

Neutral vowels and exceptional harmony in Western Yugur

Sichen Lyu*

Abstract. This study examines the highly lexicalized vowel harmony system of Western Yugur (Siberian Turkic, Gansu province, China). Specifically, it explores the representation of ‘exceptional roots’ which trigger harmony of the opposite value than indicated by surface phonetic clues (/k^hun+lAr/ → [k^hunlir]), as well as the representations of neutral (i.e. opaque and transparent) suffixal vowels. This study adopts Bellik’s (2015) Feature Domain Theory (FDT), which provides a unified account of facts regarding both neutral vowels and exceptional harmony in Western Yugur. Attention is further drawn to the phonetic value of the transparent vowel [u], which falls outside typological expectations. A tentative explanation is offered within the framework of FDT.

Keywords. Western Yugur; vowel harmony; phonological feature representation; Turkic

1. Introduction. While vowel harmony is an integral phonological process in many languages, those in which vowel harmony operates maximally across the entire lexicon are nearly impossible to come by. Disharmony is often viewed as exceptional rather than a natural component of harmony systems (Harrison 1999). This perspective led earlier theoreticians of, for instance, Turkish vowel harmony to the conclusion that all disharmonic roots must be loanwords marked with [–native] in the speakers’ mental lexicon (Lightner 1972, as cited in Clements & Sezer 1982). We would soon notice, however, that these conclusions overlook crucial aspects of the phonological representation, such as the status of transparent and opaque segments in vowel harmony (Clements & Sezer 1982; Kiparsky & Pajusalu 2003; Bellik 2015, among many more). Therefore, a theory of vowel harmony should only be regarded as complete when both regular and exceptional processes are accounted for.

On the other hand, the extent to which ‘exceptions’ are exceptional also varies from language to language. For a language like Uzbek (Boeschoten 2021) in which vowel harmony is no longer productive, what seems like exceptions to vowel harmony are instead a different, regularized system. In this study, I present data from Western Yugur, a Siberian Turkic language spoken in northwestern China, which is characterized by an abundance of roots triggering ‘exceptional harmony’ – that is, they spread the opposite feature value than what is realized on root vowels on the surface. In addition, suffixes in Western Yugur can be harmonic, opaque, or transparent. Based on the data, I argue that a feature domain-based account (cf. Bellik 2015) is capable of explaining these peculiarities in a unified fashion.

The structure of this paper is as follows. Section 2 provides an overview of the Western Yugur data. Section 3 introduces the version of the domain-based representation I adopt entitled Feature Domain Theory, ensued by an Optimality Theoretic analysis of the data. Section 4 discusses alternatives to my account. Section 5 highlights an additional issue regarding the markedness of the transparent vowel in Western Yugur. Lastly, section 6 concludes the paper.

* I would like to thank Ryan Bennett and Rachel Walker for their guidance and feedback on this project. I also thank the audience at TU+8 and an anonymous reviewer for their comments and help in my developing this paper. Authors: Sichen Lyu, UC Santa Cruz (slyu15@ucsc.edu).

2. Vowel (dis-)harmony in Western Yugur. Like many Turkic languages, Western Yugur has an eight-vowel inventory, shown in table 1, which can be symmetrically grouped by height, backness, and rounding.

	-Back		+Back	
	-Round	+Round	-Round	+Round
+High	i	y	ɯ	u
-High	e	ø	ɑ	o

Table 1: Vowel inventory of Western Yugur

In Western Yugur, backness harmony is productive across root-suffix boundary and propagates among suffixes (Roos 2000): in the regular harmonic process, the backness feature of the root spreads onto the following suffix, as shown below in the data in (1). For disharmonic roots like (1e), only the final vowel in the root can be identified as the trigger of harmony.

- (1) a. /sart-lAr/ [sart-lar] ‘Muslims’
 b. /jer-lAr/ [jer-lir] ‘places’
 c. /i^ht-lAr/ [i^ht-lir] ‘meat’
 d. /uʃt-lAr/ [uʃt-lar] ‘dog’
 e. /jiltus-lAr/ [jiltus-lar] ‘stars’

The high back unrounded vowel [ɯ] is notated differently by different authors. Roos (2000) and Zhong (2019) both describe it as a back vowel, though the latter also reports a lowered, more centralized allophone. Miao (2019) transcribes the vowel as a schwa-like central vowel. Here, I treat it as a back vowel [ɯ] for two reasons: [ɯ] regularly triggers [+back] harmony, as illustrated by (1d) and (1e). Moreover, Roos (2000) shows that historically, Common Turkic high rounded vowels in non-initial syllables tend to deround to their unrounded counterparts, such that Common Turkic *[u] develops into [ɯ] in Western Yugur (§3.1).

2.1 LEXICALIZED EXCEPTIONAL ROOTS. While historically, like most other Turkic languages, harmony among root-internal vowels is observed by native Turkic words in Western Yugur, this is no longer a feature of the synchronic lexicon, as exemplified by disharmonic items such as (1e) above. Root-internal harmony has been obscured by loanwords, compounding, as well as sound changes indiscriminately targeting the native Turkic vocabulary. Furthermore, although such sound changes may affect the backness values of vowels, Western Yugur roots tend to maintain their historical harmony classes. The two sets of processes combined have led to an apparent proliferation of *exceptional roots* in modern Western Yugur: these roots spread the opposite feature value than what is specified on the surface. As shown in the data below, the final vowels in the roots of (2a) and (2b) are specified for [+back] on the surface, yet they take the front suffix. On the contrary, while the final vowels in the roots of (2c) and (2d) are [-back] on the surface, they appear to trigger back harmony.

- (2) a. /k^hun-GA/ [k^hun-ki] ‘day-DAT’
 b. /uʃ-GA/ [uʃ-ki] ‘inside-DAT’
 c. /q^huti-GA/ [q^huti-ɣɑ] ‘Chinese person-DAT’
 d. /jim-GA/ [jim-qa] ‘to close’

A survey of the vocabulary section of Roos (2000:278-405) shows that root-final [ɯ] and [u] are the most frequent triggers of exceptional harmony. This lexical trend is a consequence of the strong diachronic tendency for Common Turkic (henceforth, CT) high front vowels to retract

regardless of roundness; hence even CT *[i] has a regular modern reflex [u] (Roos 2000:§3.1) – I further discuss this apparently unnatural direction of sound change in section 5. Only after (alveolo-)palatal segments (e.g. [j], [tɕ], [tɕʰ]) do high vowels uniformly neutralize to front vowels. Additionally, as mentioned earlier, the occurrence of rounded high vowels is restricted beyond the initial syllable. Therefore, not only do earlier [u] deround to [ʊ], but historical front rounded [y] also undergoes both retraction and derounding *[y] > *[u] > [ʊ].

As a result, in the synchronic lexicon of Western Yugur, [ʊ] is the most common root-internal exceptionality-triggering vowel. On the contrary, instances of root-final front vowels triggering exceptional back harmony are rather rare and mostly occur in predictable fronting environments. It is far more common for root-final back vowels to trigger front harmony. I will return to this observation in section 5.

To sum up, in Western Yugur, spreading of vowel harmony across the root-suffix boundary is a largely lexicalized process; roots frequently contain no surface phonetic cues to signal its harmony class.

2.2 NEUTRAL SUFFIXAL VOWELS. In contrast, derived vowels permeate regular vowel harmony; specifically, suffixal vowels that show alternation in different backness environments spread their derived backness values onto the next vowel, as shown in (3).

- (3) Regular vowel harmony among suffixes
- | | | |
|----|----------------------------|-------------------|
| a. | [k ^h un-nir-ti] | ‘day-PL-LOC’ |
| b. | [tu ^h t-pa-lar] | ‘work.hold-VN-PL’ |

These vowels are conventionally represented by archiphonemes in the underlying representations (the /A/ in (1)), under the assumption that they are underspecified for the harmonic feature. However, there also exist a large number of *neutral* suffixes which seemingly fail to undergo vowel harmony. Of these suffixes, the ones containing the vowel [ʊ] are described as being *transparent* to vowel harmony, in that they neither harmonize nor trigger harmony – they are simply ‘skipped.’ Examples are given in (4).

- (4) Transparent suffixal vowel [ʊ]
- | | | |
|----|-------------|----------------|
| a. | [ɑl-uun-qa] | ‘take-RFL-CCO’ |
| b. | [køz-uŋ-ki] | ‘eye-POS3-DAT’ |

The rest of the suffixes contain non-alternating [ɑ], [e], [i], [o], and [u] and are described as *opaque* – they do not harmonize with preceding vowels but spread their own feature values (Walker 2012). Examples are given in (5).

- (5) Opaque suffixal vowels
- | | | |
|----|----------------------------------|------------------------|
| a. | [e ^h t-mi] | ‘take-NEG’ |
| b. | [e ^h t-o-ma] | ‘take-RFL-NEG’ |
| c. | [e ^h t-us-uwal-ma-ɣɑ] | ‘take-REC-INC-NEG-CCO’ |

Note that (5c) also contains a ‘mixed’ suffix [-uwal] that involves both the transparent suffixal vowel [ʊ] as well as a non-alternating [ɑ]. I adopt the assumption that alternating suffixal vowels are underspecified in the UR, while neutral suffixal vowels are underlying. Therefore, the non-alternating [ɑ] in [-uwal] would differ underlyingly from the alternating [ɑ], which is underspecified for backness and thus harmonizes with preceding vowels. The underlying form of (5b) would then be represented as /e^ht-us-uwal-mA-GA/.

In summary, there are three types of suffixes in Western Yugur (cf. Roos 2000). I posit that the harmonic suffix represented by /A/ is underspecified in the UR; the transparent and the opaque suffixal vowels are underlyingly specified.

3. A feature domain-based analysis. To account for the existence of neutrality in vowel harmony, several Optimality theoretic (Prince & Smolensky 1993) approaches have been proposed that utilize domain-based representations (Smolensky 1993; Bellik 2015). In these frameworks, individual features are represented as domains overlaid on one or more segments specified for that feature, and adjacent segments sharing the same feature value collapse into one feature domain. Harmony is then represented in terms of featural domains which optimally span the entire word. This intuition is in line with experimental findings that vowel disharmony (across word boundary) facilitates target word (boundary) detection for speakers of Turkish, a language exhibiting active, phonologized vowel harmony, but not for speakers of French, a language in which no such phonological processes exist (Kabak et al. 2010).

On the other hand, stem-internal disharmonic vowels, such as [i] in (1e) [jiltus] or the transparent [u] in (4b) [køzuŋki], are incorporated into the featural-domain representation via *embedding*: a featural domain of specification [α BACK] can tolerate the embedding of another domain of the opposite specification [$-\alpha$ BACK] within its span, such that the dominant domain still spans over the embedded segments without being broken up. In the remainder of this section, I provide an analysis of the Western Yugur data under a domain-based theory proposed by Bellik (2015), entitled Feature Domain Theory (FDT).

3.1 OVERVIEW OF FDT. Feature Domain Theory differs from previous domain-based theories primarily in its assumption of *headless (unheaded) feature domains*, which divorces feature domains from one single head segment. Bellik (2015) notes that the FDT representation is, in principle, isomorphic with traditional feature theory, and that for languages without vowel harmony, the size of each domain can simply be individual segments. While for languages with vowel harmony, which require adjacent vowels to have identical values for the harmonizing feature, the size of each domain is optimally that of the whole word – this is formalized as the constraint *FEATUREDOMAIN.

(6) FEATUREDOMAIN (*FD): Assign one violation for each feature domain in the output.

Bellik (2015) argues that the constraint must be highly ranked for VH languages that attempt to minimize the number of FDs by forcing agreement between adjacent vowels inside the same word. The fully faithful mappings in (7)¹ serve to demonstrate phonological representations under FDT.

(7)	<i>input</i>	<i>output</i>
a.	/(kAtAk) _F /	[(ketek) _F]
b.	/(bIdAt) _B /	[(budat) _B]
c.	/(kA) _F (tAk) _B /	[(ke) _F (tak) _B]

Disharmonic Turkish roots, provided in (8), illustrate the embedding mechanism of FDT. For instance, in (8a), the leftmost vowel which surfaces as [a] is contained in a back domain embedded within a front domain.

¹ Subscripted B and F correspond to [+back] and [-back] domains respectively.

- (8)
- | | |
|------------------------------|--|
| <i>input</i> | <i>output</i> |
| a. /dIk(kAt) _B F/ | [(dik(kat) _B) _F] |
| b. /k(A) _B brF/ | [(k(α) _B bir) _F] |

3.2 FDT ANALYSIS OF LEXICALIZED HARMONY AND NEUTRALITY. In FDT, the contrast between underspecified and underlying segments is reformulated as whether a segment is underlyingly contained within feature domains. Hence, I assume a family of faithfulness constraints notated *MAX-FD* which protects against the deletion of underlying FDs and segments contained therein. Lastly, a constraint *SPECIFY* must be undominated.

- (9) *SPECIFY*: Assign one violation for each segment not fully specified in the output.

With the three constraints introduced thus far, we can account for regular harmony in Western Yugur. As shown in Tableau 1, Spreading is driven by *SPECIFY*, and **FD* ensures that an existing domain – the root domain in this case – spreads, resulting in regular harmony spreading throughout the word.

/(sArt) _B +lAr/	SPECIFY	MAX-FD	*FD
☞ a. (sartlar) _B			*
b. (sart) _B (lir) _F			**!
c. (sart) _B lAr	*!		*

Tableau 1

The embedding mechanism is formalized in Bellik (2015) as a general **EMBED* constraint penalizing embedding and a family of **EMBED(MARKED)* constraints, drawing from Kiparsky & Pajusalu’s (2003) markedness-based typology of disharmony, which argues that transparent vowels must themselves as unmarked as possible to avoid cumulative violations of both harmony-favoring agreement constraints as well as markedness constraints. Additionally, Gafos and Dye’s (2011) perception study concludes that transparent vowels ‘do participate in vowel harmony articulatorily’ but only fail to be perceived as such, because the coarticulatory effects do not trigger significant enough perceptual difference for the vowels to be categorized differently. Under this view, transparent vowels are then optimally those which are ‘perceptually stable’² such that they withstand greater coarticulatory pressures. From these findings, Bellik notes that the **EMBED(MARKED)* constraints may also be interpreted as **EMBED(UNSTABLE)* constraints which penalize the embedding of ‘perceptually unstable’ segments. In fact, the predictions of **EMBED(MARKED)* and **EMBED(UNSTABLE)* differ in that the former predicts that, in a given language, embeddable segments should be treated as less marked than unembeddable segments across all phonological processes, I examine these predictions in section 5.

The existence of such constraints predicts an implicational hierarchy of surface vowels’ compatibility with embedding: The more perceptually stable (or less marked) a vowel, the more amenable it is to embedding within a certain language. Therefore, if a language permits the embedding of less perceptually stable (or more marked) vowels, it must also permit the embedding of more perceptually stable (or less marked) vowels. For the sake of demonstration, I will apply **EMBED* to individual vowels instead of referring to natural classes. This is also because transparent and opaque vowels in Western Yugur do not group together cleanly, since

² The idea of ‘perceptual stability’ is exemplified by a vowel like [i], which might be retracted more before it becomes perceptually distinct (Gafos & Dye 2011).

while [u] is transparent in suffixes, opaque vowels consist of both front and back vowels. The definitions of the *EMBED constraints are as follows:

(10) *EMBED: Assign one violation for each segment in the output contained in more than one feature domain.

(11) *EMBED(x): Assign one violation for each segment [x] in the output contained in more than one feature domain.

In Tableau 2 the input which triggers exceptional harmony is shown as being embedded in a feature domain of the opposite backness value. While the winning candidate a incurs violations from *FD and the *EMBED(u), avoidance of such violations inevitably results in MAX-FD violation. Therefore, the faithfulness constraint MAX-FD must be highly ranked to eliminate any change made to underlyingly specified segments.

/((k ^h Un) _B) _F +lAr/	MAX-FD	*FD	*EMBED(u)
☞ a. ((k ^h un) _B lir) _F		**	*
b. (k ^h unlar) _B	*!	*	
c. (k ^h ynlir) _F	*!	*	

Tableau 2: MAX-FD undominated

Given a highly ranked MAX-FD, neutral vowels, which I propose are specified in the UR, must either be embedded – resulting transparency – or otherwise spread its own domain, since its feature domain cannot be deleted or embedded – this is the case of opacity. However, the current constraint set favors opacity because derived embedding always incurs more violations. This means that in order to differentiate between the two types of neutrality, there needs to be another constraint driving the alignment of featural domains to edges of the word. The constraints relevant to progressive suffixal harmony of Western Yugur are ALIGN-R(–BK) and ALIGN-R(+BK).

(12) ALIGN-R(±BK): Assign one violation for each segment between the right edge of a [±back] domain and the right edge of the word.

Suffixal transparency and opacity are derived in the following tableaux. In Tableau 3, *EMBED(u) is outranked by ALIGN-R(–BK) which requires progressive spreading of the front domain of the root. Because *EMBED(u) is lower ranked, the root domain spreads over [u] at the expense of embedding it. General *EMBED must dominate ALIGN-R(+BK) to ensure that the transparent segment (which is [+back]) does not also spread rightwards.

/(kOz) _F -(In) _B -GA/	MAX-FD	ALIGN-R(–BK)	*EMBED	ALIGN-R(+BK)	*EMBED(u)
☞ a. (køz(uŋ) _B ki) _F			**	**	*
b. (køziŋki) _F	*!				
c. (køz) _F (uŋqa) _B		****!			
d. (køz(uŋqa) _B) _F			****!		

Tableau 3. Suffixal transparency

For an opaque suffixal vowel like [o], the corresponding *EMBED(o) constraint must outrank ALIGN-R(–BK) so that the root domain avoids embedding [o] at the expense of not spreading to the right word edge. As a result, [o] spreads its feature domain instead.

$/(E^{ht})_F-(O)_B-mA/$	MAX-FD	*EMBED(o)	ALIGN-R(-BK)	*EMBED	ALIGN-R(+BK)
☞ a. $(e^{ht})_F(om\alpha)_B$			***		
b. $(e^{ht}\theta mi)_F$	*!				
c. $(e^{ht}(o)_B mi)_F$		*!		*	**

Tableau 4. Suffixal opacity

In summary, regular vowel harmony takes place when a feature domain permeates through underspecified vowels. As for neutral suffixal vowels, the relative ranking of *EMBED(x) below or above ALIGN(FD) with reference to individual segment types results in that segment's being transparent or opaque respectively.

4. Alternatives. In this section, I consider two alternative approaches to the FDT account provided above. The first is the 'opaque consonant' approach, proposed by Clements & Sezer (1982) in their autosegmental analysis of disharmony in Turkish. Clements & Sezer argue that the root-final consonant is opaque in Turkish exceptional roots, such as those in (13), which trigger front harmony despite the final root vowels' being a back vowel [ɑ], and that these opaque segments are underlyingly [-back]. They then posit a feature-dissociating rule that targets the [-back] feature on syllable-final, non-sonorant segments. Application of this rule leaves the [-back] feature afloat, which is deleted on the surface.

- (13)
- | | | |
|-----------------|-----------------|----------------------|
| <i>nom. sg.</i> | <i>acc. sg.</i> | <i>gloss</i> |
| a. dikkat | dikkat-i | 'attention' |
| b. saat | saat-i | 'time; hour' |
| c. heli'ak | heli'a:ki-i | 'death; destruction' |
| d. emli'ak | emli'a:ki-i | 'estate, property' |

When the root-final segment is a sonorant, however, the [-back] feature is phonetically realized as palatalization, as shown by the data in (14).

- (14)
- | | | |
|--|---|----------------|
| <i>nom. sg.</i> | <i>acc. sg.</i> | <i>gloss</i> |
| a. petrol ^ɨ | petrol ^ɨ -y | 'petrol' |
| b. rol ^ɨ | rol ^ɨ -y | 'role' |
| c. hil ^ɨ al ^ɨ | hil ^ɨ a:l ^ɨ -i | 'crescent' |
| d. istik ^ɨ l ^ɨ al ^ɨ | istik ^ɨ l ^ɨ a:l ^ɨ -i | 'independence' |

Currently, there is insufficient data on Western Yugur for determining whether the root-final opacity approach is supported in this language. But the surface facts in (13) and (14) can also be accounted for under a domain-based account (see Bellik 2015 for detailed treatment). Here, I highlight another problem with C&S's (1982) analysis. For transparent [u]-suffixes in Western Yugur, one would need to posit underlying opaque elements on a case-by-case basis. For instance, when the root triggers front harmony, as in (15a), if an underlying [-back] is posited for the transparent suffix, then after back-vocalic roots, the feature must be deleted; otherwise the opaque segment incorrectly triggers front harmony, as in (15b).

- (15) a. køz-uŋ-ki b. *at-uŋ-ki (expected: at-uŋ-qa)
- | | | | | |
|-----|----|----|-----|-----|
| -B | +B | -B | | |
| | | | | |
| køz | +u | ŋ | +kA | *at |

The second approach I consider is Smolensky's (1993) feature domain-based representation, which derives transparency with featural markedness constraints, ALIGN, and *EMBED. Under this approach, transparency is derived from gaps in the surface vowel shaped by markedness pressures. As shown in Tableau 5, in a hypothetical language, the [+back] counterpart of [i] is absent from the surface inventory due to highly ranked *u. As a repair, an additional front domain must be epenthesized to accommodate the less marked counterpart [i] which surfaces as transparent.

/ (O) _B IO /	*u	*DEP-FD	*i	ALIGN-R(+BK)	*EMBED
a. (o(i) _F O) _B		*	*		*
b. (ouo) _B	*!				
c. (o(iø) _F O) _B		*	*		**!
d. (o) _B (iø) _F		*	*	**!	

Tableau 5. Transparency from inventory gap

By switching the positions of ALIGN-R(+BK) and *EMBED in Tableau 5, we can derive opacity as in candidate (d) instead of transparency as in candidate (a). Nevertheless, one major shortcoming of the gap approach is that it is insufficient to account for the existence of neutral vowels in symmetrical systems such as the inventory of Western Yugur, where [u] is transparent in the surface inventory, and opaque vowels include both [e] and [a] of matching height and opposite [back] values.

This flaw can be amended by conjoining featural markedness constraints (such as *u) with *EMBED and limiting the domain of the conjoined constraint evaluation to single segments. That is, violations of *u and *EMBED must overlap on one segment to constitute a violation of *u & *EMBED. However, a conjoined constraint *u & *EMBED so defined is identical to *EMBED(u) whose definition is given in (11). Hence, I treat this as an argument in favor of FDT's constraint family *EMBED(x).

5. Vowel markedness-sensitivity in Western Yugur. In this section, I turn to an additional problem regarding the particular values of the neutral vowels.

Kiparsky & Pajusalu (2003) argues that the conjoining of AGREE and *featural markedness constraints* derives transparency in vowel harmony. The two featural markedness constraints they propose are as follows.

(16) [-LO -RD] → [+BK]: Assign one violation to any vowel that is back and [-Low -Round]. Mnemonically referred to as *u, *ɣ.

(17) [+BK] → [-LO -RD]: Assign one violation to any vowel that is front and [+Low] or [+Round]. Mnemonically, *æ, *ø, *y.

The conjoined constraint *æ, *ø, *y & AGREE(BACK) is violated by a sequence like [a i æ] but respects both [a i a] and [æ i æ], forcing front vowels to be fully transparent. As its mirror image, the constraint *u, *ɣ & AGREE(BACK) derives transparent back vowels. Both conjoined constraints preclude full transparency for marked vowels referenced by either featural markedness constraint. To illustrate, the sequence [ø u i] attested in (4b) [køuŋki] is a violation of both constraints because it contains marked vowels of both [+back] and [-back] specifications – once the transparent vowel is marked, the presence of another marked, harmonizing vowel of the opposite [back] value in the sequence would result in a violation of both conjoined constraints.

The data from Western Yugur then presents a counterexample to Kipsarky & Pajusalu's (2003) typology, which argues that all markedness constraints could, in principle, be conjoined with AGREE (p.239). One way to work around this contradiction is to decompose the featural markedness constraints into constraints targeting individual featural and evaluate the relative markedness of each individual sound. This approach calls for an examination of [u] in other processes of the language. In the section below, I present an analysis of the alternating vowel /A/ in Western Yugur and suggest that the height asymmetry between its two surface variants is driven by markedness.

5.1 HARMONY ALTERNATION AS MARKEDNESS DRIVEN. The harmonic vowel /A/ has two surface realizations [ɑ] and [i] which alternate in accordance with the harmony environment (Roos 2000). Importantly, the two surface variants mismatch in height, unlike many other Turkic languages where the front counterpart of [ɑ] is [e], or Turkish with lowering of [e] to [æ] before [r] (Göksel & Kerlake 2005). In those cases, both variants share the specification [-high]; whereas for Western Yugur, the height and backness specifications for [ɑ] and [i] are complementary: [ɑ] is [-high +low], and [i] is [+high -low].

In the remainder of this section, I pursue an account of the high asymmetry as driven by markedness, such that the two surface forms are the least marked members of both [±back] classes. This requires us to turn to other known facts about the vowel inventory of Western Yugur. To begin with, Western Yugur has three mid vowels [e] [ø] and [o]. Roos (2000) notes an ongoing merger between mid vowels and their high vowel counterparts, though robustness differs – mid vowel raising is more commonly observed for the front vowels than for the back vowel.

- (18) a. [kø^hkʷus ~ ky^hkʷus] 'loin'
 b. [e^ht ~ i^ht] 'to reach'
 c. [pottuu ~ puttuu] 'became'

Assuming a diachronic view, it might seem intuitive to regard the height asymmetry in Western Yugur as the result of mid vowel raising applied to a more symmetric [ɑ]~[e] pattern. Still, under this view, one needs to explain why raising applies so uniformly to harmonizing vowels, even in contexts where raising is dispreferred. For instance, compare the singular and plural in (19), [e] tends to retain its height before sonorant segments like [r], but the plural suffix /-lir/ always surfaces as [-lir] in a front environment.

- (19) a. [ezer] 'saddle'
 b. [ezer-lir] 'saddles'

The facts surrounding mid vowel raising hints at a formal constraint *MID which disprefers mid vowels. Moreover, in Western Yugur, the occurrence of rounded vowels is restricted beyond the initial syllable (Roos 2000:14). Accordingly, I formalize this as a constraint *ROUND against round vowels. The preservation of round vowels in initial syllables could be due to a higher-ranked positional faithfulness (Beckman 1997) constraint IDENT-σ1(BACK) targeting initial syllables.

Lastly, I adopt the two featural markedness constraints from Kiparsky & Pajusalu (2003). Given an underspecified vowel input A, these four constraints will only favor the front high vowel and the back low vowel. As shown in tableaux 6 and 7 below, for the purposes of demonstration, it is assumed that backness harmony has already operated via higher-ranked constraints on the vowel /A/, which is underspecified for rounding, height and backness.

/[+Back]/	*MID	*ROUND	*u, *ɤ	*æ, *ø, *y
☞ a. a				
b. o	*!	*		
c. u		*!		
d. u			*!	

Tableau 6. Least marked back vowel

/[-Back]/	*MID	*ROUND	*u, *ɤ	*æ, *ø, *y
☞ a. i				
b. e	*!			
c. æ				*!
d. y		*!		*

Tableau 7. Least marked front vowel

The Emergence of the Unmarked analysis provides support that all four markedness constraints are an active part of Western Yugur’s phonology, which indicates that [u] is not a less marked vowel than [a]. It then confirms the challenge posed to Kiparsky & Pajusalu’s (2003) typology and contradicts the predictions of Bellik’s (2015) *EMBED(MARKED) formulation, because *EMBED(u) must rank below *EMBED(a) to derive the transparency of underlying [u] and the opacity of underlying [a]. The account given above then further favors a *EMBED(UNSTABLE) interpretation of the *EMBED(X) constraints, where markedness does not map straightforwardly to perceptual instability.

5.2 LEXICAL PRESSURE OR UNIVERSAL MARKEDNESS? Recall that in the synchronic lexicon, [u] is the most frequent vowel in the root that triggers exceptional harmony. This is attributed to a seemingly unnatural sound change which retracts Common Turkic *[i] to Western Yugur [u], coupled with the frequent retention of a root’s historical harmony class even after sound change has erased all its surface cues. A feature domain representation can shed some light on the reason why a seemingly marked vowel might be transparent – if exceptional roots necessarily involve embedding, then given the number of disharmonic roots involving [u], the most common target of embedding would naturally also be the vowel [u]. In a model of error-driving learning, whenever the learner encounters an instance of embedded [u], they would make a small adjustment to the ranking of *EMBED(u) accordingly. It seems reasonable that exposure to a large quantity of embedded [u] should result in a lower ranked *EMBED(u), which then leads to transparency when it is ranked below the ALIGN constraint.

The question remains as to how an unnatural sound change from *[i] > [u] could occur in the first place. Beguš (2022) argues that any complex, decomposable sound change must consist of natural atomic changes which are necessarily motivated in some way. Hence, the seemingly unnatural change from *[i] to [u] may have been an accumulation of otherwise motivated changes, beginning with a stage of laxing and centralization of *[i] to a more schwa-like quality, followed by repatterning with the back high vowel in certain prosodic positions. Although without further acoustic or crosslinguistic studies, this view remains speculative.

6. Conclusion. In this paper, I argued for a unified account of exceptional vowel harmony and neutral vowels within the domain-based featural representational framework of Feature Domain Theory (Bellik 2015). Via its embedding mechanism, transparency and opacity in vowel harmony can be derived from the interaction between *EMBED(X) constraints and constraints driving

spreading. I have also argued for the centrality of an embedding constraint which targets segments through comparison with alternative approaches. Furthermore, for a language with a symmetrical inventory, in order to account for neutrality without recourse to inventory gaps, the contrast between underspecified and underlying segments becomes crucial – neutral and exceptional elements are uniformly specified in the UR, while most regular vowel harmony processes take place among derived, that is, underspecified segments. The increasing limitation of regular harmonic processes from the root-morpheme domains into suffixes should be regarded as characteristic of the decay of phonological vowel harmony in the language.

References

- Beckman, Jill N. 1997. Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology* 14(1). 1–46.
- Beguš, Gašper. 2022. Distinguishing cognitive from historical influences in phonology. *Language* 98(1). 1–34. <https://doi.org/10.1353/lan.2021.0084>
- Bellik, Jennifer. 2015. Feature Domains and Covert Harmony in Turkish: The Exceptional Transparency of /a/. Santa Cruz, CA: UC Santa Cruz MA thesis.
- Boeschoten, Hendrik. 2021. Uzbek. In Lars Johanson & Éva Ágnes Csató (eds.), *The Turkic Languages* (Routledge Language Family Series), 388–408. Second edition. London ; New York: Routledge, Taylor & Francis Group.
- Clements, George N & Engin Sezer. 1982. Vowel and consonant disharmony in Turkish. *The Structure of Phonological Representations* 2. 213–255.
- Gafos, Adamantios I & Amanda Dye. 2011. Vowel harmony: Opaque and transparent vowels. *Wiley Online Library* 1–26. <https://doi.org/10.1002/9781444335262.wbctp0091>.
- Göksel, Aslı & Celia Kerslake. 2005. *Turkish: A Comprehensive Grammar*. New York: Routledge.
- Harrison, K David. 1999. Vowel harmony and disharmony in Tuvan and Tofa. *Citeseer*.
- Kabak, Barış, Kazumi Maniwa & Nina Kazanina. 2010. Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Laboratory Phonology* 1(1). <https://doi.org/10.1515/labphon.2010.010>.
- Kiparsky, Paul & Karl Pajusalu. 2003. Towards a typology of disharmony. *The Linguistic Review* 20(2–4). <https://doi.org/10.1515/flir.2003.009>.
- Lightner, Theodore M. 1972. *Russian phonology and Turkish phonology*. Linguistic Research Incorporated.
- Miao, Dongxia. 2019. *Gansu sunan xibuyuguyu [Southern Gansu Western Yugur language]*. Beijing: Shang wu yin shu guan.
- Prince, Alan, and Paul Smolensky. 1993/2002. Optimality Theory: Constraint Interaction in Generative Grammar. *Rutgers Optimality Archive*: ROA 537.
- Roos, Martina Erica. 2000. *The Western Yugur (Yellow Uygur) Language: Grammar, Texts, Vocabulary*. Leiden: Rijksuniversiteit Leiden.
- Walker, Rachel. 2012. Vowel harmony in Optimality Theory. *Language and Linguistics Compass* 6(9). 575–592. <https://doi.org/10.1002/lnc3.340>.
- Zhong, Yarjis Xueqing. 2019. *Rescuing a language from extinction: Documentation and practical steps for the revitalisation of (Western) Yugur*. Caberra: The Australian National University.