 2001, 2011; Finley 2008a). Kaun (1995:1) argues that "bad vowels spread", proposing a strong relationship between a vowel's perceptual weakness and its propensity to trigger harmony. Typologically, Kaun connects cross-linguistic patterns of labial harmony with the phonetic properties of round vowels (cf. Cole & Trigo 1988; Nevins 2010; Ko 2012 for analyses without functional grounding).

In many languages, restrictions on labial harmony are assumed to be featural, but Vaux (1993) suggests that labial harmony is also constrained by systemic factors. He contends that, in Turkic, harmony targets non-high vowels in languages with three distinct vowel heights but not in languages with only two phonological heights. According to Vaux, the harmony pattern is, in part, derivable from the nature of the inventory. The use of systemic considerations will play a crucial role in the analysis presented in §5.

3.2 Typological Generalizations

As noted above, Kaun argues that perceptually weak vowels are better triggers for harmony. Conversely, she contends that perceptually salient vowels are better targets. In this way, harmony serves to extend the temporal span of a weak contrast to improve chances of reliable discrimination (Suomi 1983). She proposes the following five generalizations to account for cross-linguistic patterns of labial harmony.

(6) Kaun's generalizations (2004:92)

Conditions favoring harmony:	Relevant constraint
a. The trigger is [-hi].	ALIGN-L/R([RD]/[-HI])
b. The trigger is [-bk].	ALIGN-L/R([RD]/[-BK])
c. The target is [+hi].	*[+RD,-HI]
d. Trigger and target agree in height.	GESTURALUNIFORMITY[RD]
e. The trigger is short.	ALIGN-L/R([RD]/[-LONG])

I will discuss Kaun's generalizations in order. Addressing the first two generalizations, (6a,b), Linker's (1982) articulatory study finds less lip compression and protrusion for front and non-high vowels. In other words, front and non-high vowels are less rounded articulatorily than back and high vowels. Kaun also references two perception-related findings from Terbeek (1977). First, Terbeek reports that high vowels were perceived as more rounded than non-high vowels. Second, back vowels were perceived as more rounded than front vowels. Kaun thus argues that where the articulatory lip rounding gesture is smaller in magnitude, its acoustic consequences are more subtle, leading to greater perceptual difficulty. In sum, because non-high and front vowels are produced with less lip rounding, they are perceptually less salient than high and back vowels, making them better triggers of harmony.

As for Kaun's third generalization (6c), she contends that, in contrast to trigger asymmetries, target asymmetries derive from perceptual salience. Since high rounded vowels involve more lip rounding, and high vowels are perceived as more rounded, Kaun argues that high vowels better cue the roundness of the triggering vowel.

Fourth, Kaun motivates the dispreference for cross-height labial harmony (6d) in the phonetic implementation of the feature, [round] (see also Steriade 1981:5). She argues that labial harmony involves the multiple linkage of a single [round] feature, and because lip gestures differ according to vowel height, the phonetic modulation of the lips to accommodate these differences militates against cross-height harmony (1995:§5.6). The importance of this generalization is corroborated in McCollum (2017), which demonstrates in a statistical analysis of 61 languages with labial harmony that same-height harmony is, by far, the most important of Kaun's typological generalizations. The other generalizations in (6) are more genetically and areally restricted, but the preference for same-height harmony spans the typology.[4]

In cases like Turkish labial harmony, there is no apparent preference for same-height harmony, since round vowels of both heights trigger harmony on high vowels suffixes. However, in a study on harmony in loanwords, Kaun (1999:97) finds that high vowels are better triggers for regressive labial harmony than non-high vowels in Turkish (see also Yavaş 1980; Clements & Sezer 1982:247). Kaun finds that, for all speakers, high vowels obligatorily triggered harmony, but non-high vowels obligatorily triggered harmony for only two of nine speakers. Kaun construes this as evidence for a same-height harmony preference in Turkish not manifest in the native lexicon.

Lastly, the length asymmetry noted in (6e) accounts for several length-based asymmetries reported in Li (1996) and Harrison (2000). In general, Kaun's generalization may relate to phonetic studies by Bennett (1968), Ainsworth (1972) and Hillenbrand et al. (2000), which suggest that shorter vowels are more difficult to identify, particularly in crowded portions of the vowel space.

In sum, perceptual weakness is argued to undergird trigger-related asymmetries, while perceptual salience motivates target asymmetries. In the next section, I analyze gradient labial harmony in Kazakh using the feature-based constraints in (6), which I then compare with an alternative, dispersion-based analysis in §5.

4 Feature-based analysis

4.1 Data Review

The data presented in (3-5) is briefly summarized in Table 3. Within roots, front and short vowels typically trigger harmony on short vowels while the long back vowel does not. Outside of roots, the short front vowel, /ʏ/, optionally triggers harmony on short suffixes while the long and back vowels do not (cf. 4e-h). The data from the converbial suffix in (5), though, show a more fixed hierarchy of trigger strength, ʏ > ɔ > ʏː > ɔː, or perhaps ʏ > {ɔ,ʏː} > ɔː. Lastly, when the target is long, harmony rarely obtains, regardless of morphological context.

[4] Only eastern Mongolian dialects allow same-height harmony where trigger and target disagree in backness.

Table 3: Synopsis of Kazakh data

Context	Target	Hierarchy	Examples (frequencies)
Root-Internal	[-long] target	ʏ, ɔ, ʏː > ɔː	qɔɫɔn (16/17) vs. qɔːzɔ (1/10)
	[+long] target		tʏːbʏː (0/11) vs. bɔːɫɔːt (0/10)
Suffix with underlying V	[-long] target	ʏ > ɔ, ʏː, ɔː	kʏldʏ (7/22) vs. ʏːldʏ (1/20)
	[+long] target		kʏl-sɪː (0/15) vs. qɔrsɑː (0/13)
Converbial suffix (CVB)	epenthetic target	ʏ > ɔ > ʏː > ɔː	kʏlʏp (26/31) vs. qɔːsɔp (1/15)

4.2 Formal feature-based analysis

4.2.1 Constraint set

The data from §2.2.2 are intuitively amenable to Kaun's analysis, since both short and front vowels are better triggers for harmony. In Kaun (2004), these preferences are encoded via conditional alignment constraints. A general alignment constraint (McCarthy & Prince 1993; Kirchner 1993) is defined in (7).

(7) ALIGN-R([RD],WD) assign a violation to every vowel that is not associated with the [rd] feature of the left edge of the word

Kaun conditions alignment upon the presence of certain features in the trigger. I use the features [-bk] and [-long] to define the set of conditional alignment constraints necessary for the feature-based analysis (cf. Walker 2011:247-251). The constraint in (8) encodes the preference for front vowel triggers. The constraint in (9) encodes the preference for short triggers.

(8) ALIGN-R([RD]/[-BK],WD) if the features [+rd] and [-bk] co-occur at the left edge of the word, assign a violation to every vowel that is not associated with the [rd] feature of the left edge of the word

(9) ALIGN-R([RD]/[-LONG],WD) if the features [+rd] and [-long] co-occur at the left edge of the word, assign a violation to every vowel that is not associated with the [rd] feature of the left edge of the word

Three markedness constraints will also figure prominently in the analysis. First, CRISPEDGE[RD,MORPH] bans the extension of a labial gesture across a morpheme boundary (Itô & Mester 1999; Walker 2011).

(10) CRISPEDGE[RD,MORPH] assign a violation to every segment that is [+rd] that is associated with more than one morpheme.

Harmony processes are often sensitive to morphological boundaries. For instance, in Yeyi, a Bantu language of southern Africa, regressive assimilation of [i] to [u] consistently obtains within roots, but is variable across a root boundary (Seidel 2008:47-48). Further, categorical blocking by morphology is reported in a variety of languages (Archangeli & Pulleyblank 2007:365; Rose & Walker 2011:§4.2.2). Additionally, some evidence indicates that articulatory gestures are more tightly coordinated within morphemes than across morphological boundaries (Cho 2001).

The second markedness constraint used in the analysis is *[+RD,+LONG], which penalizes long round vowels. In effect, this constraint replaces Kaun's *[+RD,-HI] constraint in (6), since the historical mid vowels are now long vowels. A similar resistance to harmony is exhibited by long vowels in Maltese (Puech 1978), where short vowels undergo harmony but long vowels do not.

(11)	*[+RD,+LONG]	assign a violation to every segment bearing the features, [+rd] and [+long].

Third, a general markedness constraint against [+rd] vowels penalizes harmony on underlying as well as epenthetic vowels.

(12)	*[+RD]	assign a violation to every instance of the feature [+rd]

In addition to the above markedness constraints, one IDENT constraint (McCarthy & Prince 1995) is used in the analysis.

(13)	IDENT-IO[RD]	assign a violation to every input-output pair that disagree for the feature, [rd].

This constraint penalizes harmony on underlying vowels, like the past tense suffix, but not on epenthetic vowels, like the converbial suffix.[5] Epenthetic vowels often pattern differently from underlying vowels in harmony systems (Vaux 1998:169; Finley 2008a). This is evident in the related language, Uyghur, where labial harmony targets epenthetic vowels only (Hahn 1991:49-51).

4.2.2 Gradience and Maximum Entropy Harmonic Grammar

Gradience is problematic for canonical Optimality theory (Prince & Smolensky 1993/2004) since strict domination entails categorical outcomes. Early attempts to model gradience include Anttila (1997), which uses partial ordering of constraints, and Boersma & Hayes (2001), which gives each constraint a range of overlapping strictness bands. More recent work has used weighted constraints in Harmonic Grammar (HG; Legendre et al. 1990; Pater 2009) to derive probabilistic predictions (Goldwater & Johnson 2003; Coetzee & Pater 2011; Boersma & Pater 2016). The particular version of HG used throughout the analysis is Maximum Entropy

[5] It is also possible to differentiate these two because harmony spans one consonant in CVB but two in PST (Xiangru & Hahn 1989:272-273 for Ili Turki). Harmony does span multiple consonants within roots like [bʏrkʏt] 'golden eagle,' though.

Harmonic Grammar (MaxEnt HG; Goldwater & Johnson 2003). MaxEnt models are statistical models with well-understood mathematical properties that have been widely used in a variety of disciplines.

MaxEnt HG assigns each output candidate a harmony score, H. The harmony of each candidate is calculated using the following equation (Hayes & Wilson 2008:383):

(14)

$$H(x) = \sum_{i=1}^{N} w_i \cdot C_i(x)$$

where

w_i is the weight of the i^{th} constraint,

$C_i(x)$ is the number of times that x violates the i^{th} constraint, and

$\sum_{i=1}^{N}$ denotes summation over all constraints $(C_1 \ldots C_N)$

Thus, the Harmony score for each output candidate is the summed total of constraint violations multiplied by their respective weights. All constraint violations are assumed to be negative.

From the Harmony score, $H(x)$, the probability of output candidate x, P_x, is calculated by taking the exponent of the harmony score normalized by the summed score of the exponents of all competing output candidates, y, $(Y_1 \ldots Y_M)$, in (15).

(15)

$$P_x = e^{H(x)} \Big/ \sum_{y=1}^{M} e^{H(y)}$$

Harmony scores with more negative values produce smaller numerators, which result in lower probabilities. Thus, the ideal candidate would have a harmony score of zero, incurring no violations of faithfulness or markedness constraints. The relation between $H(x)$ and P_x is contingent upon the set of output competitors under examination. Therefore, a harmony score of, for instance, -5, may represent a high probability outcome if other outputs have significantly lower harmony scores. On the other hand, a score of -5 may indicate an extremely low probability candidate if other outputs have harmony scores greater than -5.

To exemplify how MaxEnt models compute probabilities, consider the illustrative tableau below. In (16), Candidate (a) violates ALIGN-R twice, because two non-initial vowels do not undergo harmony while Candidate (b) violates ALIGN-R only once. These violations are multiplied by the weight associated with ALIGN-R, 1.5, which yields -3 and -1.5, respectively. As for *[+RD], Candidate (a) incurs one violation while Candidate (b) incurs two. Since *[+RD] has been assigned a weight of 1, the Candidates incur penalties of -1 and -2. The sum total of all penalties incurred from these two constraints result in each candidate's Harmony score.

The probability of each candidate is derived from the equation in (15). The probability of Candidate (a) would be: $e^{-4}/(e^{-4}+e^{-3.5})$, which is .38. Similarly, the probability of Candidate (b) would be: $e^{-3.5}/(e^{-4}+e^{-3.5})$, which is .62. Thus, (16) predicts that Candidate (b) will surface 62% of the time.

(16)

	/ɔ...ə...ə/	ALIGN-R	*[+RD]	Harmony	Predicted probability
	Weights	1.5	1		
Candidate a.	ɔ...ə...ə	-2 * 1.5 = -3	-1 * 1= -1	-3 + -1 = -4	0.38
Candidate b.	ɔ...ɔ…ə	-1 * 1.5= -1.5	-2 * 1 = -2	-1.5 + -2 =-3.5	0.62

4.3 Constraint weights

The seven constraints presented in (7-13) were used to assess input-output mappings. Each constraint was assigned a weight in order to maximize log-likelihood of the model. The log-likelihood is the summed product of log probabilities, $\log(P_x)$, and the attested frequencies of the candidates under examination, which results in a negative number with a large absolute value.[6] As the search space for MaxEnt models is free of local maxima (Della Pietra et al. 1997), a variety of optimization algorithms converge on one solution. Accordingly, I used Excel's Solver add-in (Fylstra et al. 1998)[7] to find the maximum (least negative) log likelihood using the Generalized Reduced-Gradient algorithm.[8]

The optimal weights assigned to the training data are shown in (17).

(17) Constraint weights

CRISPEDGE(RD,MORPH)	3.35
ALIGN-R([RD],WD)	3.14
IDENT-IO[RD]	3.13
*[+RD,+LONG]	2.19
ALIGN-R([RD]/[-BK],WD)	1.82
ALIGN-R([RD]/ [-LONG],WD)	1.42
*[+RD]	1.16

Given the weights in (17), two tableaux exemplify the analysis below. In (18), Kaun's model fits the [kʏlʏp]~[kʏlɪp] data remarkably well, diverging from the actual probability by only .027.

[6] For a description of the math, see the supplemental materials for McPherson & Hayes (2016), at: http://www.dartmouth.edu/~mcpherson/papers-and-handouts/Harmony_supplements.html

[7] In Wilson & George's (2009) MaxEnt Grammar Tool, constraint violations can only be integers, limiting the precision of the calculations in §5, which rely on non-integer values.

[8] Solver does not impose a penalty on large weights (cf. Goldwater & Johnson 2003).

(18)

/kʏl-p/	CRISP EDGE	ALIGN-R	ID-IO-[RD]	*[+RD, +LONG]	ALIGN-R/[-BK]	ALIGN-R/[-LONG]	*[+RD]	Har-mony	Predicted probability	Observed frequency
Weights	3.35	3.14	3.13	2.19	1.82	1.42	1.16			
kʏl-ʏp	-3.35						1.16 *-2 = -2.32	-3.35 + -2.32 = -5.67	0.866	0.839
kʏl-ɪp		-3.14			-1.82	-1.42	-1.16	-3.14 + -1.82 + -1.42 + -1.16 = -7.54	0.133	0.161

In (19), output forms, [qɔłɔn] and [qɔłən] from input /qɔłən/ 'colt' are considered. In contrast to (18), the fit from Kaun's analysis differs substantially from the observed frequency. Harmony obtains 94% of the time in this word, but Kaun's analysis predicts that harmony should obtain only 57% of the time.

(19)

/qɔłən/	CRISP EDGE	ALIGN-R	ID-IO-[RD]	*[RD, +LONG]	ALIGN-R/[-BK]	ALIGN-R/[-LONG]	*[+RD]	Harmony	Predicted probability	Observed frequency
Weights	3.35	3.14	3.13	2.19	1.82	1.42	1.16			
qɔłɔn			-3.13				-2 * 1.16= -2.32	-3.13 + -2.32 = -5.45	0.567	0.941
qɔłən		-3.14				-1.42	-1.16	-3.14 + -1.42 + -1.16 = -5.72	0.433	0.059

One significant idealization was made- only two output candidates were evaluated, following Hayes & McPherson (2016). In a fuller analysis, additional (highly-weighted) constraints against unrounding root vowels and complex codas would penalize the potential but unattested forms, [kɪl-ɪp] and [kʏl-p].

Figure 4 shows the error for Kaun's analysis, with under/overpredicted probabilities represented by horizontal bars. The length of each bar shows the error on each word.

Figure 4: By-word error of the feature-based analysis

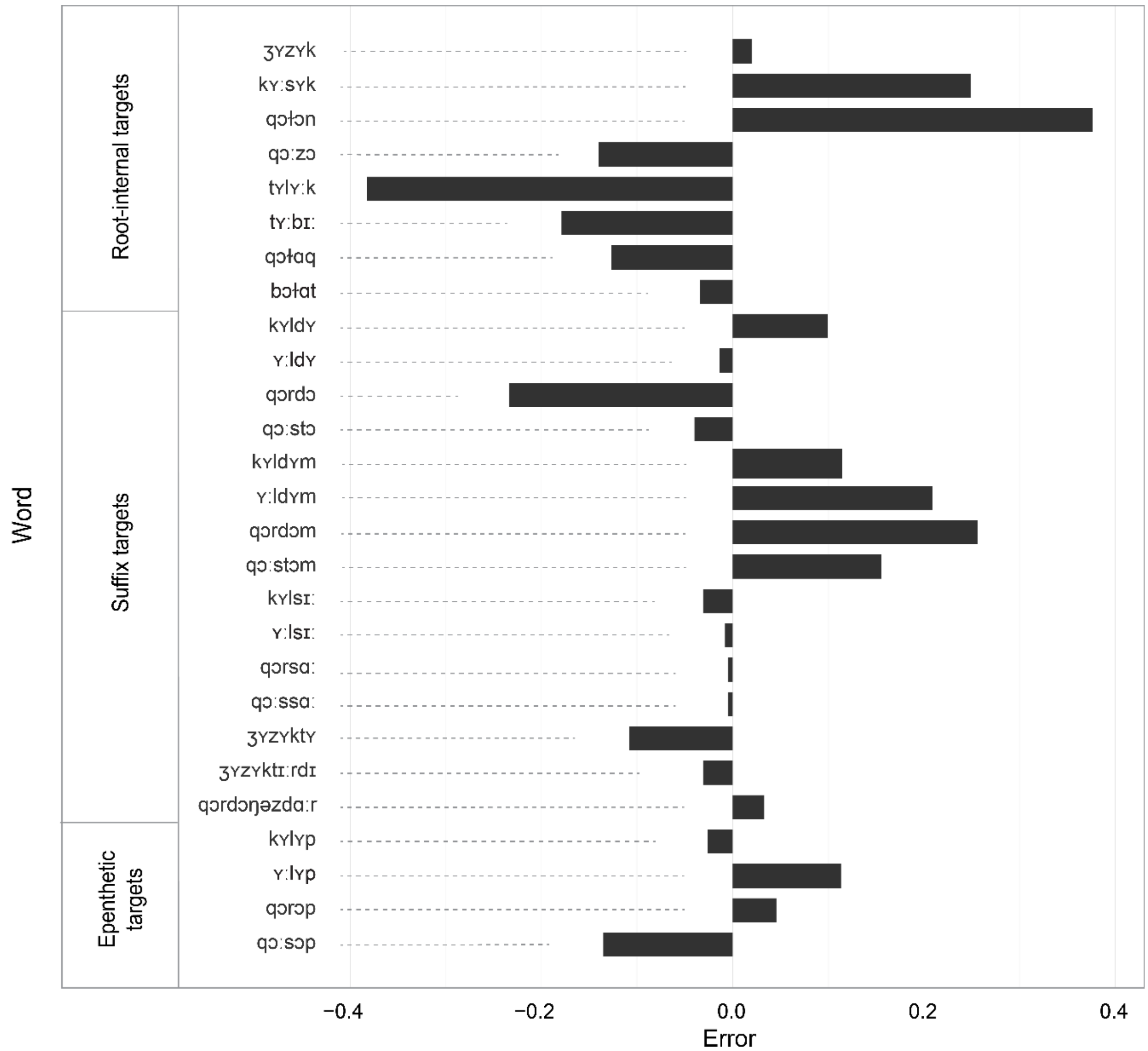


Table 4: Summary of the feature-based analysis

Analysis (*k*=number of parameters)	Log-likelihood (null = -335.48)	Pseudo-r^2
Feature-based (*k*=7)	-177.57	.64

Table 4 summarizes the global fit of the feature-based model. A null model was created to assess pseudo-r^2 (Nagelkerke 1991), an indicator of variance explained by a logistic model (e.g. MaxEnt). While the feature-based analysis accounts for a good portion of the data, it does not tightly fit a number of items, which is evident in Figure 4.[9] In the next section I present a dispersion-based analysis that more successfully accounts for the Kazakh data than Kaun's analysis.

5 Dispersion-based analysis

This section argues that dispersion provides a better operationalization of weakness than features, using dispersion-based perceptual weakness to motivate an alternative analysis. In §5.1, I introduce Dispersion Theory (Liljencrants & Lindblom 1972; Lindblom 1986). In §5.2, I compute perceptual distances for the Kazakh inventory. These distances then inform the analysis presented in §5.3.

5.1 Dispersion Theory

Dispersion theory (henceforth DT; Liljencrants & Lindblom 1972; Lindblom 1986) attempts to model the shape and typological frequency of inventories rather than individual segments. DT models the vowel space without explicit reference to features, arguing instead that vowels should be maximally dispersed. DT uses the formant values of each phonemic vowel to assess the fitness of an inventory. Following Liljencrants & Lindblom (1972), most work in DT has sought to explain the particular shapes that vowel inventories often take (Lindblom 1975, 1986; Schwartz et al. 1997; Flemming 2002). For example, the frequency of the three-vowel system containing /i a u/ is explained by the large distances between the three vowels in the system. In Flemming (2002), this dispersionary motivation is formalized with a set of Optimality theoretic constraints favoring large distances between contrasts in opposition to constraints maximizing the number of contrasts. The interaction of these conflicting constraints define the shape of the vowel space for a given language. Whereas previous work used dispersion to predict inventories, Flemming's work (2002, 2004, 2008a,b) assumes that speakers synchronically access prototype-like representations. Figure 5 shows Flemming's (2002; cf. Flemming 2004, 2006) representational system. Perceptual distances are calculated in spatial rather than featural terms by counting grid spaces between vowels.

Figure 5: Flemming's vowel space (2002:30)

F2

5	4	3	2	1	
i	y	ɨ	ɯ	u	1
ɪ	ʏ			ʊ	2
e	ø	ə	ɤ	o	3
	ɛ	œ	ʌ	ɔ	4
		a	ɑ		5

F1

Within Optimality Theory, DT has also been used to account for phenomena beyond inventory structure. Padgett (2004), uses an ERB (Equivalent Rectangular Bandwidth) scale to define the Russian vowel space, with SPACE constraints militating against category prototypes insufficiently

[9] Kaun's proposal was set in OT with strict dominance, as opposed to HG, so perhaps her analysis, specifically her constraint set, would've differed if it had been cast in HG.

dispersed in the acoustic space. Flemming (2006, 2008a) address positional markedness, enhancement and surface phonotactics within a dispersion-based framework. This analysis extends the influence of DT into another realm of phonology, vowel harmony.

5.2 Perceptual distance

In this section, I apply the methodologies used in previous DT work to the Kazakh vowel inventory to compute perceptual distances for pairs that alternate in harmony. Specifically, the perceptual distances calculated below are used to operationalize perceptual weakness, where weaker (i.e. more similar) contrasts correspond to smaller weighted Euclidean distances between harmonic counterparts. By-vowel mean F1-F2 with standard deviations are reported in Table 5.[10]

Table 5: Mean root vowel F1-F2 (Hz) with one standard deviation

Vowel	F1 (SD)	F2 (SD)
ɑː	749 (117)	1333 (198)
ɔː	493 (61)	1015 (167)
ə	540 (124)	1395 (166)
ɔ	507 (107)	1094 (171)
ɪː	402 (54)	2286 (294)
ʏː	407 (55)	1467 (306)
ɪ	419 (87)	1915 (231)
ʏ	387 (72)	1664 (271)

First, raw formant values were transformed into ERB using the equation provided in Reetz & Jongman (2008:245-246).[11] After conversion to ERB, data was normalized (Lobanov 1971) to reduce inter-speaker physiological differences, allowing for between-speaker comparisons. Results from this transformation are shown in Table 6.

After normalization, the vowel space was warped using the parameter, λ, from Schwartz et al. (1997) such that the ratio of F2-to-F1 ranges was 0.75-0.5:1 (see also de Boer 2000; Padgett 2004; Padgett & Tabain 2005). As in Lindblom (1975), this weights F1 over higher formant contrasts, since differences in F1 are more salient than differences in F2 (Flanagan 1955). This warping of the vowel space reflects this asymmetry in perception. To achieve the 0.625:1 midpoint in this range, λ was set to 0.723.[12] Note that Flemming (2002:29-32) uses a 1:1 vowel space in tandem with constraints favoring F1 contrasts over F2. Although the dimensions are different in Flemming's analysis, the constraints ranking F1 over F2 differences accords with the results elsewhere in DT.

[10] F3 is not incorporated into the model for the following reasons: first, F3 did not significantly improve vowel discrimination in a discriminant analysis; second, F3 was highly variable within-speakers, and third, the model performed poorer using a perceptual second formant affected by F3 (Schwartz et al. 1997:263; de Boer 2000:448).

[11] Without actual perceptual data these findings must suffice, but research on vowel perception is needed for a fuller understanding of labial harmony.

[12] In principle, the target ratio could have fallen anywhere in the proposed range. The distances, if computed without adjusting F2, would be- ɪ-ʏ=0.447, ə-ɔ=0.777, ɪː-ʏː=1.138, ɑː-ɔː=1.521.

In Table 6, the perceptual distance between the harmonic counterparts, /ɪ/ and /ʏ/, is 0.365, much smaller than all other distances. Moreover, the relative distances between all four harmonic pairs correspond to the trigger strength hierarchy described in §2.2.2, ʏ > ɔ > ʏː > ɔː.

Table 6: root vowel F1-F2 in normalized ERB with standard deviations

Vowel	F1 (SD)	F2 (SD)	Weighted Euclidean Distance (λ(F2)=0.723)
ɪ	-0.542 (0.532)	0.745 (0.227)	0.365
ʏ	-0.788 (0.459)	0.372 (0.388)	
ə	0.117 (0.657)	0.103 (0.272)	0.568
ɔ	-0.007 (0.587)	-0.664 (0.353)	
ɪː	-0.637 (0.375)	1.196 (0.265)	0.823
ʏː	-0.595 (0.399)	0.059 (0.482)	
ɑː	1.339 (0.452)	-0.169 (0.298)	1.451
ɔː	-0.032 (0.354)	-0.828 (0.340)	

The trigger strength hierarchy, after conversion to ERB, normalization, and warping differs from the hierarchy obtained from raw formant values. If perceptual distances are calculated based on raw formant values, without any transformations, ɪ-ʏ= 253.2, ə-ɔ = 302.7, ɑː-ɔː= 408.5, ɪː-ʏː= 819.1. This difference in the ranking of /ɔː/ and /ʏː/ derives from the independently motivated transformations discussed above. These transformations highlight the difference between the acoustic and perceptual properties of vowels, and this difference underlies the change in relative rankings for these two vowels.

In Figure 6, the vowel space is depicted along with the weighted Euclidean distances between each harmonic pair. Recall from above that /ʏ/ is the best trigger and /ɔː/ is the worst trigger of harmony, with /ɔ/ and /ʏː/ somewhere in between. This empirical generalization lines up with the distances shown in Figure 6 and their correlation with frequency of harmony, shown in Figure 7.

Figure 6: Perceptual distances between labial harmonic counterparts

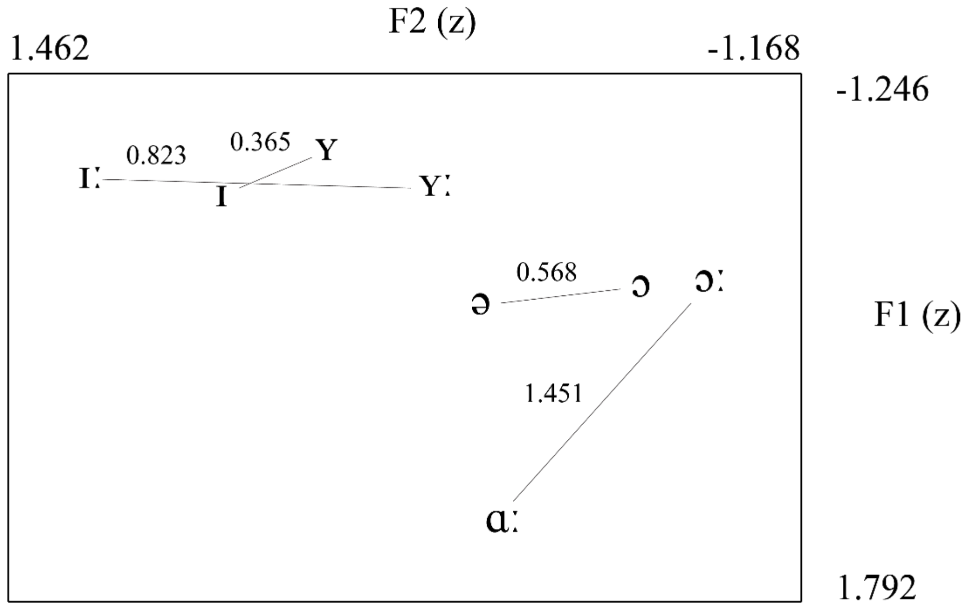
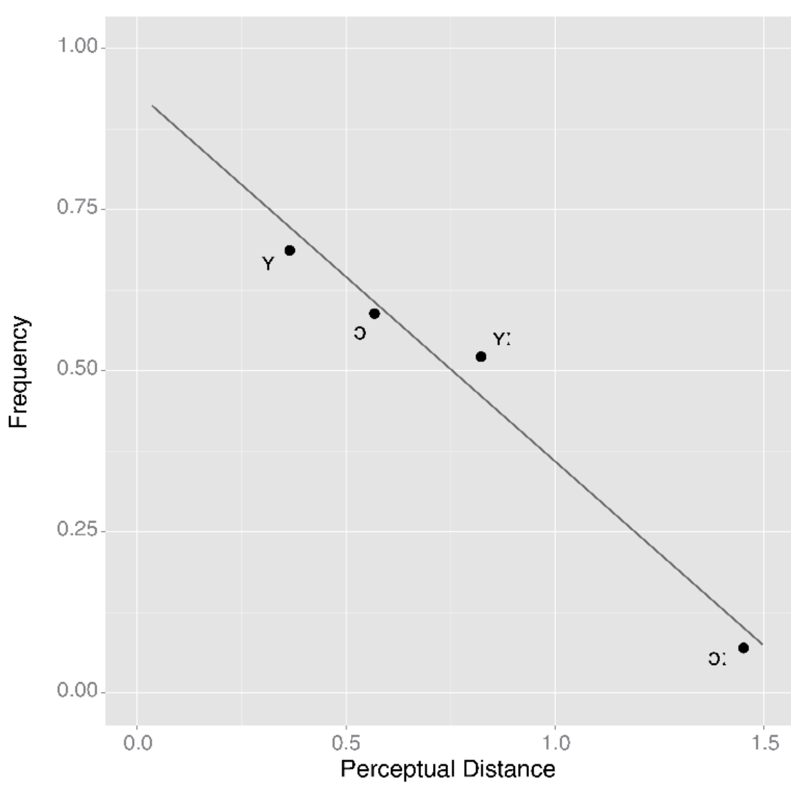


Figure 7: Trigger perceptual distance and rate of harmony on short vowels averaged across morphological contexts (ʏ=0.365z, ɔ=0.568z, ʏː=0.823z, and ɔː=1.451z)



In §5.3.1 below, I use the distances derived in this section to motivate a scalar alignment constraints for the dispersion-based analysis presented in §5.3.2.

5.3 Formal dispersion-based analysis

5.3.1 Scalar alignment

The perceptual distances just calculated create one simple metric for comparison. Perceptual distance is formally implemented with one scalar constraint, which replaces the three alignment constraints in the feature-based analysis. Scalar constraints necessitate only a slight modification to HG (e.g. Hsu & Jesney 2015). In HG, all constraints are assigned weights. A scalar constraint multiplies the penalty incurred by that constraint in relation to some metric, like sonority, sequential distance, or in this case, perceptual distance. This introduces a new mechanism, *S*, into the equation from (14) used to determine Harmony scores.

(20)

$$H(x) = \sum_{i=1}^{N} w_i \cdot C_i(x) \cdot S_i(x)$$

where

w_i is the weight of the i^{th} constraint,

$C_i(x)$ is the number of times that x violates the i^{th} constraint,

$S_i(x)$ is a scale applied to $C_i(x)$, and

$\sum_{i=1}^{N}$ denotes summation over all constraints ($C_1 \dots C_N$)

Thus, for the relevant constraint, ALIGN-R([RD]), constraint violations are multiplied by the general weight of ALIGN-R([RD]), and by the scaled weight assigned to each trigger vowel, using the perceptual distances calculated in §5.2.

(21) Definition: SCALAR ALIGN-R(RD,WD)
given the constraint, ALIGN-R(RD,WD), the constraint weight, $w_{\text{ALIGN-R}}$, Δ*z*, the perceptual distance between [rd] vowel, *z*, and its harmonic counterpart, and a scale, *S*:

For every vowel to which the [rd] feature of the initial vowel with perceptual distance Δ*z* is not aligned, assign a weighted violation, $w_{\text{ALIGN-R}} \cdot S(\Delta z)$, where *S(Δz)* is the scaled value derived from Δ*z*.

To implement this constraint, it was first necessary to determine the optimal scale. Through trial and error, the optimal scale was 2.3:1. As shown in Table 7, the weight of ALIGN-R([RD]) after /ʏ/, 1, was 2.3 times greater than after /ɔː/, 0.435. Using the scaling factor below, the values weights associated with the two intermediate values, /ɔ/ =0.894 and /ʏː/ = 0.762 were linearly interpolated between the two endpoints, /ʏ/ and /ɔː/. Thus, the only free parameter below is the slope of the scale, 2.3, with 1.19 and -0.52 being derived entirely from the scale's slope and the perceptual distances calculated above.

Table 7: Scaled weights for each vowel

Vowel (V)	Perceptual distance (Δ)	Scaling factor $= 1 - \frac{1-\frac{1}{2.3}}{\Delta(\text{ʏ}) - \Delta(\text{ɔː})} \cdot (\Delta(\text{ʏ}) - \Delta(V))$ $=1.19 - 0.52 \cdot \Delta(V)$
ʏ	0.365	1.000
ɔ	0.568	0.894
ʏː	0.823	0.762
ɔː	1.451	0.435

One conceptual idealization made throughout the analysis is that harmonic pairings are assumed. Since featural specifications are not used to derive harmonic counterparts, some tool, either representational or grammatical, must exist to maintain the correct pairings. Throughout the paper I abstract away from this issue, but it deserves a fuller analysis.

Returning to the example from (16), given a scalar alignment constraint, as defined in (21), then the probability of Candidate (b), with partial harmony, in (22) is 0.58 because the weight of Align-R is scaled by 0.894, the value derived in Table 7 for /ɔ/.

(22)

	/ɔ...ə...ə/	ALIGN-R	*[+RD]	Harmony	Predicted probability
	Weights	1.5	1		
		SCALE: ʏ = 1, ɔ =.894, ʏː =.762, ɔː =.435			
Candidate a.	ɔ...ə...ə	-2 * 1.5 * .894 = -2.682	-1 * 1 = -1	-2.682 + -1 = -3.682	0.42
Candidate b.	ɔ...ɔ…ə	-1 * 1.5 * .894 = -1.341	-2 * 1 = -2	-1.341 + -2 = -3.341	0.58

5.3.2 Formal analysis

Using the perceptual distances and their implementation in a scalar alignment constraint from the previous subsection, this subsection presents the dispersion-based analysis, showing that the dispersion-based account is superior to the feature-based analysis in §4.

Recall that the feature-based analysis used seven constraints: ALIGN-R[RD/-BK], ALIGN-R[RD/-LONG], ALIGN-R[RD], CRISPEDGE(RD, MORPH), *[+RD], *[+RD,+LONG] and IDENT-IO[RD]. The dispersion-based analysis presented below uses the same set of constraints, except the three conditional alignment constraints are replaced by the scalar alignment constraint introduced in the previous section. Thus, each analysis differs only in its harmony-driving constraint(s). Since all other factors were held equal, we can readily compare the different results derived from Kaun's three feature-based ALIGN constraints with my single dispersion-based scalar ALIGN constraint.

The four constraints, SCALAR ALIGN-R([RD]), CRISPEDGE(RD, MORPH), *RD and IDENT-IO[RD] were optimized to fit the observed frequencies in the data. The weights assigned to each constraint are shown in (23), and exemplified in (24-25).

(23) Constraint weights

SCALAR ALIGN-R(RD,WD)	5.84
*[+RD,+LONG]	5.67
CRISPEDGE(RD,MORPH)	3.50
IDENT-IO[RD]	2.40
*[+RD]	1.58

In (24), the dispersion-based model predicts the frequency of harmony from the input, /kʏl-p/ with relative accuracy. The summed number of violations are multiplied by the relevant weights, equaling -5.14 and -6.66, respectively. The exponents of these two harmony scores result in the probabilities- $P_{kʏlʏp}$= 0.821 and $P_{kʏlɪp}$= 0.179, which correspond fairly well with attested frequency. For this particular example, the dispersion-based analysis is slightly more accurate than the feature-based analysis, with errors of .018 and .024, respectively.

(24)

/kʏl-p/	ALIGN-R	*[+RD,+LONG]	CRISP EDGE	ID-IO[RD]	*[+RD]	Harmony	Predicted probability	Observed frequency
	SCALE: ʏ = 1, ɔ = .894, ʏː = .762, ɔː =							
Weights	5.84	5.67	3.50	2.40	0.82			
kʏl-ʏp			-3.50		0.82 * -2 = -1.64	-3.50 + -1.64 = -5.14	0.821	0.839
kʏl-ɪp	5.84 * 1 * -1 =-5.84				-0.82	-5.84 + -0.82 = -6.66	0.179	0.161

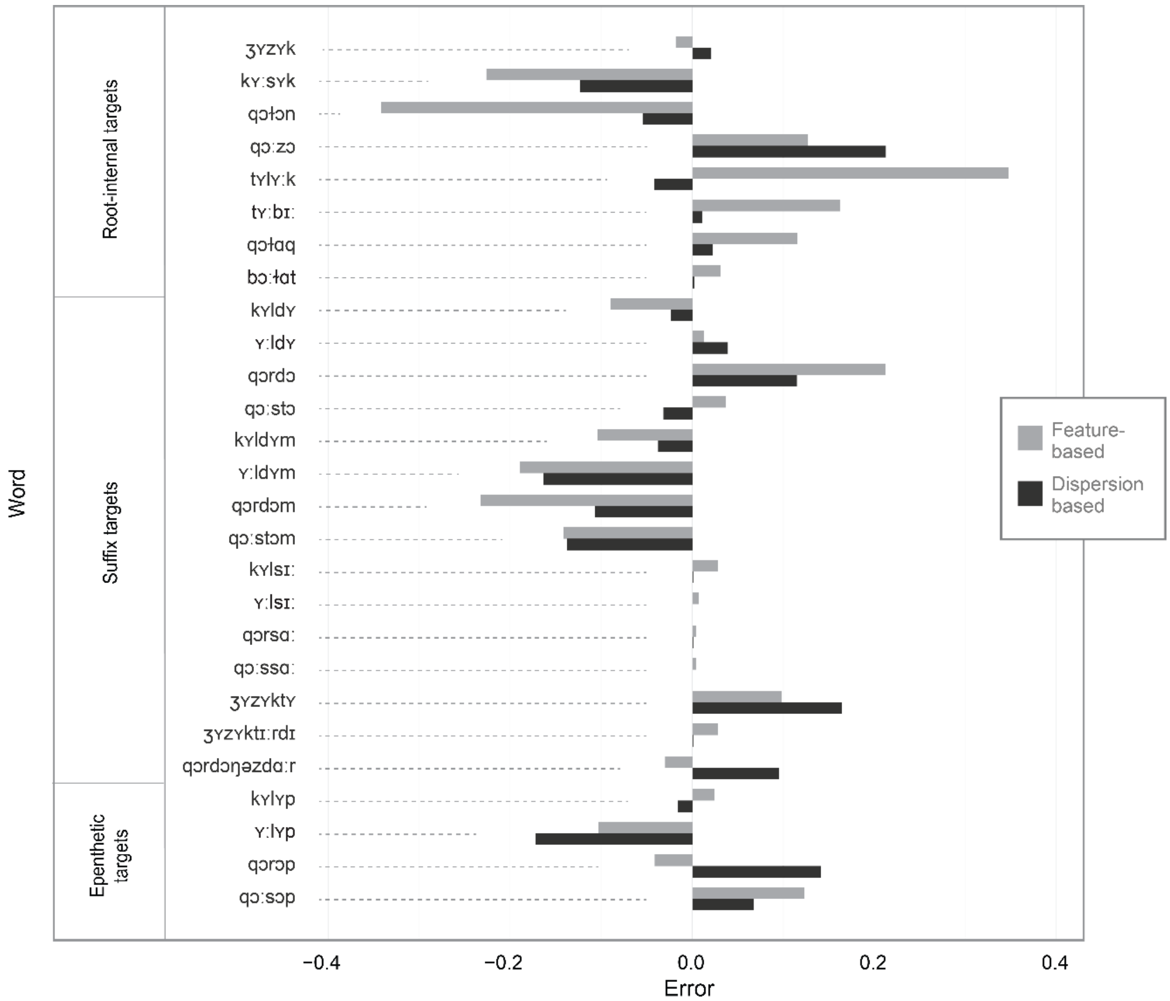
Recall that the feature-based analysis failed to accurately predict the frequency of harmony from the input, /qɔɬən/. In (25), the harmonic candidate, [qɔɬɔn], violates ID-IO[RD] once and *[+RD] twice. The disharmonic candidate, [qɔɬən], violates SCALAR ALIGN-R and *[+RD] once each. The two candidates' resulting Harmony scores are -4.04 and -6.04. From these values, the model predicts harmony 88% of the time. This is far closer to the actual frequency of harmony than the feature-based prediction in (19), 57%.

(25)

/qɔłən/	ALIGN-R	*[+RD, +LONG]	CRISP EDGE	ID-IO[RD]	*[+RD]	Harmony	Predicted probability	Observed frequency
Weights	SCALE: ʏ = 1, ɔ = .894, ʏː = .762, ɔː = .435							
	5.84	5.67	3.50	2.40	0.82			
qɔłɔn				-2.40	0.82 * -2 = -1.64	-2.40 + -1.64 = -4.04	0.881	0.941
qɔłən	5.84 * 0.894 * -1= -5.22				-0.82	-5.22 + -0.82 = -6.04	0.119	0.059

Figure 8 compares under/overpredicted probabilities of the data in (3-5) for each analysis. The overall length (error) of the black bars is smaller, which indicates that the dispersion-based model fits the data better than the feature-based model. Of the 27 items compared below, the dispersion-based analysis is more accurate on 20, while Kaun's analysis is more accurate on 7.

Figure 9: By-word error of both analyses



In Table 8 below, the differences in model fit shown in Figure 8 are compared using log-likelihood, AIC (Akaike Information Criterion) and pseudo-r^2. Significantly, AIC is a log-likelihood-derived measure of model fit that is affected by model complexity. The dispersion-based model is simpler, using five constraints, including one alignment constraint, plus a scale, *k*=6, whereas the feature-based analysis uses seven constraints, including three alignment constraints, *k*=7. This increase in complexity is penalized by AIC. Burnham & Anderson (2004:270-271) suggest that in model comparison, a difference in AIC (ΔAIC) of less than 2 is insignificant, models differing by 4 or more is significant, and a difference of more than 10 is highly significant.

On every metric in Table 8, the dispersion-based analysis outperforms the feature-based analysis. In particular, the ΔAIC of the two analyses is over 50, which is evidence for the superiority of the dispersion-based analysis. The dispersionary analysis is simpler and still accounts for more data than Kaun's analysis. The relative complexity of the feature-based model results in an AIC increase of 2. The remaining ΔAIC comes from the difference in model accuracy To illustrate this, if the number of model parameters were held constant, the log-likelihood necessary for the featural model to fare better than the dispersionary analysis is -150.59, a difference of 26.98 from the actual model. Viewed alternatively, if log-likelihoods were held constant, any dispersion-based analysis with fewer than 33 parameters would result in a lower AIC than the feature-based analysis.

Table 8: Model comparisons

Analysis (*k*=number of model parameters)	Log-likelihood (LL, null = -335.48)	AIC (null = 670.97)	Pseudo-r^2
Dispersion-based (*k*=6)	-151.59 (ΔLL=25.98)	315.19 (ΔAIC=53.94)	.71
Feature-based (*k*=7)	-177.57	369.13	.64

I also used cross-validation to compare the two models and check for overfitting the data. A model may learn to describe the data well but may perform poorly when asked to predict outputs from unseen data. It is thus common to separate the dataset into training and test sets to assess the model's ability to predict probabilities of novel data. I randomly assigned 25% of the dataset to test, and the remaining 75% was used for training. Four runs were conducted, with each data point occurring in exactly one test set. Results from each trial are shown in Table 9. See the appendix for a full list of training and test items.

This additional method for model comparison corroborated the results in Table 8. In all but one run, Test 2, the dispersion-based analysis outperformed Kaun's analysis. Summed log-likelihoods in Table 9 show a similar difference to that in Table 8, suggesting the dispersion-based analysis is a superior predictive model.

Table 9: Cross-validation results for both analyses

<table>
<tr><th rowspan="2">Analysis</th><th colspan="7">Log-likelihood (LL)</th></tr>
<tr><th></th><th>1</th><th>2</th><th>3</th><th>4</th><th>Sum of Test</th><th>Mean of Test</th></tr>
<tr><td rowspan="2">Dispersion-based (k=6)</td><td>Train</td><td>-110.61</td><td>-103.22</td><td>-120.31</td><td>-118.47</td><td rowspan="2">-167.36 (ΔLL=27.94)</td><td rowspan="2">-41.84</td></tr>
<tr><td>Test</td><td>-42.00</td><td>-59.76</td><td>-31.98</td><td>-33.62</td></tr>
<tr><td rowspan="2">Feature-based (k=7)</td><td>Train</td><td>-132.00</td><td>-122.86</td><td>134.28</td><td>-137.03</td><td rowspan="2">-195.30</td><td rowspan="2">-48.83</td></tr>
<tr><td>Test</td><td>-49.99</td><td>-57.18</td><td>-46.51</td><td>-41.62</td></tr>
</table>

Why did the dispersionary analysis perform poorly on Test 2? Actually, both analyses performed poorly on Test 2. The log-likelihood of both models on the training set was far better than other training sets. When the tightness of training fit is juxtaposed to poorer performance on test data, it seems clear both analyses overfit the training data here. The weight of *[+RD,+LONG] in Test 2 for the dispersionary model was 16.68, far higher than any other constraint in any of the other trials. I reduced the weight of *[+RD,+LONG] to 5 after Solver ran, resulting in a log-likelihood of -104.48 on training, only a slight decrease in performance from -103.22. I then substituted 5 for 16.68 on test data, producing a notable increase in performance, from -59.76 to -49.2.

Importantly, these two analyses derived output mappings from inputs with non-initial vowels specified as [-rd] (cf. Steriade 1995). The Richness of the Base hypothesis, however, argues that the set of possible inputs is not constrained by the grammar, which is subverted if non-initial inputs are obligatorily [-rd] (Prince & Smolensky 1993/2004). To test the effect of non-initial [+rd] inputs, I re-ran each model once with all non-initial inputs being specified as [+rd] and twice with [+rd] being randomly assigned to 50% of non-initial inputs. Throughout, I maintained the same partition of training and test items. In all trials the dispersion-based model outperformed the feature-based model on test data, even though on two runs the dispersion-based model performed poorer on training data. Thus, I conclude that the present analysis doesn't unduly benefit from the assumption that non-initial vowels are underlyingly specified as [-rd], although this contravenes Richness of the Base.

As one reviewer notes, the dispersion-based analysis performs better because it captures a more salient generalization than the feature-based analysis. Within the feature-based analysis, the probability of harmony depends on the additive structure of featural representations. The probability of harmony after each vowel is dependent on the weight, *w*, of ALIGN-R([RD]), as well as the weights of the other two harmony-driving constraints, ALIGN-R([RD]/[-BK]) and ALIGN-R([RD]/[-LONG]), in (26). The vowel, /ʏ/, is both front and short, so all three alignment constraints motivate harmony. In contrast, harmony after the long back vowel, /ɔː/, is only motivated by ALIGN-R([RD]), so harmony is less frequent.

(26) By-vowel influence of featural alignment constraints

ʏ $w_{\text{ALIGN}} + w_{\text{ALIGN(FRONT)}} + \mathrm{w}_{\text{ALIGN(SHORT)}}$

ɔ $w_{\text{ALIGN}} + \mathrm{w}_{\text{ALIGN(SHORT)}}$

ʏː $w_{\text{ALIGN}} + w_{\text{ALIGN(FRONT)}}$

ɔː w_{ALIGN}

Crucially, because the featural analysis encodes a length distinction, the difference between /ʏ/ and /ɔ/ is assumed to equal to the difference between their long counterparts, /ʏː/ and /ɔː/. However, the difference in the rate of harmony, as seen in Figure 7, is far greater for the long vowels. In contrast, these differences in rate of harmony are highly correlated the perceptual distances, which is central to the dispersion-based analysis.

Briefly, both analyses do not account for all the variance encountered in the data. Other factors, including age, gender and education were all found to play a contributing role in explaining the remaining variance, but are not relevant for comparing these two analyses.

In summary, the dispersion-based analysis fits the Kazakh data better than Kaun's feature-based analysis, which is evidence for a dispersion-based reinterpretation of perceptual weakness in Kaun (1995). The next two sections extend this argument to other languages, demonstrating that the dispersion-based account can both account for other data in Kaun (1995, 2004), as well as data that is exceptional under her analysis.

6 Testing the analysis

The previous section showed that the dispersion-based analysis outperformed Kaun's feature-based analysis for Kazakh. Whereas the previous sections focused exclusively on Kazakh, this section demonstrates that the dispersion-based analysis can account for a labial harmony in four languages with a harmony pattern that is significantly different from that in Kazakh.

In languages that restrict harmony in some way, two distinct sub-types based on vowel height emerge from Kaun's typology: the [+hi] target type, where targets are [+hi] irrespective of trigger height, and the [-hi] trigger-target type, where both trigger and target are obligatorily [-hi]. Most Turkic languages exemplify the [+hi] target type, although [+hi] has been reinterpreted as [-long] in Kazakh. The [-hi] trigger-target type is attested in Mongolic and Tungusic languages, as well as a number of African languages (Essegbey & McCollum 2017). This section demonstrates that the dispersion-based analysis can also account for the [-hi] trigger-target pattern with data from four Altaic languages.

6.1 Categorical dispersion-based constraints

The scalar alignment constraint used in the MaxEnt analysis in §5 captured a large amount of variance in the gradient Kazakh data. However, many descriptions of harmony delineate the set of triggers and targets categorically. This is usually accomplished by constraint weighting alone, but is complicated by the use of a scalar constraint. To derive categorical rather than gradient predictions using a scalar constraint, a MaxEnt model would need to produce a sigmoid curve with a steep slope (see McPherson & Hayes 2016; Hayes 2017 for sigmoids), and additionally, participants and non-participants would need to occur at or near the asymptotes of the curve Unless positions on the curve are guaranteed to occur near the function's minimum or maximum, then weighting is insufficient, in and of itself, to differentiate categorical from gradient using scalar constraints in MaxEnt. I sidestep this issue, as it is not central, and demonstrate my point by using strictly-ranked constraints enforcing a perceptual distance threshold.

In this section, I use Kaun's conditional alignment constraints to compel harmony. While Kaun's constraints drive harmony when the trigger is a member of some featural class, the alignment constraint in (27) motivates harmony if the perceptual distance between the trigger and its harmonic counterpart is less than some threshold, *m*. This threshold is language-specific. This constraint predicts that the set of triggers in a given language will be the set of [+rd] vowels who are closest to their harmonic counterparts.

(27) ALIGN-L/R([RD]/$\Delta_{xy}<m$) align the feature [+rd] to the left/right edge of the word if the distance between the [+rd] trigger, *x*, and its [-rd] harmonic counterpart, *y*, Δ_{xy}, is less than *m*.

Similarly, I define the set of targets by perceptual distance. Whereas Kaun contends that the set of targets in a language will be those where the feature [rd] is most perceptually salient, instead, I argue that the set of targets, like the set of triggers, is defined by weakness. This parallels the analysis of triggers- pairs separated by smaller perceptual distances will be more likely to surface via harmony. This is formalized via an IDENT constraint below, although it is also possible to encode the preference for minimally salient alternations via *MAP constraints that penalize output-output alternations rather than input-output mismatches (Zuraw 2013). The argument that perceptual weakness determines the set of targets follows directly from Steriade's (2009) P-Map proposal in two ways. First, Steriade contends that speakers have access to the relative perceptibility of a set of alternations. Second, she argues that phonotactically-illicit forms are repaired by the least-salient possible alternation. For instance, a prohibition on word-final voiced obstruents is always repaired via devoicing, although numerous other repairs are possible. She argue that devoicing is privileged in this way because it a voicing alternations involves a smaller perceptual change than other possible alternations. In a similar vein, the vowels most likely to undergo harmony are those that involve a less salient alternation.

(28) IDENT-IO ([RD]/Δ_{vw}>*n*) assign a violation to every input-output [rd] pair, *v* and *w,* with a perceptual distance, Δ_{vw}, greater than some threshold, *n*.

Two things are worth noting at this point. First, using input-output correspondence to curtail harmony involves a significant idealization since in the vast majority of languages with labial harmony there exists some other harmony pattern. Thus, affixes typically undergo four-way alternations, and not just the two-way alternations addressed here. For any four-way alternation, evaluating perceptual distances is potentially problematic. Second, input-output correspondence constraints with phonetic detail entail phonetically-specified inputs (Flemming 2008b).

6.2 Khalkha Mongolian

Khalkha Mongolian has the following seven phonemic vowels: /a e i ɔ o ʊ u/. Labial harmony in Khalkha is triggered by non-high vowels and targets non-high vowels only (Svantesson 1985:318-320; Kaun 1995:48-53), shown in (29). Non-high suffix vowels surface as unrounded after unrounded or high round vowels (29a-d), but after non-high round vowels, these non-high suffix vowels undergo rounding (29e-h).

(29) Khalkha

a.	jav-laː	'go-PST'
b.	deːl-eːr	'coat-INST'
c.	tʊːlai-gaːr	'hare-INST'
d.	uz-leː	'jump-PST'
e.	ɔr-lɔː	'enter-PST'
f.	ɔr-ɔːd	'enter-PFV'
g.	og-loː	'give-PST'
h.	tor-oːd	'be born-PFV'
i.	ɔr-ʊːl-aːd	'enter-CAUS-PFV'
j.	tor-uːl-eːd	'be born-CAUS-PFV'
k.	xɔt-ɪːxɔː	'town-REFL.GEN'
l.	tomr-iːxɔː	'iron-REFL.GEN'

The high unrounded vowel, /i/, as well as its [-ATR] allophone, /ɪ/ (Svantesson 1985:301), is transparent to harmony (29k,l), but the high round vowels, /u/ and /ʊ/, block harmony (29i,j). As the high round vowels do not trigger harmony (29c,d) nor do they surface via harmony from underlying /i/, the dispersion-based analysis predicts that the perceptual distance between /u/ and /i/ should be greater than the perceptual distances between /e/-/o/ and /a/-/ɔ/.

Svantesson (1985:290-291) recorded three repetitions of each long phonemic vowel in initial position from one Khalkha speaker. Using the procedures detailed in §5.2.2, I calculated perceptual distances between each relevant vowel pair from Svantesson's data.[13] Only attested (i.e. ATR and [high] harmonic) pairings were considered.

Mean F1-F2 and corresponding perceptual distances for each harmonic pair are shown in Table 10. Since non-high vowels trigger and surface via harmony, the distances between /a/-/ɔ/ and /e/-/o/ should be smaller than the distances between /ɪ/-/ʊ/ and /i/-/u/.

Table 10: Mean F1-F2 (ERB) and perceptual distances in Khalkha

Vowel	F1	F2	Harmonic Pairing	Perceptual Distance (λ(F2)=0.473)	Perceptual Threshold $2.2 < m < 2.5$
e	9.919	18.429	e-o	1.135	Participant
o	9.964	16.03			
a	13.604	17.827	a-ɔ	2.196	
ɔ	11.69	15.549			
ɪ	10.467	20.008	ɪ-ʊ	2.597	Non-participant
ʊ	10.031	14.591			
i	8.39	20.909	i-u	3.361	
u	8.232	13.804			

As is evident above, the perceptual distances between the non-high pairs is less than the distances between the two high vowel pairs. Within these possible alternations, the attested labial harmony alternations, as predicted, are those with the smallest perceptual distances.

6.3 Inner Mongolian dialects

In dialects of Mongolian spoken in China, like Khalkha, non-high vowels trigger harmony on following non-high vowels. However, in these dialects, roughly corresponding to the Shuluun Höh data in Kaun (1995:53-58), the vowel inventory is larger than the Khalkha inventory. These dialects possess the vowels /æ œ ø ə/ in addition to the seven phonemic vowels in Khalkha.[14] Note that Svantesson finds no evidence of /ø/ during data collection (1985:290).

One interesting complication to labial harmony in Inner Mongolian dialects is the prohibition on front rounded vowels in suffixes, exemplified in (30). In (30a,b), /œ/ triggers labial harmony on the instrumental suffix, which is always [+back]. In (30c,d), harmony does not obtain when the

[13] Like in §5.2, F3 was ignored for calculating perceptual distances.
[14] Svantesson (1985) suggests that these vowels derive from historical umlaut.

suffix is underlyingly [-back], like the comitative suffix. In these cases, disharmony occurs due to the restriction that suffixal [+rd] vowels must be [+bk].

(30) Inner Mongolian
a. nɔxœː-gɔːr ‘dog-INST’
b. mœrj-ɔːr ‘horse-INST’

c. ɔd-tæː ‘star-COM’
d. obs-teː ‘grass-COM’

Svantesson (1985:290-291) presents acoustic data from two Inner Mongolian dialects, Baarin and Šiliingol. Like in Khalkha, the distances between the non-high round vowels and their harmonic counterparts should be smaller than the distances between the high vowel pairs. Mean F1-F2 and perceptual distances are shown in Tables 11 and 12 below. The asterisks below mark that the vowel, /œ/, only occurs within roots, so it does not alternate for harmony.

Table 11: Mean F1-F2 (ERB) and perceptual distances in Baarin

Vowel	F1	F2	Harmonic Pairing	Perceptual Distance (λ(F2)= 0.495)	Perceptual Threshold 3.5 < *m* < 4.3
ə	9.576	16.725	ə-o	1.062	Participant*
o	10.208	15.002			
æ	11.766	19.425	æ-œ	1.202	
œ	10.699	20.547			
a	14.464	17.827	a-ɔ	3.413	
ɔ	11.306	15.214			
i	7.962	22.015	i-u	4.279	Non-participant
u	7.629	13.773			
ɪ	8.801	22.027	ɪ-ʊ	4.323	
ʊ	10.208	13.773			

Once again, the high vowel pairs are separated by greater perceptual distances than the non-high vowel pairs in both dialects. The relative distance between the two high vowel pairings differs across dialects, but both pairs are the most dispersed in each dialect.

Table 12: Mean F1-F2 (ERB) and perceptual distances in Šiliingol

Vowel	F1	F2	Harmonic Pairing	Perceptual Distance (λ(F2)= 0.441)	Perceptual Threshold 2.5 < *m* < 3.0
æ	12.625	19.605	æ-œ	1.196	Participant*
œ	12.487	16.912			
ə	8.851	18.844	ə-o	1.59	
o	9.505	15.562			
a	13.834	17.329	a-ɔ	2.46	
ɔ	11.442	16.03			
ɪ	10.186	21.455	ɪ-ʊ	3.023	Non-participant
ʊ	9.85	14.647			
i	9.025	21.402	i-u	3.546	
u	8.258	13.557			

6.4 Solon

In addition to the Mongolian dialects above, Svantesson (1985:296) presents acoustic data from one speaker of Solon, a Tungusic language spoken in China. The Solon inventory consists of ten vowels: /ɛ e ɪ i a ə ɔ o ʊ u/. Labial harmony in Solon, demonstrated in (31), is similar to that of Khalkha. High vowels do not trigger or undergo harmony (31a-d). The non-high back vowels both trigger and undergo harmony (31i,j). Unlike the other three languages, Solon has non-high vowels that do not participate in harmony. All such vowels are [-bk, -rd], so they cannot trigger harmony (31g), but they also do not undergo harmony (31h).[15] High vowels may occur non-initially regardless of initial vowel roundness (31e,f).

(31) Solon (Li 1996:104; Tsumagari 2009)

a. ʊlda ‘quilt’
b. uldə ‘meat’
c. ɪɪlda ‘soot’
d. iigu ‘grindstone’
e. ʤaxʊn ‘eight’
f. xoŋgi ‘bucket’
g. bɛɛga ‘moon’
h. omoleː ‘grandson’
i. ɔʃxɔ ‘fish’
j. toʃʃo ‘cloud’

Based on the harmony pattern, the back non-high vowel pairs, /ə/-/o/ and /a/-/ɔ/, should be less dispersed than the high vowel pairs, as well as the potential but unattested pairings, /e/-/o/ and /ɛ/-/ɔ/. Table 13 shows mean F1-F2 and perceptual distances for Solon.

[15] No instances of short /e/ were found in the descriptions used.

Table 13: Mean F1-F2 (ERB) and perceptual distances in Solon

Vowel	F1	F2	Harmonic Pairing	Perceptual Distance (λ(F2)=0.298)	Perceptual Threshold $1.0 < m < 2.1$
ə	10.403	16.314	ə-o	0.443	Participant
o	10.317	14.853			
a	12.329	18.261	a-ɔ	0.904	
ɔ	11.69	16.114			
ɛ	10.552	21.277	ɛ-ɔ	1.913	Non-participant
ɪ	10.208	21.468	ɪ-ʊ	2.151	
ʊ	10.295	14.249			
e	8.876	22.178	e-o (potential, but unattested)	2.614	
i	8.044	22.349	i-u	2.698	
u	8.364	13.353			

Like the three languages above, this prediction is borne out in the perceptual distances calculated from Svantesson (1985)- the back non-high pairs are less dispersed than other, both actual and potential, harmonic pairs.

One reviewer questions the typological consequences of allowing λ to differ between languages, as it does in all the languages above. In each language, the dimensions of the vowel space differ, so λ is set to create a normalized 0.625:1 perceptual space based only on each language's range of F2-to-F1 contrasts. It seems relatively clear that F1 must be weighted over F2, both for psycholinguistic and typological reasons (see Schwartz et al 1997:§4). It is not clear, however, exactly how much to weight F1 over F2. Schwartz et al. (1997) suggests a range of 0.5-0.75:1 because their model performs better under this assumption (Padgett & Tabain 2005). If λ were fixed, it would allow inventories to exhibit a wide range of shapes. Fitting each vowel space's F2-to-F1 contrast ratio to 0.625:1 doesn't involve changing a parameter, but rather holding a parameter (the ratio of F2-to-F1 contrasts) constant, which requires changing λ, since λ is just a function of that ratio and the language-specific vowel space. This is similar to the uniform auditory space used in Flemming (2002), which would be defeated if lambda did not vary across languages. Since it is an open question how much λ should vary, if at all, I recalculated the perceptual distances above using 0.723, the lambda value from §5.2, and the results did not differ.

6.5 Analysis of Solon

In this subsection I briefly sketch an analysis of Solon with strict ranking. The alignment constraint in (32) and the faithfulness constraint in (33), both of which use a threshold of 1 (see Table 11), can account for the distribution of non-initial vowels in Solon.

(32) ALIGN-R([RD]/Δ_{xy}< 1) align the feature [rd] to the right edge of the word if the distance between the [rd] trigger, *x*, and its harmonic pair, *y*, Δ_{xy}, is less than 1.

(33) IDENT-IO ([RD]/Δ_{xy}> 1) assign a violation to every input-output [rd] pair, *x* and *y,* with a perceptual distance, Δ_{xy}, greater than 1.

The first tableau, (34), shows that ALIGN-R must dominate ID-IO[RD]. This ranking dictates that harmony occurs after perceptually weak triggers.

(34)

	/toʃʃə/	ALIGN-R([RD]/Δ< 1)	ID-IO[RD]
	toʃʃə	*!	
☞	toʃʃo		*

The tableau in (35) shows that the dispersion-based IDENT constraint must outrank ALIGN-R to prevent assimilation of high vowels.

(35)

	/xoŋgi/	ID-IO([RD]/Δ> 1)	ALIGN-R ([RD]/Δ< 1)	ID-IO[RD]
☞	xoŋgi		*	
	xoŋgu	*!		*

In (36), /u/ does not trigger harmony on a following non-high vowel because it does not satisfy the conditional requirement of ALIGN-R([RD]/Δ< 1) since $\Delta_{\text{u-i}}$ is 2.7, leaving faithfulness to militate against assimilation.

(36)

	/uldə/	ID-IO([RD]/Δ> 1)	ALIGN-R([RD]/Δ< 1)	ID-IO[RD]
☞	uldə			
	uldo			*!

The general ranking instantiated above, ID-IO([RD]/Δ>*n*) >> ALIGN-R([RD]/Δ< *n*) >> ID-IO[RD], holds for all four languages described in this section. The perceptual threshold, *n*, for each language varies, but the ranking schema remains constant.

In this section, I have exemplified how the dispersion-based analysis made the correct predictions for four Altaic languages using categorical alignment and faithfulness constraints. §5 demonstrated that the dispersion-based analysis more successfully accounts for the Kazakh data, and this section extended the scope of the dispersion-based claim to four non-Turkic languages. The best test of the proposal would be languages with gradient harmony, more similar to Kazakh. In the absence of such data, categorical patterns examined in this section still corroborate the Kazakh results. The next section argues that the dispersion-based analysis constructs a better typology than Kaun's analysis.

7. The typology of labial harmony

In this section, I discuss the typological predictions of the analysis. Before addressing the larger typology, though, I discuss labial harmony in Mayak, which undermines the larger typological predictions of Kaun's analysis, yet receives a straightforward analysis under the dispersion-based analysis. In §7.2, I introduce two constraints linking perceptual weakness with duration. Using the set of constraints introduced in §6.1 and §7.2, I lay out typological claims of the dispersion-based analysis, comparing them with the typology in Kaun (1995, 2004).

7.1 Problematic cross-height harmony

7.1.1 The prediction

Following Steriade (1981), Kaun's typology makes one claim very clearly: cross-height harmony targeting a non-high back vowel is highly marked. An implicational universal emerges from Kaun (1995), if cross-height harmony targets a non-high back vowel, then harmony is unrestricted. Recall from §3.2 the five constraints used in her analysis, reproduced in (37).

(37) Kaun's typological generalizations (2004:92)

Conditions favoring harmony:	Relevant constraint
1. The trigger is [-hi].	ALIGN-L/R([RD]/[-HI])
2. The trigger is [-bk].	ALIGN-L/R([RD]/[-BK])
3. The target is [+hi].	*[+RD,-HI]
4. Trigger and target agree in height.	GESTURALUNIFORMITY[RD]
5. The trigger is short.	ALIGN-L/R([RD]/[-LONG])

If labial harmony in a given language, like Akan (Dolphyne 1988), is unrestricted, then a general ALIGN([RD]) constraint outranks all relevant markedness and faithfulness constraints. If, however, harmony is restricted in some way, then harmony must either target front vowels, high vowels, or must occur only when trigger and target agree in height. If harmony is unrestricted among front vowels, as in older descriptions of Kazakh (Korn 1969:101-102), this is generated by the ranking: ALIGN([RD]/[-BK]) >> *[+RD,-HI]). This ranking allows harmony on non-high vowels despite the markedness constraint, *[+RD,-HI]. If harmony targets only high vowels, as in Turkish, then *[+RD,-HI] >> ALIGN([RD]). In both cases, harmony does not target non-high back vowels. If harmony targets non-high back vowels, then trigger and target must both be [-hi]. The [-hi] trigger-target patterns in §6 derive from the ranking, GESTURALUNIFORMITY[RD] >> ALIGN([RD]/[-HI]) >> *[+RD,-HI]. Crucially, there is only one ranking that generates cross-height harmony on a non-high vowel, and that is the unrestricted pattern, where ALIGN([RD]) >> *[+RD,-HI], as in Akan.

7.1.2 Mayak

Mayak, a western Nilotic language spoken in South Sudan, possesses the following eight phonemic vowels: /i ɪ ɛ ʌ a ɔ ʊ u/, with two additional surface vowels, [e o] (Andersen 1999). Mayak exhibits both bidirectional ATR harmony and optional regressive labial harmony. In

(38a,b), a following /u/ optionally triggers assimilation of /ʌ/ to [o]. In contrast, /i/ does not undergo harmony in the same environment (38c). None of the [-ATR] vowels trigger or undergo harmony (38d-g). In short, /u/ is the only trigger and /ʌ/ is the only target of harmony. Harmony is thus cross-height and targets only a non-high back vowel.

(38)

a.	tʌk-uð-i	~	tok-uð-i		'wash.AP-PST-SUF'
b.	ʔʌm-uð-i	~	ʔom-uð-i		'eat.AP-PST-SUF'
c.	ʔið-u		*ʔuð-u		'shape.with.axe-PST'
d.	mɪɣ-ʊk		*mʊɣ-ʊk		'spider-PL'
e.	wɪl-ɔl		*wʊl-ɔl		'guest-SG'
f.	wɛlɣ-ɔn		*wɔlɣ-ɔn		'rib-SG'
g.	maað-ɔnɔn		*mɔɔð-ɔnɔn		'drink-1P.EXC'

There is no ranking of Kaun's constraints that can motivate the assimilation of /ʌ/ to [o] before /u/ since harmony in Mayak is not unrestricted. The context that Steriade (1981) and Kaun (1995) explicitly exclude from the typology, cross-height harmony on a non-high back vowel, is thus attested in Mayak. The same exceptional pattern is also attested in closely-related Kurmuk (Andersen 2007). If one modifies Kaun's analysis, either by an additional alignment constraint favoring [+hi] triggers or a constraint favoring [-hi] targets, then the restrictiveness of Kaun's analysis is undermined. Crucially, every featural combination of triggers and targets is possible, since Mayak fills in the one gap excluded by Kaun's analysis.

Given a pattern like Mayak, we are either forced to conclude that there are no substantive restrictions on labial harmony, or that these restrictions cannot be expressed via features. If there are no restrictions on harmony, there is no explanation for the overwhelming similarity between labial harmony patterns in unrelated languages. If we maintain that there are real restrictions on harmony, then cases like Mayak demonstrate that these restrictions cannot be featural in nature. McCollum (2017) argues this point, using both the auditory representations in Flemming (2002) and phonetic data from related Maa (Guion et al. 2004), to show that the Mayak pattern falls out from a dispersion-based restriction on targets. Conceived this way, all [+rd] vowels may trigger harmony, but a restriction on salient [rd] alternations limit the set of targets to /ʌ/. Perceptual distances from both calculations are shown in Table 14 below.[16] Observe in Table 14 that the only attested alternation, ʌ-o, is the least salient one. The Mayak pattern, like those in Kazakh and in the four languages in §6, falls out from perceptual weakness, defined in systemic dispersionary terms.

[16] The mid vowel, [o], is phonotactically banned from non-initial positions, so may not occur as a trigger.

Table 14: Perceptual distances in Mayak (McCollum 2017:6-7)

Pairing	Perceptual Distance (Flemming)	Perceptual Distance (Maa, in ERB)
ʌ-o	1.18	1.97
a-ɔ	1.6	2.23
ɛ-ɔ	1.88	2.36
ɪ-ʊ	2.5	2.74
e-o	2.5	2.92
i-u	2.5	3.35

Thus far we have seen that the dispersion-based analysis outperforms the feature-based analysis for Kazakh. We have also seen that the dispersion-based analysis can account for harmony in non-Turkic languages. This subsection has argued that the dispersion-based account can handle data from Mayak that is exceptional under Kaun's analysis. In the following subsection, I contend that the dispersion-based analysis generates better typological predictions than the feature-based analysis.

7.2 Typological predictions

7.2.1 Duration and perceptual weakness

I have claimed that the best triggers are also the best targets for harmony, in accordance with Kaun's GESTERALUNIFORMITY constraint. In all four of the languages discussed in §6, high vowels were always more perceptually distinct than the non-high vowels. In many languages, this is likely the case. For this reason, Kaun (1995) claims that [+hi] vowels are intrinsically better targets for harmony because they better signal the [rd] feature of the trigger vowel. Kaun's claim accounts for languages like Turkish, where regardless of trigger height, only high vowels undergo harmony (see also Ultan 1973:44-47). Using only the perceptual similarity-based constraints introduced thus far, this analysis cannot account for Turkish, since it seems likely that high vowels are more distinct from their harmonic counterparts than non-high vowels (Kiliç & Öğüt 2004).

If the [+hi] target pattern does not derive from target salience, then one alternative is that it results from duration. For targets, shorter vowels are better targets for the same reason that auditorily similar vowels are better targets- less noticeable alternations are preferred (Steriade 2009). Moreover, there is evidence that suggests duration plays a significant role in vowel perception, particularly in more crowded regions of the vowel space (Bennett 1968; Ainsworth 1972; Hillenbrand et al. 2000). If shorter vowels are more difficult to correctly discriminate, then they may preferentially participate in harmony. As in §6.1, the constraint in (39) establishes a threshold to demarcate the set of undergoers from non-undergoers.[17]

(39) IDENT([RD]/ $T_v > m$) — let v be an output vowel, with duration, T, and w be its input correspondent, if $T_v >$ some threshold, m, assign a violation to every w-v pair that disagree for [rd].

[17] This likely relates to diachronic claims on the emergence of vowel harmony (Ohala 1994; Przezdziecki 2005).

Kazakh, as well as Maltese (Puech 1978), shows that contrastive length may play a role in harmony. I predict, however, that non-contrastive length may also factor into harmony. Many Turkic languages exhibit duration differences between the putative high and non-high vowels that mirror the differences in Kazakh, although there is no obvious length distinction in many of these languages. Additionally, high (or alternatively, short) vowels undergo elision in many Turkic languages, further suggesting the intrinsic shortness of these vowels (e.g. Poppe 1964; Kavitskaya 2013; Washington 2016). Given the shortness of high vowels in Turkic, a constraint like (39), offers an account for this genetic skew towards high vowel targets in Turkic. This pattern, which motivates Kaun's *[+RD,-HI] constraint, is almost entirely confined to Turkic (see Derbyshire 1979; Casali 1995 for non-Turkic examples).

In addition to defining the set of targets, duration may play a crucial role in determining the set of triggers. In Kazakh, the best triggers are not only the least dispersed vowels. They are also the short vowels. In Kachin Khakass (Korn 1969:102-103), harmony only obtains when a high vowel triggers assimilation of another high vowel. If, as noted above, the high vowels are far shorter than the non-high vowels in Kachin Khakass, then harmony may be initiated by short vowels, targeting those same short vowels. I define a duration-related alignment constraint in (40).[18]

(40) ALIGN-L/R([RD]/ $T_x < n$) align the feature [rd] to the left/right edge of the word if the duration of the trigger, x, is less than some threshold, n.

Before proceeding to the typology, note that the analysis curtails harmony by faithfulness. This stands in contrast to Kaun's analysis, which limits harmony by universal markedness constraints, like *[+RD,-HI]. A universal dispreference for non-high round vowels is not consistent with the dispersion-based analysis because perceptual weakness hinges, not on universal properties of each vowel, but rather on system-internal factors. If a systemic notion of perceptual weakness motivates asymmetric trigger relations, then it is more parsimonious if the same force underlies the relevant target asymmetries, too. This position cannot be captured by universal markedness. Instead, perceptual similarity must be defined within each system, and therefore requires language-specific factors (though these are not necessarily faithfulness-related) rather than feature co-occurrence restrictions, like *[+RD,-HI], to play a role in the analysis.

Framed in language-specific terms, the set of targets in Kazakh is not defined by a ban on long round vowels, as is encoded by the constraint, *[+RD,+LONG], used in §4-5. Like *[+RD,-HI], a feature co-occurrence constraints against long round vowels, which is typologically suspect, should be replaced by constraints using language-specific detail. For Kazakh then, long vowels fail to undergo harmony due to an IDENT([RD]/ $T_v > m$) constraint, like (39), set to some value around 50 ms. Alternatively, the long vowels may fail to undergo harmony in Kazakh because they are the most dispersed. Either way, using faithfulness to militate against harmony depends on assuming [-rd] non-initial inputs, contra Richness of the Base. If a non-initial vowel is underlyingly [+rd], as it may be if inputs are not restricted, then an IDENT constraint like (39)

[18] It is likely that different languages respond in distinct ways to length differences, although in this paper this second dimension is proposed as merely a convenient way to account for length differences that are not well-understood.

will problematically compel that vowel to surface as [+rd], and thus avoid a change, rather than compelling unrounding. Thus, perceptual distance- and duration-based IDENT constraints like those used in the analysis can only promote faithfulness and not a dispreference for harmony unless non-initial inputs are restricted.

The importance of systemic factors in harmony is also supported by recent work on the typology of ATR harmony in African languages. For instance, earlier work argued for a universal dispreference for [+hi,-ATR] vowels (e.g. Archangeli & Pulleyblank 1994). This conclusion, though, is undermined by the prevalence of [+hi,-ATR] vowels outside of West Africa (Casali 2008, 2014). Casali (2008, 2014) contends that language-specific inventories dictate the nature of ATR harmony, in parallel with the claim in this paper. What were construed as universal factors affecting harmony may, in reality, derive from system-internal considerations.

7.2.2 An outline of the typology

Below, I present a typology of labial harmony recast in system-internal terms. The typology was generated using two threshold-based alignment constraints motivating harmony from less-dispersed and from shorter [+rd] vowels, and two threshold-based faithfulness constraints under strict ranking. First, though, several caveats are necessary. I limit the number of [+rd] vowels participating in harmony to four, partly out of convenience, and partly because even in languages with more than four [+rd] vowels, like Baarin and Šiliingol, additional vowels are typically marginal or derive historically from other vowels in the inventory. I additionally assume that a language only makes a two-way distinction for duration, even when length is non-contrastive. This stems partly from convenience but also from the fact that languages with contrastive length tend to make only one length distinction (Remijsen & Gilley 2008). The numbers 1-4 in the typology represent the least to the most perceptually salient round vowels, while [short] represents the shortest (two) round vowels in a given language, even if length is non-contrastive. The predictions below involve only two harmony-driving constraints without reference to possible stringency rankings, scalar implementations, or similar formal mechanisms.

The typology predicts up to 25 patterns, given the stipulations just noted, which is comparable to the 24 harmony patterns predicted by Kaun's (1995) set of five constraints.[19] Of the 24 possible patterns in Kaun (1995), 9 are attested. Similarly, about nine patterns are attested in the proposed typology. Superficially, this proposal does not provide better empirical coverage than Kaun's constraint set. However, some key generalizations emerge from the typology in Table 15 that are not evident in Kaun's work. Moreover, Kaun assumes the gaps in her typology are accidental (2004:109), but the typology below suggests that certain gaps are very principled.

First, the set of targets is never larger than the set of triggers. There are no known labial harmonies where the set of targets is not a subset of the set of triggers. The unattested Type 16-19 patterns all violate the trigger-target subset relation. Of the 6 unattested patterns among Types 1-15, Types 4, 8, and 9 also violate this trigger-target subset relation. Further, of the (potentially) unattested Types 20-25, two patterns, Type 25 (and potentially Type 20), also violate this relation. These gaps suggest that a theory of labial harmony cannot depend entirely

[19] Kaun (2004:109) notes the addition of a sixth constraint, *[+RD,-BK], increases the typological space to 36. With a seventh constraint, ALIGN-L/R([RD]/[-LONG]), the number of distinct patterns increases to 131.

on trigger restrictions (see also Nevins 2010). As argued in McCollum (2017), the typology of labial harmony does appear to depend more on GESTURALUNIFORMITY, the constraint on trigger-target relations, than constraints on triggering vowels alone.

Second, languages tend to use either duration or perceptual distance, but not both, to restrict harmony, as in Types 20-25. With this in mind, the Kazakh data is particularly interesting, since triggers are defined by perceptual distance, but targets are definable either by perceptual distance or duration. Like Kazakh, the Type 21 Tofa pattern (Harrison 2000), is analyzable without reference to duration, which would render it an example of the Type 3 pattern. So, it is unclear if any languages clearly exhibit a pattern that depends on both perceptual distance and duration. This observation may connect with some speech perception research that has argued that hearers preferentially attend to only one auditory cue, even when multiple cues are available, in accordance with the claim here (Flege & Hillenbrand 1986; Goudbeek et al. 2008).

If these biases exist, then the proposed model fits the typology very well. If not, the typology is at least as empirically adequate as Kaun (1995), and more so than Kaun (2004). Lastly, the gaps in her typology could be construed as accidents, but the gaps in Table 13 fall into several classes that all suggest plausible reasons for their non-existence. The proposed typology thus offers new insights into labial harmony to motivate further typological, experimental, and psycholinguistic work.

Table 15: A schematic typology of labial harmony (unexpected patterns are shaded gray, languages with multiple possible analyses are entered twice and marked with ?)

Macrotype	Type	Number	Triggers (1=least dispersed, 4=most dispersed)	Targets (1=least dispersed, 4=most dispersed)	Example
Unrestricted	Unrestricted	1	unrestricted	unrestricted	Akan (Dolphyne 1988), Altai Tuvan (Harrison 2000)
Same parameter (unexpected languages that violate the trigger-target subset relation are shaded gray)	Least Dispersed Trigger-Least Dispersed Target	2	[1]	[1]	
		3	[1,2]	[1]	Tofa (Harrison 2000)?
		4	[1]	[1,2]	
		5	[1,2]	[1,2]	Khalkha, Solon
		6	[1,2,3]	[1]	
		7	[1,2,3]	[1,2]	Kazakh?
		8	[1]	[1,2,3]	
		9	[1,2]	[1,2,3]	
		10	[1,2,3]	[1,2,3]	Kyzyl Khakass
	Short Trigger-Short Target	11	[short]	[short]	Kachin Khakass, Yeyi (Seidel 2008)
Target restrictions	Unrestricted Trigger- Least Dispersed Target	12	unrestricted	[1]	Mayak, Meadow Mari (Vaysman 2009; Estill 2012)
		13	unrestricted	[1,2]	
		14	unrestricted	[1,2,3]	Older Kazakh
	Unrestricted Trigger- Short Target	15	unrestricted	[short]	Turkish, Crimean Tatar (Kavitskaya 2013)
Trigger Restrictions (unexpected because these violate the trigger-target subset relation)	Least Dispersed Trigger-Unrestricted Target	16	[1]	unrestricted	
		17	[1,2]	unrestricted	
		18	[1,2,3]	unrestricted	
	Short Trigger-Unrestricted Target	19	[short]	unrestricted	
Different parameters (unexpected because languages tend to use only one cue)	Least Dispersed Trigger- Short Target	20	[1]	[short]	
		21	[1,2]	[short]	Tofa?
		22	[1,2,3]	[short]	Kazakh?
	Short Trigger-Least Dispersed Target	23	[short]	[1]	
		24		[short]-[1,2]	
		25		[short]-[1,2,3]	

8 Discussion

Traditionally, DT has been used to model vowel inventories. Within Optimality theory, though, DT has been used to account for phenomena beyond inventory structure. Padgett (2004) uses DT to analyze vowel reduction in Russian. Flemming (2006, 2008a) uses DT to address positional markedness, enhancement, and surface phonotactics. The present work continues Optimality theoretic work, extending the influence of DT into another realm of phonology.

Given the similarity between Flemming's work and the present analysis, it is important to address how they differ. Using Flemming's grid of auditory representations, we could position the Kazakh vowels as shown in Figure 9. In Figure 9, the distance between a contrastive pair is equal to the number of gridlines between each member of the pair for a given acoustic dimension (e.g. F1 or F2). Thus, /ɪː/ and /ʏː/ differ in F2 by two, while /ɪ/ and /ʏ/ differ by only one. Among the back vowels, /ə/ and /ɔ/ are also separated by only one gridline, whereas /ɑː/ and /ɔː/ differ by one F2 gridline and two F1 gridlines. The Euclidean distance between /ɑː/ and /ɔː/ is approximately 2.2. So, the distances would give us the general hierarchy, ʏ, ɔ > ʏː > ɔː because the distances between the long vowels are greater than those differentiating the short vowels, but the distance between /ɪ/ and /ʏ/ is identical to that between /ə/ and /ɔ/. Flemming's representational schema captures the same general insight, but is constrained by the granularity of its representations. By using greater acoustic detail, the present work is able to capture crucial generalizations that are not observable from featural or discretized auditory representations (see also Padgett 2004), like the increased frequency of harmony after /ʏ/ relative to /ɔ/.

Figure 9: Flemming's (2002) vowel space for the Kazakh vowels (cf. Flemming 2004, 2006)

F2

5	4	3	2	1	
					1
ɪː	ɪ	ʏ / ʏː			2
			ə	ɔ / ɔː	3
					4
			ɑː		5

F1

In Flemming (2006, 2008a), dispersionary constraints not only drive inventory selection but also interact with surface-oriented markedness and faithfulness constraints in the phonology. This paper, however, has chosen, like Padgett (2004), to address surface phonological processes without reference to how the inventory is shaped. For various reasons, vowel inventories don't always appear optimal and the complex interaction of forces that underlie other aspects of contrast maintenance is beyond the purview of this paper.

While this work diverges from Flemming's work representationally, it fundamentally agrees that perceptual distance is active in the synchronic grammar. Although Lindblom's (1986) formulation of DT is couched in evolutionary terms, both Flemming's analysis and this paper maintain that speakers access dispersion synchronically. This has been contested by various authors (e.g. Boersma & Haman 2008; Hall 2011; Vaux & Samuels 2015). These authors object to the teleological underpinnings of Flemming's MinDist and Padgett's Space constraints, instead contending for non-phonological factors in the development of vowel spacing. However, the accessibility perceptual distances in the grammar does not necessarily entail a causal relationship between dispersion and harmony. The learner may establish a correlation between observed perceptual distances and rates of harmony without any attempt to optimize their vowel space.

9 Conclusion

In this paper I presented novel data from colloquial Kazakh. From these data, I presented a novel framework through which to articulate Kaun's observation that "bad vowels spread." This paper, though, operationalized perceptual weakness in systemic, dispersionary terms rather than vowel-intrinsic, feature-based terms. I demonstrated that dispersionary constraints outperform Kaun's featural constraints in a MaxEnt analysis of colloquial Kazakh. I then examined four Altaic languages, showing that, in each case, the dispersion-based predictions were borne out in each language's labial harmony pattern. I then argued that this analysis provides an account for data from Mayak that is problematic for Kaun's analysis. Lastly, I have contended that the dispersion-based analysis offers new typological insights as well as a testable means by which to evaluate the present claims in order to further our understanding of vowel harmony, as well as the types of representations accessible to the grammar.

References

Abuov, Zhoumaghaly (1994). The phonetics of Kazakh and the theory of synharmonism. *UCLA Working Papers in Phonetics* **88**:39-54.

Ainsworth, William (1972). Duration as a cue in the recognition of synthetic vowels. *JASA* 51:648-651.

Andersen, Torben (1999). Vowel harmony and vowel alternation in Mayak. *Studies in African Linguistics* **28**:1-30.

Andersen, Torben (2007). Kurmuk phonology. *Studies in African Linguistics* **36**:29-90.

Anttila, Arto (1997). Deriving variation from grammar. *Amsterdam Studies in the Theory and History of Linguistic Sciences* **4**:35-68.

Archangeli, Diana, and Douglas Pulleyblank (1994). *Grounded phonology*. MIT Press.

Archangeli, Diana, and Douglas Pulleyblank (2007). Vowel Harmony. *The Cambridge handbook of phonolog*y, Paul de Lacy (ed.), 353-378.

Beckman, Jill (1997). Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology* **14**:1-46.

Bennett, David (1968). Spectral form and duration as cues in the recognition of English and German vowels. *Language and Speech* **11**:65-85.

de Boer, Bart (2000). Self-organization in vowel systems. *JPh* **4**:441-465.

Boersma, Paul, and Silke Hamann (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* **25**:217-270.

Boersma, Paul, and Bruce Hayes. (2001) Empirical tests of the gradual learning algorithm. *LI* **32**:45-86.

Boersma, Paul, and Joe Pater (2016). Convergence properties of a gradual learning algorithm for Harmonic Grammar. *Harmonic Serialism and Harmonic Grammar*, John McCarthy and Joe Pater (eds.), 389-434. Equinox.

Bowman, Samuel, and Benjamin Lokshin (2014). Idiosyncratically Transparent Vowels in Kazakh. *Proceedings of the 2013 Annual Meeting on Phonology*.

Boyce, Suzanne (1990). Coarticulatory organization for lip rounding in Turkish and English. *JASA* **88**:2584-2595.

Casali, Roderic (1995). Labial opacity and roundness harmony in Nawuri. *NLLT* **13**:649-663.

Casali, Roderic (2008). ATR harmony in African languages. *Language and Linguistics Compass* **2**:496-549.

Casali, Roderic (2014). Assimilation, markedness and inventory structure in tongue root harmony systems. unpublished ms.

Cho, Taehong (2001). Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica* **58**:129-162.

Clements, George N. and Engin Sezer (1982). Vowel and consonant disharmony in Turkish. *The structure of phonological representations* **2**:213-255.

Coetzee, Andries, and Joe Pater (2011). The place of variation in phonological theory. *The Handbook of Phonological Theory,* 401-434.

Cole, Jennifer, and Charles Kisseberth (1994). An optimal domains theory of harmony. *Studies in the Linguistic Sciences* **24**:101-114.

Cole, Jennifer, and Loren Trigo (1988). Parasitic harmony. *Features, segmental structure and harmony processes,* 19-38.

Della Pietra, Stephen; Vincent Della Pietra, and John Lafferty (1997). Inducing features of random fields. *Pattern Analysis and Machine Intelligence, IEEE Transactions* **19**:380-393.

Derbyshire, Desmond (1979). *Hixkaryana*. Lingua Descriptive Studies.

Dolphyne, Florence Abena (1988). *The Akan (Twi-Fante) language: Its sound systems and tonal structure*. Ghana Universities Press.

Dzhunisbekov, A. (1972). *Glasnye kazakhskogo jazyka*. Nauka.

Dzhunisbekov, A. (1980). *Singarmonizm v kazakhskom jazyke*. Nauka.

Essegbey, James, and Adam G. McCollum (2017). Initial prominence and progressive harmony in Tutrugbu. unpublished ms.

Estill, Dennis. 2012. Revisiting the Meadow Mari vocalic system. *Linguistica Uralica* **48**:228-237.

Flanagan, James. (1955). A difference limen for vowel formant frequency. *JASA* **27**:613-617.

Finley, Sara (2008a). The interaction of vowel harmony and epenthesis. *CLS* **44**:95-109.

Finley, Sara (2008b). Formal and cognitive restrictions on vowel harmony. PhD dissertation, Johns Hopkins University.

Flege, James, and James Hillenbrand (1986). Differential use of temporal cues to the/s/–/z/ contrast by native and non-native speakers of English. *JASA* **79**:508-517.

Flemming, Edward (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* **18**:7-44.

Flemming, Edward (2002). *Auditory representations in phonology*. Routledge.

Flemming, Edward (2004). Contrast and perceptual distinctiveness. *Phonetically-based phonology*: 232-276.

Flemming, Edward (2006). *The role of distinctiveness constraints in phonology*. unpublished ms.

Flemming, Edward (2008a). Asymmetries between assimilation and epenthesis. paper presented at *The 82nd Annual Meeting of the Linguistic Society of America, Chicago*, Ill.

Flemming, Edward. (2008b). *The realized input*. unpublished ms.

Fylstra, Daniel; Leon Lasdon, John Watson, and Allan Waren. (1998). Design and use of the Microsoft Excel Solver. *Interfaces* **28**:29-55.

Goldsmith, John A., Jason Riggle, and C. L. Alan Yu, eds. (2011). *The handbook of phonological theory*. Blackwell.

Goldwater, Sharon, and Mark Johnson (2003). Learning OT constraint rankings using a maximum entropy model. *Proceedings of the Stockholm workshop on variation within Optimality Theory*, 111-120.

Goudbeek, Martijn, Anne Cutler, and Roel Smits (2008). Supervised and unsupervised learning of multidimensionally varying non-native speech categories. *Speech communication* **50**:109-125.

Guion, Susan, Mark Post, and Doris Payne (2004). Phonetic correlates of tongue root vowel contrasts in Maa. *JPhon* **32**:517-542.

Hahn, Reinhard (1991). *Spoken Uyghur*. University of Washington Press.

Hall, Daniel Currie (2011). Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* **28**:1-54.

Harrison, K. David (2000). *Topics in the Phonology and Morphology of Tuvan*. PhD dissertation. Yale University.

Harrison, K. David and Abigail Kaun (2000). Pattern-responsive lexicon optimization. *NELS* **30**.

Hayes, Bruce (2017). Varieties of Noisy Harmonic Grammar. *Proceedings of the 2016 Annual Meeting of Phonology*.
Hayes, Bruce, Robert Kirchner, and Donca Steriade, (eds.) (2004). *Phonetically based phonology*. Cambridge University Press.
Hayes, Bruce, and Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *LI* **39**:379-440.
Hsu, Brian, and Karen Jesney (2015). Scalar positional markedness and faithfulness in Harmonic Grammar. Paper presented at *CLS* **51**.
van der Hulst, Harry, and Beata Moskal (2013). Pattern of defective labial harmony in Turkic languages. *Dilbilim Araştırmaları Dergisi* **1**:17-52.
van der Hulst, Harry, and Jeroen van der Weijer (1995), Vowel harmony. In John Goldsmith (ed.), *The handbook of phonological theory*.
van der Hulst, Harry, and Norval Smith (1988a). Tungusic and Mongolian vowel harmony: A minimal pair. *Linguistics in the Netherlands*: 79-88.
van der Hulst, Harry, and Norval Smith, eds. (1988b). *Features, segmental structure and harmony processes.* Foris.
Itô, Junko, and Armin Mester (1999). Realignment. In Rene Kager et al. (eds.), *The prosody-morphology interface*: 188-217.
Johanson, Lars (1998). The history of Turkic. *The Turkic Languages*, 81-125.
Johanson, Lars, and Éva Csató, (eds.) (1998). *The Turkic Languages*. Routledge.
Jurgec, Peter (2011). *Feature spreading 2.0: A unified theory of assimilation*. PhD dissertation, University of Tromsø.
Kabak, Barıs, and Silke Weber (2013). Markedness, context, and directionality in Turkish harmony: A corpus study on vowel co-occurrence patterns. *Dilbilim Araştırmaları Dergisi* **1**:53-85.
Kara, David Somfai (2002). *Kazak*. Lincom Europa.
Kaun, Abigail (1995). *The typology of rounding harmony: An optimality theoretic approach*. PhD dissertation, UCLA.
Kaun, Abigail (1999). Epenthesis-driven harmony in Turkish. *BLS* **25**:95-106.
Kaun, Abigail (2004). The typology of rounding harmony. *Phonetically based phonology*, 87-116.
Kavitskaya, Darya (2013). Segmental inventory and the evolution of harmony in Crimean Tatar. *Turkic Languages* **17**:86-114.
Kiliç, Mehmet Akif, and Fatih Öğüt. 2004. A high unrounded vowel in Turkish: is it a central or back vowel?. *Speech communication* 43: 143-154.
Kirchner, Mark (1998). Kazakh and Karakalpak. *The Turkic languages*, 318-332.
Kirchner, Robert (1993). Turkish vowel harmony and disharmony: An Optimality Theoretic account. *Rutgers Optimality Workshop I.*
Ko, Seongyeon (2012). *Tongue root harmony and vowel contrast in Northeast Asian languages*. PhD dissertation, Cornell University.
Korn, David (1969). Types of labial vowel harmony in the Turkic languages. *Anthropological Linguistics* **11**:98–106.
Ladefoged, Peter (2001). *Vowels and consonants: An introduction to the sounds of languages*. Wiley-Blackwell.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky. (1990). *Harmonic grammar: A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations*. University of Colorado, Boulder.
Li, Bing (1996). *Tungusic vowel harmony: Description and analysis*. Holland Academic Graphics.
Liljencrants, Johan, and Björn Lindblom (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Lg* ***48***:839-862.
Linker, Wendy. (1982). Articulatory and acoustic correlates of labial activity in vowels: A cross-linguistic study. *UCLA Working Papers in Phonetics* **56**.
Lindblom, Björn (1975). Experiments in sound structure. *Revue de Phonétique Appliquée Mons* **51**: 155-189.
Lindblom, Björn (1986). Phonetic universals in vowel systems. *Experimental phonology*, 13-44.
Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *JASA* **49**:606–608.
McCarthy, John, and Alan Prince (1993). *Generalized alignment*. Springer.
McCarthy, John, and Alan Prince (1995). Faithfulness and reduplicative identity. unpublished ms.
McCollum, Adam G. (2015). Labial harmonic shift in Kazakh: Mapping the pathways and motivations for change. *BLS* **41**:329-352.
McCollum, Adam G. and Si Chen (2016). Kazakh. unpublished ms.
McCollum, Adam G. (2017). Mayak and the typology of labial harmony. *The Supplemental Proceedings of the 2016 Annual Meeting on Phonology*.
McPherson, Laura, and Bruce Hayes (2016). Relating application frequency to morphological structure: The case of Tommo So vowel harmony. *Phonology* **33**:125-167.
Menges, Karl (1947). *Qaraqalpaq grammar*. New York: King's Crown Press.
Muhamedowa, Raihan (2015). *Kazakh: A Comprehensive Grammar*. Routledge.
Nevins, Andrew (2010). *Locality in vowel harmony*. MIT Press.
Ohala, John J. (1994). Towards a universal, phonetically-based, theory of vowel harmony. *3rd International Conference on Spoken Language Processing*, 491-494.
Padgett, Jaye (2004). Russian vowel reduction and Dispersion Theory. *Phonological Studies* **7**:81-96.
Padgett, Jaye, and Marija Tabain (2005). Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica* **62**:14-54.
Pater, Joe (2009). Weighted constraints in generative linguistics. *Cognitive Science* **33**:999-1035.
Poppe, Nicholas (1964). *Bashkir Manual: Descriptive grammar and texts with a Bashkir-English glossary.* Indiana University.
Prince, Alan, and Paul Smolensky (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar.* Blackwell.
Przezdziecki, Marek (2005). *Vowel harmony and coarticulation in three dialects of Yoruba: Phonetics determining phonology*. PhD dissertation, Cornell University.
Puech, Gilbert (1978). A cross-dialectal study of vowel harmony in Maltese. *CLS* **14**:377-390.
Reetz, Henning, and Allard Jongman (2008). *Phonetics: Transcription, production, acoustics, and perception.* Wiley-Blackwell.
Remijsen, Bert, and Leoma Gilley (2008). Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect). *JPhon* **36**:318-344.

Rose, Sharon, and Rachel Walker (2011). Harmony systems. *The handbook of phonological theory*, 240-290
Schwartz, Jean-Luc; Louis-Jean Boe, Nathalie Valee, and Christian Abry (1997). The dispersion-focalization theory of vowel systems. *JPh* **25**:255-286.
Seidel, Frank (2008). *A grammar of Yeyi: A Bantu language of southern Africa*. Köppe.
Steriade, Donca (1981). Parameters of metrical harmony rules. unpublished ms.
Steriade, Donca (1995). Underspecification and markedness. *The handbook of phonological theory*, 114-174.
Steriade, Donca (2009). The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In: *The Nature of the Word: Studies in Honor of Paul Kiparsky.* MIT Press. 151-179.
Suomi, Kari (1983). Palatal vowel harmony: A perceptually motivated phenomenon? *Nordic Journal of Linguistics* **6**:1-35.
Svantesson, Jan-Olof (1985). Vowel harmony shift in Mongolian. *Lingua* **67**:283-327.
Terbeek, Dale (1977). A cross-language multidimensional scaling study of vowel perception. PhD dissertation, UCLA.
Tsumagari, Toshiro (2009). A sketch of Solon grammar. *Journal of the Center for Northern Humanities* **2**:1-21.
Ultan, Russell (1973). Some reflections on vowel harmony. *Working Papers on Language Universals* **12**:37-67
Vajda, Edward (1994). Kazakh phonology. *Opuscula Altaica: Essays Presented in Honor of Henry Schwarz*, 603-650. Western Washington University.
Vaux, Bert (1993). The origins of Altaic labial attraction. *Harvard Working Papers in Linguistics* **3**, Thrainsson et al. (eds.), 228-237.
Vaux, Bert (1998). *The phonology of Armenian*. Oxford University Press.
Vaux, Bert, and Bridget Samuels (2015). Explaining vowel systems: dispersion theory vs natural selection. *The Linguistic Review* **32**:573-599.
Vaysman, Olga (2009). *Segmental alternations and metrical theory*. PhD dissertation, MIT.
Walker, Rachel (2001). Round licensing, harmony, and bisyllabic triggers in Altaic. *NLLT* **19**:827-878.
Walker, Rachel (2011). *Vowel patterns in language*. Cambridge University Press.
Washington, Jonathan North (2016). *An investigation of vowel anteriority in three Turkic languages using ultrasound tongue imaging*. PhD dissertation, Indiana University.
Wilson, Colin and Benjamin George (2009). Maxent grammar tool [software]. http://www. linguistics.ucla.edu/people/hayes/MaxentGrammarTool/.
Xiāngrú, Zhào, and Reinhard F. Hahn (1989). The Ili Turk people and their language. *Central Asiatic Journal* **33**:260-289.
Yavaş, Mehmet (1980). Some pilot experiments on Turkish vowel harmony. *Research on Language & Social Interaction* **13**:543-562.
Zuraw, Kie (2013). *Map constraints. unpublished ms.

Appendix

Word	Partition	Gloss	Harmony
ʒʏzʏk~ʒʏzɪk	1	ring	10/11
kʏːsʏk~kʏːsɪk	2	desert carrot	10/11
qɔłɔn~qɔłən	3	colt	16/17
qɔːzɔ~qɔːzə	2	lamb	1/10
tʏlʏːk~tʏlɪːk	2	graduate	1/11
tʏːbɪː	4	hill	0/11
qɔłɑːq	4	ear	0/11
bɔłɑːt	1	steel	0/11
kʏlʏp~kʏlɪp	4	laugh-CVB	26/31
ʏːlʏp~ʏːlɪp	2	die-CVB	10/18
qɔrɔp~qɔrəp	1	construct-CVB	13/18
qɔːsɔp~qɔːsəp	3	add-CVB	1/15
kʏldʏ~kʏldɪ	1	laugh-PST.3	7/22
ʏːldʏ~ʏːldɪ	2	die-PST.3	1/20
qɔrdɔ~qɔrdə	3	construct-PST.3	1/18
qɔstɔ~qɔstə	2	spit up-PST.3	4/30
qɔːstɔ~qɔːstə	4	add-PST.3	1/20
qɔːjdə	4	put-PST.3	0/11
kʏldʏm~kʏldɪm	4	laugh-PST-1	4/12
ʏːldʏm~ʏːldɪm	3	die-PST-1	3/11
qɔrdɔm~qɔrdəm	4	construct-PST-1	3/10
qɔːstɔm~qɔːstəm	1	add-PST-1	2/10
qɔrdɔŋəzdar~qɔrdəŋəzdar	3	construct-PST-2-FORM-PL	1/12
kʏlsɪː	1	laugh-COND	0/15
ʏːlsɪː	1	die-COND	1/14
qɔrsɑː	3	build-COND	0/13
qɔːssɑː	2	add-COND	0/13
kʏlgɪːn	4	laugh-PFV	0/2
ʏːlgɪːn	1	die-PFV	0/1
kʏndʏ~kʏndɪ	2	day-ACC	3/10
ʒʏzʏktʏ~ʒʏzʏktɪ	1	ring-ACC	1/9
ʏːttʏ~ʏːttɪ	3	gall bladder- ACC	2/12
kʏndɪː	2	day-LOC	0/7
ʏːttɪː	4	gall bladder-LOC	0/11
ʒʏzʏktɪːrdɪː	3	ring-PL-ACC	23/23