The grammaticalization of the Turkic comitative(s): Mismatches in harmony and stress

Joshua Dees
University of Illinois Urbana-Champaign

1 Introduction

In this paper, building on Dees' (2023, 2024) Cophonologies by Phase (CBP) analysis of the disharmonic Kazakh (South Kipchak Turkic) comitative/instrumental marker, I provide a CBP analysis of the Turkish (West Oghuz Turkic) comitative/instrumental marker. The two respective analyses capture observed mismatches in harmony and stress between the Kazakh and Turkish comitative/instrumental markers. Based on these analyses, and tying in Qaraqalpaq (South Kipchak Turkic) and Azerbaijani (West Oghuz Turkic), I argue that the South Kipchak comitative/instrumental marker and the West Oghuz comitative/instrumental marker have been grammaticalized (or semi-grammaticalized) from distinct proto-elements. Furthermore, I make novel observations about CBP-style enhanced vocabulary items, which seem to suggest that (supra)segmental content and prosodic subcategorizations are subject to change over time, but Cophonologies (or reweighting of constraints) remain intact.

It is well-established that palatal harmony is a highly productive process across Turkic languages. Furthermore, palatal harmony in Turkic is iterative – the [±back] feature of the root vowel(s) will spread from morpheme to morpheme. For example, the [±back] feature spreads from the root vowel(s) to the plural morpheme to the accusative morpheme in Kazakh (1) and the [±back] feature spreads from the root vowel(s) to the plural morpheme to the genitive morpheme in Turkish (2).

(1) *Iterative palatal harmony in Kazakh: root + PL + ACC (Dees 2023: 74)*:
  a. **ɔ**:q**uw**ʃ**ə**-l**a**:r-d**ə** 'student-PL-ACC'
  b. k**ʏsʏ**k-l**ie**r-d**ɛ** 'puppy-PL-ACC'

(2) *Iterative palatal harmony in Turkish: root + PL + GEN (Nevins 2010: 28):*
  a. k**i**z-l**a**r-**i** 'student-PL-GEN'
  b. **i**p-l**e**r-**i**n 'puppy-PL-GEN'

Turkic languages are also typically characterized as having word-final stress [see, e.g., Kornfilt (1997) on Turkish; Mohamedowa (2016) on Kazakh; Wurm (1951) on Qaraqalpaq]. For example, in Kazakh, the word *ba:la:* 'child' has stress on the final syllable (3a). When the plural morpheme is added, the stress shifts to the morpheme (3b), as the plural morpheme attaches to the noun and forms a prosodic word (PWd). The same is true for Turkish, as is demonstrated in (4).

(3) *Kazakh word stress (Dees 2023: 78):*

a.	b.
ba:l**á**:	ba:la:-l**á**:r
'child'	child-PL
	'children'

* Many thanks to Aigerim and Zhazira for all of the hours spent teaching me about your language. Additionally, I want to thank my advisor Dr. Aida Talić (University of Illinois Urbana-Champaign), Dr. Hannah Sande (University of California Berkeley), Dr. José Ignacio Hualde (University of Illinois Urbana-Champaign), the attendees of the 29th Mid-Continental Phonetics and Phonology Conference, the attendees of the 2024 Annual Meeting on Phonology, and the members of the University of Illinois' Phonetics and Phonology forum - The feedback from all of these people played an important role in the development of this paper.

Proceedings of *AMP 2023/2024*

(4) *Turkish word stress (Kabak & Vogel 2001: 316):*

a.	kitaplík	b.	kitaplik-lár
	'bookcase'		bookcase-PL
			'bookcases'

The PWd is standardly considered the domain for vowel harmony in Turkic languages (Kabak & Vogel 2001). This is where the Turkic comitative/instrumental marker(s) present an interesting puzzle. Specifically, there seem to be mismatches in harmony and stress between the Kazakh and Turkish comitative/instrumental markers. In Kazakh, the comitative/instrumental marker is realized as /-mien/ (5). In Turkish, the comitative can be realized in its full lexical form /ile/ (6a) or a cliticized version /-le/ (6b).

(5) *Kazakh comitative/instrumental marker (Dees 2023:75)*
a:da:m-**mien**
person-COM
'with person'

(6) *Turkish comitative/instrumental marker (Kabak & Vogel 2001: 331)*

a.	Nihan **ile**	b.	Nihan-**la**
	Nihan **COM**		Nihan-**COM**
	'with Nihan'		'with Nihan'

In Kazakh, stress shifts from the final syllable of the noun to the comitative/instrumental marker (7).[1] However, in Turkish, when the comitative/instrumental marker is in its full lexical form, the noun and the comitative/instrumental marker each bear their own word-final stress (8a). When it is a clitic, stress does not shift from the noun to the comitative/instrumental marker (8b).

(7) *Word stress and the Kazakh comitative/instrumental marker*
a:da:m-m**ié**n
person-COM
'with person'

(8) *Word stress and the Turkish comitative/instrumental marker (Kabak & Vogel 2001: 330)*

a.	Nih**á**n il**é**	b.	Nih**á**n-la
	Nihan COM		Nihan-COM
	'with Nihan'		'with Nihan'

Because the Kazakh comitative/instrumental marker forms a PWd with the root, one might assume that it would participate in palatal harmony. On the other hand, because the Turkish comitative/instrumental marker does not form a PWd with the root, it should not participate in palatal harmony. However, it turns out to be the exact opposite: The Kazakh comitative/instrumental marker is disharmonic (cf. (9a) and (9b)) and the Turkish comitative/instrumental marker (when cliticized) is harmonic (cf. (10a) and (10b)).

(9) *Disharmonic Kazakh comitative/instrumental marker (Dees 2023: 75)*

a.	bɵ:pie-m**ié**n	b.	a:da:m-m**ié**n
	baby-COM		person-COM
	'with baby'		'with person'

(10) *Harmonic Turkish comitative/instrumental marker (Kabak & Vogel 2001: 331)*

a.	çekíç=l**e**	b.	Nihán=l**a**
	hammer=INSTR		Nihan=COM
	'with hammer'		'with Nihan'

Therefore, the primary goal of this paper is to account for this apparent mismatch between the harmony and

[1] See Dees (2023) for more on this stress shift.

stress of the Kazakh comitative/instrumental marker and the Turkish comitative/instrumental marker.

The paper is structured as follows: §2 provides an overview of Cophonologies by Phase (CBP) and Dees' (2023, 2024) CBP analysis of the Kazakh comitative/instrumental marker. §3 provides a CBP analysis of the Turkish comitative/instrumental marker. In §4, I provide a diachronic proposal regarding the Turkic comitative/instrumental marker(s), providing additional data from Qaraqalpaq and Azerbaijani. Concluding remarks and directions for future research are provided in §5.

2 Cophonologies by Phase and the Kazakh comitative/instrumental marker

In this section, I provide a very brief summary of what Cophonologies by Phase (CBP) is. A more detailed summary is provided in Dees (2023), and the reader can find full descriptions of the model in Sande & Jenks (2018), Sande (2019), Sande et al. (2020). After the summary of CBP, I demonstrate the findings of Dees' (2023) CBP analysis of the disharmonic Kazakh comitative/instrumental case suffix.

2.1 *A brief overview of CBP* CBP, as introduced by Sande & Jenks (2018), is a model of the morphosyntax-phonology interface. Under this model, it is assumed that morphology and phonology are interpreted via the hierarchical output of syntax. These output forms are evaluated by a constraint-based phonological component (Sande 2019). Important assumptions of CBP include:

- The assumption of Distributed Morphology (DM). Insertion of vocabulary items takes place late in the derivation at PF.
- Syntactic phases are the domain for phonologization. Voice, C, and D constitute phase heads (i.e. Chomsky 2000, 2001, Marvin 2002), as do categorizing heads (i.e. Arad 2003, Embick 2010). Importantly all phase heads are phonologized with their complements.
- DM-style vocabulary items associate morphosyntactic features with three phonological components: An underlying phonological representation ($\mathcal{F}$), a prosodic subcategorization frame ($\mathcal{P}$), and, most importantly, a reweighting of phonological constraints ($\mathcal{R}$). This is referred to as *Enhanced Vocabulary Items*.

Some elaboration on enhanced vocabulary items and phase-based spell-out is required. The underlying representation ($\mathcal{F}$) may consist of a segmental or suprasegmental representation. Alternatively, like the other phonological components, it can be null. The prosodic subcategorization frame ($\mathcal{P}$) determines whether $\mathcal{F}$ is a free-standing prosodic word, attaches to elements, etc. It also determines the direction of attachment for elements that are bound morphemes. The reweighting[2] of constraints component ($\mathcal{R}$) is one of the key innovative components of CBP. It is an extension of work on cophonology theory (see, e.g., Orgun1996, Anttila 2002, Inkelas & Zoll 2005, 2007). Crucially, only under phases in which a vocabulary item is associated with a contentful $\mathcal{R}$ will there be a reweighting of constraints. If all vocabulary items within a phase have null $\mathcal{R}$s, then the default weighting of constraints will be applied.

As in Sande (2019), enhanced vocabulary items will be expressed as extended DM-style vocabulary items (as in (11)).

(11) *CBP-style enhanced vocabulary item (Sande 2019: 462)*:

$$\text{[Syntactic features]} \leftrightarrow \left\{\begin{array}{l}\mathcal{F}\text{: (Supra)segmental content}\\ \mathcal{P}\text{: Prosodic subcategorization}\\ \mathcal{R}\text{: Constraint reweighting: B(+3)}\end{array}\right\}$$

As was previously mentioned, any of the three components in (11) may be null. In (11), since $\mathcal{R}$ is associated with a constraint reweighting, it interacts with the default weighting of constraints for the language. However, phonology does not take place at vocabulary insertion. Instead, it takes place at the phase boundary. Thus, the $\mathcal{R}$s of all vocabulary items within a phase will interact simultaneously with the default weighting. Put more simply, reweightings take scope over spell-out domains (in the case of Sande 2019, this is the entire

[2] Note that Sande & Jenks (2018) utilize constraint rerankings. Either approach works for this paper. However, to maintain consistency with the most recent literature on CBP, reweightings will be used in the analysis. The constraint weights are arbitrarily set to account of the constraint ranking in the grammar.

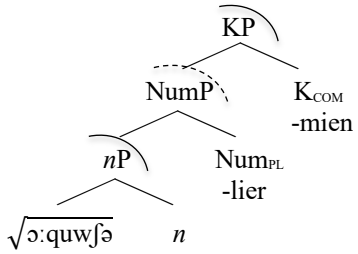
phase). If no other vocabulary items within the same phase as the vocabulary item in (11) are associated with a contentful $\mathcal{R}$, that entire phase will be spelled out based on the constraint reweighting in (11). For example, let us assume the default weighting of concerned constraints for a given language is A = 2, B = 1, and C = 3. The constraint weighting for the phase in which (13) is spelled out will then become A = 2, B = 4, C = 3. Crucially, the reweightings of constraints cannot affect phonological evaluation of material outside the domain (Sande 2019).

2.2 *Modelling the disharmonic Kazakh comitative* As previously noted, Dees (2023, 2024) demonstrates that the disharmonic Kazakh comitative/instrumental marker can be modelled via CBP, but only under the assumption that the K head can be phonologized separately from its complement. A review of this analysis is provided below.

First, Dees (2023) argues that the Kazakh comitative is one of six grammatical cases. Thus, Dees proposes the structure in (13) for a phrase like (12).

(12) ɔ:quwʃə-la:r-mien
student-PL-COM
'with students'

(13)



Case markers have been argued to project KP in the nominal phrase (see Takahashi 2011 for Japanese, Bošković 2014, and Franks 2020 for Slavic). Therefore, Dees (2023) assumes Kazakh case markers (including the comitative) are realizations of a K head. Additionally, *n*P is assumed to be a phase (cf. Arad 2003; Embick 2010) and KP, as the highest phrase of the extended projection of the lexical category N, is assumed to be a phase (cf. Bošković 2014). Dees (2023) also posits the K head is phonologized separately from it's it complement.[3]

To account for palatal harmony in the Kazakh grammar, Dees (2023) uses two constraints: an agreement by projection constraint (BACKHARM), and a faithfulness constraint (ID-BACK). To unify accounts of Kazakh and Turkish, I have updated the set of constraints as follows:

(14) *Kazakh palatal harmony constraints*
a. ALIGN-RIGHT([back], PWd): For any feature [back] associated to a segment in a PWd, that feature has an association to the leftmost syllable of a word [adapted from Walker (2012)]
b. Ident-IO(back): Assign one violation if an output segment differs in back value from the corresponding input segment

Due to Kazakh palatal harmony being highly productive, the default weight for ALIGN-R should be greater than ID-BACK (as in (15)).

(15) *Default constraint weighting for palatal harmony in Kazakh (Dees 2023: 82)*

Constraint	Weight
ALIGN-R	3
ID-BACK	1

[3] Dees (2023) suggests that K is phonologized separately from it complement based on (Newell 2010).

Additionally, Dees (2023) posits the following enhanced vocabulary items:

(16) $[n] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & \text{ø} \\ \mathcal{P}_1: & [_\omega\ \text{X}] \\ \mathcal{R}_1: & \text{ø} \end{Bmatrix}$

(17) $[\sqrt{\text{ student }}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & \text{/ɔːquwʃə/} \\ \mathcal{P}_2: & \text{ø} \\ \mathcal{R}_2: & \text{ø} \end{Bmatrix}$

(18) $[\text{Num}_{\text{PL}}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & \text{/lier/} \\ \mathcal{P}_3: & \text{-X]}_\omega \\ \mathcal{R}_3: & \text{ø} \end{Bmatrix}$

(19) $[\text{K}_{\text{ACC}}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & \text{/dɛ/} \\ \mathcal{P}_4: & \text{-X]}_\omega \\ \mathcal{R}_4: & \text{ø} \end{Bmatrix}$

(20) $[\text{K}_{\text{COM}}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & \text{/mien/} \\ \mathcal{P}_5: & \text{-X]}_\omega \\ \mathcal{R}_5: & \text{BACKHARM}^{-2},\ \text{ID-BACK}^{+2} \end{Bmatrix}$

The $\mathcal{P}$s in both (19) and (20) specify that the phonological content of K_{ACC} and K_{COM} is a suffix that forms a prosodic word together with the content it attaches to. Furthermore, Dees (2023) posits that K_{ACC} is not associated with a reweighting of constraints ($\mathcal{R}$), whereas K_{COM} is associated with a reweighting of constraints ($\mathcal{R}$).

Assuming the phase head K is phonologized separately from its complement, Dees (2023) demonstrates that there are three spell-out domains involved in the derivation of the Kazakh KP. Vocabulary insertion occurs and then, at each spell-out domain, phonological composition takes place from the bottom up. Thus, the following phonological compositions take place for KP_{ACC}:

(21) a. *Noun after vocabulary insertion composition*
/ɔːquwʃə/$^{(\text{P}^{1+2}\text{R}^{1+2})}$ → [$_\omega$ ɔː.quw.ʃə]
b. *Noun + Num after vocabulary insertion composition*
/[$_\omega$ ɔː.quw.ʃə]-lier/$^{(\text{P}^3\text{R}^3)}$ → [$_\omega$ ɔː.quw.ʃə.laːr]
c. *KP_{ACC} after vocabulary insertion composition*
/[$_\omega$ ɔː.quw.ʃə.lar]-dɛ/$^{(\text{P}^4\text{R}^4)}$→ [$_\omega$ ɔː.quw.ʃə.laːr.də]

The tableau in (22) corresponds to the phonological evaluation in (21a). Neither of the vocabulary items is associated with a reweighting of constraints, so the default constraint weighting is applied. A PWd is formed, therefore all of the vowel segments must correspond with the [back] feature of the leftmost vowel.

(22)

/$_\omega$ ɔːquwʃə /	**ALIGN-R([back], PWd) (3)**	**ID-BACK (1)**	**H**
a.☞[$_\omega$ ɔː.quw.ʃə]			
b. [$_\omega$ ɔː.qie.ʃɛ]	6	2	**8**

The tableau in (23) corresponds to the phonological evaluation in (21b). Again, the vocabulary item is not associated with a reweighting of constraints, so the default constraint weighting is applied. The plural morpheme affixes to the noun forming a PWd. Therefore, all of the vowel segments must correspond with the [+back] feature of the noun's vowels (as per the ALIGN-R constraint).

(23)

/ [ω ɔː.quw.ʃə]-lier[4] /	**ALIGN-R([back], PWd) (3)**	**ID-BACK (1)**	**H**
a. [ω ɔː.quw.ʃə.lier]	3		**3**
b.☞[ω ɔː.quw.ʃə.laːr]		1	**1**

Finally, the tableau in (24) corresponds to the phonological evaluation in (21c). Again, the vocabulary item is not associated with a reweighting of constraints, so the default constraint weighting is applied. The accusative morpheme affixes to the plural noun forming a PWd. Therefore, all of the vowel segments must correspond with the [+back] feature of the noun's vowels (as per the ALIGN-R constraint).

(24)

/ [ω ɔː.quw.ʃə.laːr]-dɛ /	**ALIGN-R([back], PWd) (3)**	**ID-BACK (1)**	**H**
a. [ω ɔː.quw.ʃə.laːr.dɛ]	3		**3**
b.☞[ω ɔː.quw.ʃə.laːr.də]		1	**1**

The phonological compositions from (21a) and (21b) take place for KP_{COM} as well. However, the final vocabulary insertion composition is as follows:

(25) *KP_{ACC} afer vocabulary insertion composition*
[ω ɔː.quw.ʃə.laːr]-mien/$^{(P^5R^5)}$ → [ω ɔː.quw.ʃə.laːr.mien]

The tableau in (26) corresponds to the phonological evaluation in (25). This time, a reweighting of constraints takes place, weighting ID-BACK (weight = 3) more than ALIGN-R (weight = 1). The comitative/instrumental morpheme affixes to the plural noun forming a PWd. However, because ID-BACK is weighted more than ALIGN-R, the vowel segment of the output must match the input. Therefore, palatal harmony does not take place.

(26)

/ [ω ɔː.quw.ʃə.laːr]-mien /	**ALIGN-R([back], PWd) (1)**	**ID-BACK (3)**	**H**
a.☞[ω ɔː.quw.ʃə.laːr.mien]	1		**1**
b. [ω ɔː.quw.ʃə.laːr.mən]		3	**3**

As Dees (2023) argues, this CBP-style analysis of K in Kazakh accounts for the stress and palatal harmony patterns of the Kazakh comitative/instrumental marker in relation to stress and palatal harmony patterns of other Kazakh case markers.

3 A CBP analysis of the Turkish comitative/instrumental marker

In this section, I discuss the morphosyntactic structure (S-structure) and phonological representation (P-structure) of the Turkish comitative/instrumental case. Additionally, I provide a CBP analysis of the Turkish comitative/instrumental marker, capturing the stress and harmony patterns described in §1.

3.1 *The structure of the Turkish comitative/instrumental marker* In terms of the S-structure of the

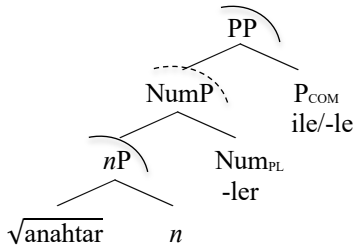
[4] As in Walker (2012), for demonstration purposes, the suffix vowels are [-back] at the input. If the suffix vowels were [+back] in the input, the same results would be achieved.

comitative/instrumental phrases in Turkish, I assume that the Turkish comitative/instrumental case is a lexical feature of the postposition (*ile*) or the postpositional clitic (*-le*) (i.e., Kabak 2006; Kabak & Vogel 2001), rather than grammatical case (i.e., Dees (2023) on Kazakh). Thus, the phrases in (27) are captured by the structure in (28).

(27) *Turkish root + PL + COM/INSTR*

a.	anahtar-lar **ile**	b.	anahtar-lar=**la**
	key-PL **INSTR**		key-PL=**INSTR**
	'with keys'		'with keys'

(28)



Again, I assume that *n*P is a phase. Since PP is the highest phrase of the extended projection of the lexical category P, it is also assumed to be a phase (i.e., Bošković 2014). Additionally, like K in Kazakh, I assume P is phonologized separately from it complement.[5]

Consider now the P-structure of the Turkish comitative/instrumental phrases. As Selkirk (1996) notes, a sequence of lexical words, [Lex Lex], is prosodized as a sequence of PWd in P-structure (29).

(29) ((*lex*)$_{PWs}$ (*lex*)$_{PWd}$)$_{PPh}$

However, since the Turkish comitative/instrumental postposition is a functional word, we must consider the possible P-structure(s) of phrases with functional words ([Lex Fnc]). According to Selkirk (1996), there are four different organizations into a prosodic word that are available for a [Lex Fnc] sequence.

(30) *Possible P-structure of [Lex Fnc] phrases (adapted from Selkirk 1996)*

a.	((*lex*)$_{PWd}$ (*fnc*)$_{PWd}$)$_{PPh}$	**Prosodic word**
b.	((*lex*)$_{PWd}$ *fnc*)$_{PPh}$	**Free clitic**
c.	((*lex fnc*)$_{PWd}$)$_{PPh}$	**Internal clitic**
d.	(((*lex*)$_{PWd}$ *fnc*)$_{PWd}$)$_{PPh}$	**Affixal clitic**

As previously demonstrated in (8a), both the noun and the free-standing comitative postposition have lexical (word-final) stress in Turkish. Thus, I categorize each as a prosodic word (30a). The fact that stress does not shift to the comitative/instrumental postpositional clitic in Turkish (8b) suggests that it is not grammatical case (i.e., Kazakh), nor is it an internal clitic (30c). Furthermore, since the clitic variant of the comitative/instrumental postposition partakes in palatal harmony (see (10)), and the PWd is the domain for harmony in Turkish, I classify it as an affixal clitic (30d). The comitative/instrumental postposition (Fnc) being located in a nested PWd structure, both sister to PWd and dominated by PWd (30d), explains why it doesn't bear stress, but partakes in palatal harmony (cf. Kabak & Vogel 2001, *Phonological Word Adjoiners*).

3.2 *A CBP analysis of the Turkish comitative/instrumental marker* Having established the S-structure and P-structure of the Turkish comitative/instrumental marker, a CBP-style analysis can be applied. The same constraints (14) and constraint weights (15) utilized in the analysis of the Kazakh comitative/instrumental marker can be applied. I posit the following enhanced vocabulary items for the

[5] Assuming P is phonologized separately from its complement is not absolutely necessary for Turkish (the results would be the same if it were phonologized with its complement). However, in §4, I will demonstrate that the Qaraqalpaq comitative is disharmonic, but seems to be a postposition. In the case of Qaraqalpaq, P would need to be phonologized separately from its complements, just like K in Kazakh.

comitative/instrumental postposition (31) and postpositional clitic (32):

(31) $[\mathrm{P_{COM}}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & /\text{ile}/ \\ \mathcal{P}: & [_{\omega}\,\mathrm{X}] \\ \mathcal{R}: & \text{ø} \end{Bmatrix}$

(32) $[\mathrm{P_{COM}}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & /\text{le}/ \\ \mathcal{P}_4: &]_{\omega}\,{=}\mathrm{X}]_{\omega}] \\ \mathcal{R}_4: & \text{ø} \end{Bmatrix}$

Since the free-standing comitative/instrumental postposition forms its own PWd, it is expected to not partake in palatal harmony with the noun. Therefore, I set it aside for the time being. Focusing on the clitic variant, the $\mathcal{P}$ in (32) specifies that the phonological content of $\mathrm{P_{COM}}$ is an affixal clitic which is located in a nested prosodic word structure. Therefore, it forms a PWd with its host in a different manner than the Kazakh comitative/instrumental marker, which explains the differences in stress patterns. Crucially, unlike $\mathrm{K_{COM}}$ in Kazakh, $\mathrm{P_{COM}}$ in Turkish is not associated with a reweighting of constraints. Additionally, for an example like (27b), the rest of the enhanced vocabulary items involved in the derivation include:

(33) $[\sqrt{\text{ key }}] \leftrightarrow \begin{Bmatrix} \mathcal{F}_1: & /\text{anahtar}/ \\ \mathcal{P}_1: & \text{ø} \\ \mathcal{R}_1: & \text{ø} \end{Bmatrix}$

(34) $[n] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & \text{ø} \\ \mathcal{P}_2: & [_{\omega}\,\mathrm{X}] \\ \mathcal{R}_2: & \text{ø} \end{Bmatrix}$

(35) $[\mathrm{Num_{PL}}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & /\text{ler}/ \\ \mathcal{P}_3: & \text{-X}]_{\omega} \\ \mathcal{R}_3: & \text{ø} \end{Bmatrix}$

Vocabulary insertion then takes place and in each spell-out domain the content at each terminal node undergoes phonological composition from the bottom up. As in Kazakh, three phonological compositions take place:

(36) a. *Noun after vocabulary insertion composition*
/anahtar/$^{(\mathrm{P^{1+2}R^{1+2}})}$ → [$_{\omega}$ an.ah.tar]
b. *Noun + Num after vocabulary insertion composition*
/[$_{\omega}$ a.nah.tar]-ler/$^{(\mathrm{P^3R^3})}$ → [$_{\omega}$ an.ah.tar.lar]
c. *$\mathrm{PP_{COM}}$ after vocabulary insertion composition*
/[$_{\omega}$ a.nah.tar.lar]-le/$^{(\mathrm{P^4R^4})}$ → [$_{\omega}$ [$_{\omega}$ an.ah.tar.lar].la]

The tableau in (37) corresponds to the phonological evaluation in (36a). Neither of the vocabulary items is associated with a reweighting of constraints, so the default constraint weighting is applied. A PWd is formed, therefore all of the vowel segments must correspond with the [back] feature of the leftmost vowel.

(37)

/ω anahtar /	**ALIGN-R([back], PWd) (3)**	**ID-BACK (1)**	**H**
a.☞[ω a.nah.tar]			
b. [ω a.neh.ter]	6	2	**8**

The tableau in (38) corresponds to the phonological evaluation in (36b). Again, the vocabulary item (35) is not associated with a reweighting of constraints, so the default constraint weighting is applied. The plural morpheme affixes to the noun forming a PWd. Therefore, all of the vowel segments must correspond with the [+back] feature of the noun's vowels (as per the ALIGN-R constraint).

(38)

/[ω a.nah.tar]-ler /	**ALIGN-R([back], PWd) (3)**	**ID-BACK (1)**	**H**
a. [ω an.ah.tar.ler]	3		**3**
b.☞[ω an.ah.tar.lar]		1	**1**

Finally, the tableau in (39) corresponds to the phonological evaluation in (36c). Unlike the vocabulary item for K_{COM} in Kazakh, the vocabulary item for the Turkish P_{COM} (32) is not associated with a reweighting of constraints, so the default constraint weighting is applied. The comitative/instrumental postposition attaches to its host (plural noun) forming a nested PWd structure. Therefore, all of the vowel segments must correspond with the [+back] feature of the plural noun's vowels (as per the ALIGN-R constraint).

(39)

/[ω a.nah.tar]-ler /	**ALIGN-R([back], PWd) (3)**	**ID-BACK (1)**	**H**
a. [ω [ω an.ah.tar.lar].le]	3		**3**
b.☞[ω [ω an.ah.tar.lar].la]		1	**1**

This CBP-style analysis captures the stress and palatal harmony patterns of the Turkish comitative/instrumental clitic, demonstrating how it differs from the Kazakh comitative/instrumental case. Next, I unify these analyses from a diachronic point-of-view, tying in data from Qaraqalpaq and Azerbaijani.

4 An analysis of mismatches in stress and harmony

Drawing from the CBP analyses I have provided in the previous sections, I propose a unified (diachronic) account of the mismatches in harmony and stress between the comitative/instrumental markers in South Kipchak Turkic languages and West Oghuz Turkic languages.

4.1 *The case cycle: Diachronic assumptions* As van Gelderen (2023) notes, there are two types of cyclical (diachronic) change: micro cycles and macro cycles. These cycles involve grammaticalization, which involves the following types of change (van Gelderen 2023: 6):

- Content item > grammatical word > clitic > inflectional affix > zero (Hopper & Traugott 2003: 7; Narrog & Heine 2021: 130)
- Free > periphrastic > affixal > fusional (Dahl 2004: 106)

Case cycles fall under the category of macro cycles. According to van Gelderen (2023), macro cycles involve grammaticalization and renewal in ways that affect the entire language and language typologies. Semantic case (e.g., comitative/instrumental case) is typically associated with the following cycle: adverb/noun/verb > preposition > case > zero.

Examining Turkic at large, we see evidence of the preposition > case stage for the comitative/instrumental case. More specifically, I argue that there is evidence of postposition > postpositional clitic > grammatical (affixal) case changes. This has already been illustrated by Dees' (2023) analysis of the Kazakh comitative/instrumental case as being a grammatical (affixal) case and my current analysis of the Turkish comitative/instrumental case being a free-standing postposition or an affixal clitic.

4.2 *Qaraqalpaq/Azerbaijani comitative/instrumental case* Like the Turkish comitative/instrumental case, both the Qaraqalpaq and Azerbaijani comitative/instrumental cases have a free-standing form and a

clitic variant. Menges (1947), for example, references a full form, *menen* (40a), and an 'abbreviated' form *men* (40b).[6] Menges also notes that, in literature, it is often written separately from the noun. However, Wilhelm Radloff, in Die alttükischen Inschriten der Mongolei, customarily conjoins nouns with the clitic variant.

(40) *Qaraqalpaq comitative/instrumental case (adapted from Menges 1947: 54)*

a.	temirdžoł	menen	b.	temirdžoł=men(en)
	railroad	INSTR		railroad=INSTR
	'by/with the railroad'			'by/with the railroad'

Likewise, Azerbaijani has the same comitative/instrumental case marking as Turkish (41).

(41) *Azerbaijani comitative/instrumental case*

a.	çekiç	ile	b.	çekiç=le
	hammer	INSTR		hammer=INSTR
	'with hammer'			'with hammer'

Based on this data, I assume that the comitative/instrumental cases in Qaraqalpaq and Azerbaijani are lexical features of a postposition or postpositional clitic. Furthermore, I suggest that the postpositional clitic in both languages is an affixal clitic.[7] The data departs, however, in terms of palatal harmony. Like the Kazakh comitative/instrumental marker, the Qaraqalpaq comitative/instrumental marker is disharmonic. Menges (1947) describes the comitative/instrumental marker in Qaralqalpaq as disharmonic, even when it is attached to the noun (42).

(42) *Disharmonic Qaraqalpaq comitative/instrumental marker (Menges 1947: 54)*

a.	temirdžoł=m**e**n(**e**n)	b.	paraxot=p**e**n(**e**n)
	railroad=INSTR		steamer=INSTR
	'by/with the railroad'		'with the steamer'

In Azerbaijani, just like Turkish, the comitative/instrumental marker does harmonize when it is an affixal clitic (43).

(43) *Harmonic Azerbaijani comitative/instrumental marker*

a.	çekiç=l**e**	b.	Nihan=l**a**
	hammer=INSTR		Nihan=COM
	'with hammer'		'with Nihan'

Based on these patterns, I propose (44) and (45) as CBP-style enhanced vocabulary items for Qaraqalpaq. The Azerbaijani enhanced vocabulary items will be identical to the ones proposed for Turkish [(31) for the free-standing postposition and (32) for the clitic].

(44) *Qaraqalpaq free-standing comitative/instrumental postposition*

$$[\mathrm{P}_{\mathrm{COM}}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & /\mathrm{menen}/ \\ \mathcal{P}: & [_{\omega}\,\mathrm{X}] \\ \mathcal{R}: & ? \end{Bmatrix}$$

[6] Hashan (Merlin) Balihaxi (p.c.) noted that Kazakh spoken in China also has a free-standing comitative/instrumental marker *menen* in addition to the affixal *mien*.

[7] Data from native speakers is required to analyze the stress patterns, supporting this assumption.

(45) *Qaraqalpaq comitative/instrumental postpositional clitic*

$$[\mathrm{P_{COM}}] \leftrightarrow \begin{Bmatrix} \mathcal{F}: & /\mathrm{men(en)}/ \\ \mathcal{P}: &]_{\omega}{=}\mathrm{X}]_{\omega} \\ \mathcal{R}: & \textsc{backharm}^{-2},\ \textsc{id-back}^{+2} \end{Bmatrix}$$

The reweighting of constraints in (45), would account for the disharmonic nature of the Qaraqalpaq comitative/instrumental marker. I leave the $\mathcal{R}$ in (44) open for the time being, as the free-standing postposition should not harmonize with the noun regardless of the $\mathcal{R}$, as it forms its own PWd.

There are some striking resemblances between the Kazakh and Qaraqalpaq comitative/instrumental vocabulary items as well as the Turkish and Azerbaijani comitative/instrumental vocabulary items. More specifically, both the Kazakh and Qaraqalpaq vocabulary items are associated with a reweighting of constraints. Furthermore, the featural content of the Qaraqalpaq comitative/instrumental pospositional clitic (/men(en)/) is similar to the featural content of the Kazakh comitative/instrumental grammatical case (/mien/). The Turkish and Azerbaijani vocabulary items, as previously noted, are actually identical. Thus, I propose that the Kazakh-Qaraqalpaq (South Kipchak) comitative/instrumental case and the Turkish-Azerbaijani (West Oghuz) comitative/instrumental case are derived from distinct proto vocabulary items.

Zooming into the Kazakh-Qaraqalpaq (South Kipchak) comitative/instrumental markers, this proposal has some key implications for CBP and diachrony. First, assuming the case cycle, the structural differences between the Kazakh comitative/instrumental case ($\mathrm{KP_{COM}}$) and the Qaraqalpaq comitative/instrumental case ($\mathrm{PP_{COM}}$) make perfect sense. Comparing Kazakh and Qaraqalpaq, it appears that the comitative/instrumental marker originates as a PWd, then it becomes an affixal clitic, and then (in the case of Kazakh) it becomes grammatical case. Throughout these changes, it is clear that the featural content of a CBP-style enhanced vocabulary item can change (/menen/ → /men/ in Qaraqalpaq). Additionally, the prosodic subcategorization changes: independent PWd ($[_{\omega}\mathrm{X}]$) > Affixal clitic ($]_{\omega}{=}\mathrm{X}]_{\omega}$) > Affix ($\text{-}\mathrm{X}]_{\omega}$). Turning to the reweighting of constraints ($\mathcal{R}$), in Qaraqalpaq, the postpositional clitic is associated with a reweighting of constraints. Likewise, in Kazakh the case affix is associated with a reweighting of constraints. Thus, I propose that the Qaraqalpaq postposition (44) is also associated with a reweighting of constraints, and that the proto vocabulary item for the South Kipchak comitative/instrumental case (adverb/noun/verb) was also associated with a reweighting of contraints. On the otherhand, the proto vocabulary item for the West Oghuz comitative/instrumental case would not have been associated with a reweighting of constraints. Therefore, I suggest that as vocabulary items go through cyclical change (i.e., the case cycle), featural content ($\mathcal{F}$) and prosodic subcategorizations ($\mathcal{P}$) are subject to change, but reweightings of constraints ($\mathcal{R}$) remain intact.

5 Conclusions

Comparing Dees' (2023, 2024) CBP analysis of the disharmonic comitative/instrumental marker in Kazakh with a novel CBP analysis of the harmonic comitative/instrumental marker in Turkish, I have captured the mismatches in harmony and stress outlined in the introduction. Furthermore, in comparing CBP-style enhanced vocabulary items for the comitative/instrumental markers in Kazakh, Qaraqalpaq, Turkish, and Azerbaijani, I have proposed that the South Kipchak (Kazakh/Qaraqalpaq) comitative/instrumental case and the West Oghuz (Turkish/Azerbaijani) comitative/instrumental case were grammaticalized from distinct proto-elements. To strengthen this final argument, synchronic data from Qaraqalpaq and Azerbaijani is necessary, as is further diachronic exploration of all four languages.

6 References

Anttila, Arto. 2002. Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20. 1–42. https://doi.org/10.1023/A:1014245408622.

Arad, Maya. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. Natural Language and Linguistic Theory 21. 737–778. https://doi.org/10.1023/A:1025533719905.

Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45(1). 27–89. https://doi.org/10.1162/LING_a_00148.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. 89–155. Cambridge, MA: MIT

Press.
Chomsky, Noam. 2001. Derivation by Phase. In Michael Kenstowicz, Michael (ed.), *Ken Hale: A Life in Language*. 1–52. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/4056.003.0004.
Dees, Joshua. 2023. Non-uniformity in phonologizing phase heads: Evidence from Kazakh. *Proceedings of the Workshop on Turkic and Languages in Contact with Turkic* 8. 74–88. https://doi.org/10.3765/6wbxsg62.
Dees, Joshua. 2024. Phase head exception in phonological spell-out: A deeper look at Kazakh and Turkic. *Proceedings of The Chicago Linguistic Society (CLS)* 59. 75–85.
Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.71?locatt=mode:legacy.
Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/9780262014229.001.0001.
Franks, Steven. 2020. *Microvariation in the South Slavic noun phrase*. Slavica Publishers.
Hopper, Mike & Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge University Press. https://doi.org/10.1017/CBO9781139165525.
Inkelas, Sharon & Cheryl Zoll. 2005. *Reduplication: Doubling in Morphology*. Cambridge University Press. https://doi.org/10.1017/S0952675708001590.
Inkelas, Sharon & Cheryl Zoll. 2007. *Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology* 45(1). 133–171. https://doi.org/10.1515/LING.2007.004.
Kabak, Bariş. 2006. An obstacle to the morphologization of postpositions. *Studies in Language* 30(1). 33–68. https://doi.org/10.1075/sl.30.1.03kab.
Kabak, Bariş. & Irene Vogel. 2001. The phonological word and stress assignment in Turkish. *Phonology* 18. 315–360. https://doi.org/10.1017/S0952675701004201.
Kornfilt, Jaklin. 1997. Turkish. Routledge. https://doi.org/10.4324/9781315823652.
Marvin, Tatjana. 2002. *Topics in the stress and syntax of* words. Cambridge, MA: Massachusetts Institute of Technology dissertation. http://hdl.handle.net/1721.1/7988.
McCollum, Adam G. 2018. The empirical consequences of data collection methods: A case study from Kazakh vowel harmony. Linguistic Discovery. 16(2), 72–110. https://doi.org/10.1349/PS1.1537-0852.A.495.
Menges, Karl H. *Qaraqalpaq Grammar: Part One: Phonology.* New York: King's Crown Press.
Muhamedowa, Raihan. 2015. *Kazakh: A comprehensive grammar*. Routledge. https://doi.org/10.4324/9781315738239.
Narrog, Heiko & Bernd Heine 2021. Grammaticalization. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199586783.001.0001.
Nevins, Andrew. 2010. *Locality in vowel harmony*. (Linguistic Inquiry Monograph 55). Cambridge, Mass: MIT Press. https://doi.org/10.7551/mitpress/9780262140973.001.0001.
Newell, Heather. 2010. Aspects of the morphology and phonology of phases. Montreal, Quebec: McGill University dissertation. https://escholarship.mcgill.ca/concern/theses/1c18dh59v.
Orgun, Cemil Orhan. 1996. *Sign-based morphology and phonology with special attention to Optimality Theory*. Berkeley, California: UC Berkeley dissertation. https://escholarship.org/uc/item/1jc5z6q9.
Sande, Hannah. 2019. A Unified Account of Conditioned Phonological Alternations: Evidence from Guébie. Language 95(3). 456–498. https://doi.org/10.1353/lan.2019.0053.
Sande, Hannah. & Peter Jenks. 2018. Cophonologies by phase. In Sherry Hucklebridge and Max Nelson (eds.) *North East Linguistic Society (NELS)* 48(2). 39–52.
Sande, Hannah, Peter Jenks & Sharon Inkelas. 2020. Cophonologies by ph(r)ase. *Natural Language and Linguistic Theory* 38. 1211–1261. https://doi.org/10.1007/s11049-020-09467-x.
Selkirk, Elisabeth. 1996. The prosodic structure of function words. In James Morgan & Katherine Demuth (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. 187–213. Lawrence Erlbaum Associates, Inc. https://doi.org/10.1002/9780470756171.
Takahashi, Masahiko. 2011. *Some consequences of Case-marking in* Japanese. Storrs, CT: University of Connecticut dissertation. https://www.proquest.com/docview/908959961.
van Gelderen, Elly. 2023. *The Linguistic Cycle*. New York, New York: Routledge. https://doi.org/10.4324/9781003272564.
Walker, Rachel. 2012. Vowel Harmony in Optimality Theory. *Language and Linguistics Compass* 6. 575 – 592. https://doi.org/10.1002/lnc3.340.
Wurm, Stefan. 1951. The Karakalpak Language. *Anthropos*, *46*(3/4), 487–610. http://www.jstor.org/stable/40449486.